

EVIDENCE-BASED HEALTH INFORMATICS

Studies in Health Technology and Informatics

This book series was started in 1990 to promote research conducted under the auspices of the EC programmes' Advanced Informatics in Medicine (AIM) and Biomedical and Health Research (BHR) bioengineering branch. A driving aspect of international health informatics is that telecommunication technology, rehabilitative technology, intelligent home technology and many other components are moving together and form one integrated world of information and communication media. The series has been accepted by MEDLINE/PubMed, SciVerse Scopus, EMCare, Book Citation Index – Science and Thomson Reuters' Conference Proceedings Citation Index.

Series Editors:

Dr. O. Bodenreider, Dr. J.P. Christensen, Prof. G. de Moor, Prof. A. Famili, Dr. U. Fors, Prof. A. Hasman, Prof. E.J.S. Hovenga, Prof. L. Hunter, Dr. I. Iakovidis, Dr. Z. Kolitsi, Mr. O. Le Dour, Dr. A. Lymberis, Prof. J. Mantas, Prof. M.A. Musen, Prof. P.F. Niederer, Prof. A. Pedotti, Prof. O. Rienhoff, Prof. F.H. Roger France, Dr. N. Rossing, Prof. N. Saranummi, Dr. E.R. Siegel, Prof. T. Solomonides and Dr. P. Wilson

Volume 222

Recently published in this series

- Vol. 221. J. Hofdijk, B. Séroussi, C. Lovis, F. Ehrler, F. Sieverink, A. Ugon and M. Hercigonja-Szekeres (Eds.), *Transforming Healthcare with the Internet of Things – Proceedings of the EFMI Special Topic Conference 2016*
- Vol. 220. J.D. Westwood, S.W. Westwood, L. Felländer-Tsai, C.M. Fidopiastis, A. Liu, S. Senger and K.G. Vosburgh (Eds.), *Medicine Meets Virtual Reality 22 – NextMed/MMVR22*
- Vol. 219. B.K. Wiederhold, G. Riva and M.D. Wiederhold (Eds.), *Annual Review of Cybertherapy and Telemedicine 2015 – Virtual Reality in Healthcare: Medical Simulation and Experiential Interface*
- Vol. 218. E.M. Borycki, A.W. Kushniruk, C.E. Kuziemsky and C.Nøhr (Eds.), *Context Sensitive Health Informatics: Many Places, Many Users, Many Contexts, Many Uses*
- Vol. 217. C. Sík-Lányi, E.-J. Hoogerwerf, K. Miesenberger and P. Cudd (Eds.), *Assistive Technology – Building Bridges*
- Vol. 216. I.N. Sarkar, A. Georgiou and P. Mazzoncini de Azevedo Marques (Eds.), *MEDINFO 2015: eHealth-enabled Health – Proceedings of the 15th World Congress on Health and Biomedical Informatics*
- Vol. 215. L. Botin, P. Bertelsen and C. Nøhr (Eds.), *Techno-Anthropology in Health Informatics – Methodologies for Improving Human-Technology Relations*

ISSN 0926-9630 (print)
ISSN 1879-8365 (online)

Evidence-Based Health Informatics

Promoting Safety and Efficiency Through Scientific Methods
and Ethical Policy

Edited by

Elske Ammenwerth

*Professor of Medical Informatics, Institute of Biomedical Informatics, UMIT –
University for Health Sciences, Medical Informatics and Technology,
Hall in Tirol, Austria*

and

Michael Rigby

*Emeritus Professor of Health Information Strategy, School of Social Science
and Public Policy, Keele University, UK*

IOS
Press

Amsterdam • Berlin • Washington, DC

© 2016 The authors and IOS Press.

This book is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License.

ISBN 978-1-61499-634-7 (print)

ISBN 978-1-61499-635-4 (online)

Library of Congress Control Number: 2016939774

Publisher

IOS Press BV

Nieuwe Hemweg 6B

1013 BG Amsterdam

Netherlands

fax: +31 20 687 0019

e-mail: order@iospress.nl

Distributor in the USA and Canada

IOS Press, Inc.

4502 Rachael Manor Drive

Fairfax, VA 22032

USA

fax: +1 703 323 3668

e-mail: iosbooks@iospress.com

LEGAL NOTICE

The publisher is not responsible for the use which might be made of the following information.

PRINTED IN THE NETHERLANDS

Foreword

Reinhold HAUX^{a,1}

^a*Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig - Institute of Technology and Hannover Medical School*

Can bad health informatics kill? A similar question has been asked a decade ago by one of the editors of this book on evidence-based health informatics [1]. And indeed, when informatics methodology and information and communication technology (ICT) are used inappropriately this can cause severe negative effects. On the other hand we will probably all agree with her, when she writes in the same article that it "is evident that the use of modern ICT offers tremendous opportunities to support health care professionals and to increase efficiency, effectiveness and appropriateness of care" [1].

Even earlier, more than 15 years ago, the other editor stated that it "is unfortunately a truism in health care informatics ... that evaluation is undertaken rarely and inadequately" and he concludes, among others, that "integrated information systems also give new opportunity to provide effective health care service evaluation, and thus a much more robust future evidence base" [2]. As perspective he writes that "a deeper and longer-term evaluation philosophy is needed which does not stop after the initial confirmation of system functioning, but continues on with a deepening into the effects on the individual clinical services, and then on the host user organisation" [2].

Both colleagues worked during recent years continuously and intensively on how to better evaluate health care processes and outcomes in the context of health information systems, so that informatics tools and information management strategies are not 'just' applied in this context, but that their evidence has also been evaluated according to current good scientific practice. It is probably no surprise to find later joint papers of them on evidence-based health informatics, reporting about their international activities there [3].

Today there is indeed still a discrepancy in making decisions on health information system architectures, infrastructures and tools, related to considerable investments for health care organizations on the one hand and much smaller investment in evaluating the evidence of these decisions on a scientifically sound basis on the other. In therapy research, this discrepancy also existed, however much earlier. In order to overcome this discrepancy, among other methods, randomized clinical trials are now forming an important part of evaluating and making decisions on good clinical practice.

Health information systems in their current form have only existed for a few decades. And they are still in continuous change. Their complexity is high and often underestimated. Insofar it is understandable that in the beginning priority has been put on successful and stable implementations and on feasible solutions concerning organizational issues. However, this initial phase should now clearly be regarded as

¹ Corresponding author: Prof. Dr. Reinhold Haux, President and Honorary Fellow of IMIA, the International Medical Informatics Association, reinhold.haux@plri.de.

finished and so the need for systematically looking for evidence must also be demanded for the practice of health informatics – for the sake of patients, of health care organizations, and for high-quality and efficient health care.

This book on evidence-based health informatics, edited by two colleagues with high international reputations in this field, is timely and very welcome. They successfully invited excellent authors worldwide to report and to discuss about the many aspects of evidence-based health informatics.

What has to be considered when reading the book?

In their preface Elske Ammenwerth and Michael Rigby report about the book's objective: It "seeks to meet the need for better understanding of the need for robust evidence to support health IT, give insight into health IT evidence and evaluation as its primary source, and to promote health informatics as the underpinning science".

They state that a reader should not expect a cook book with a few recipes on how to successfully cook some delightful meals of evidence-based health informatics. Editors and authors seek "to inform the reader on the wide range of knowledge available, and its necessary use according to circumstances".

It is good to have books helping to get better evidence in the practice of informatics, in particular in the context of health information systems.

Dear reader, please select and prepare your appropriate meal. And ...: Bon appetit!

References

- [1] E. Ammenwerth E, N.T. Shaw. Bad health informatics can kill - is evaluation the answer? *Methods Inf Med* **44** (2005), 1-3.
- [2] M. Rigby. Health informatics as a tool to improve quality in non-acute care - new opportunities and a matching need for a new evaluation paradigm. *Int J Med Inform* **56** (1999), 141-50.
- [3] M. Rigby, E. Ammenwerth, M.C. Beuscart-Zephir, J. Brender, H. Hyppönen, S. Melia, P. Nykänen, J. Talmon, N. de Keizer. Evidence Based Health Informatics: 10 Years of Efforts to Promote the Principle. Joint Contribution of IMIA WG EVAL and EFMI WG EVAL. *Yearb Med Inform* **8** (2013), 34-46.

Preface

Elske AMMENWERTH^{a,1} and Michael RIGBY^b

^a*UMIT – University for Health Sciences, Medical Informatics, Hall in Tirol, Austria*

^b*Keele University, Keele, Staffordshire, ST5 5BG, U.K.*

1. The Need for a Better Understanding of Evidence Based Health Informatics

This book seeks to meet the need for better understanding of the need for robust evidence to support health information technology (health IT), give insight into health IT evidence and evaluation as its primary source, and to promote health informatics as the underpinning science. Health IT is a major field of investment in support of healthcare delivery, either directly in clinical systems such as clinical decision support systems (CDSS), or indirectly in terms of care delivery organisation and scheduling, and the provision of electronic records, diagnostic recording and telemedicine.

Readers, patients and the general public might reasonably assume that all such applications of health IT, which are all to some degree life-critical, as well as determinants of efficacy and of resource usage, would have the same scientific underpinning and verification of safety and effectiveness in use as expected for instance with pharmaceuticals and medical devices. Sadly they do not. While many applications are efficient, welcomed by their users, and essential to modern healthcare, this does not apply to all. Unfortunately, some cause user frustration and result in inefficiency in use, some require users to ‘break the rules’ with what are known as ‘work arounds’, and unfortunately some are known to have inconvenienced patients or caused harm, including on occasions death.

How we have reached this state is a long story not to be entered into here. However, it is in part a story of systems having initially been simple, but as their complexity increased there were no moves to introduce validation mechanisms. And in part this was a mirroring of other sectors, from logistics to commerce, where systems are assumed to be efficient and effective – which in general they are.

But few IT system users are as dependent on the ‘black box’ of the computer as patients and health professionals, nor is so much at stake, while conversely pressure on resources minimised investment in testing and fine tuning. And whereas in most economies consumers have a choice of for instance online retailer and can thus select one whose system suits them, and indeed employees can choose to move, in healthcare patients and professions have systems imposed upon them by organisation policy, or by even higher policy decision. The system user is not a user by choice, but the victim of an external choice. And the cost of imperfect decisions includes clinical mistakes, user

¹ Corresponding author: Prof. Dr. Elske Ammenwerth, Institute of Biomedical Informatics, UMIT – University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, elske.ammenwerth@umit.at.

frustration in their professional work, loss of patient utility, waste of resources, and on occasions wrong treatment.

2. Welcoming a Wide Readership

This book is consequently addressed at a wide readership. Health policy makers – including politicians, national policy makers, and those heading delivery organisations – are too often overlooked. They are one of our audiences, as they make major health IT decisions. Leaders of clinical and other professional bodies might well be considered in this category.

Clinicians in service delivery, and with clinical responsibilities in care delivery organisations, are certainly a key part of the intended audience. Patients, and patients' organisations, are targets, too. They are clear stakeholders in health IT, and therefore should be aware of the importance of ensuring that health IT applications are optimal, not least with regard to safety and efficiency.

It goes without saying that also health informaticians, especially practitioners in a care delivery setting, are a key part of our intended readership. Too often theirs is seen as a technician position, selecting, installing and maintaining equipment. They should in fact have a much deeper role, leading the acquisition and dissemination of knowledge about what is possible, what should be done, and how best to ensure full achievement.

The next group is those in the academic health informatics community who are essential to the health informatics discipline and the fostering and application of the science. They undertake much of the research to create the knowledge; they also lead the teaching of health informatics not just to informaticians, but hopefully to clinicians and policy makers.

And finally, members of the vendor industry – shareholders, directors, managers, and developmental informaticians, would be well counselled to consider more deeply and carefully the science underpinning their domain, but above all the science measuring their success in person impact terms as much as ones of economics or market share – recognising of course that scientifically proven beneficial outcomes are far more powerful a marketing tool, as well as far more ethical, than unproven claims and aspirations.

3. The Purpose of the Book

This is not a health IT evidence cook book on how to serve delightful dishes. No, it seeks to inform the reader on the wide range of knowledge available, and its necessary use according to circumstances. It is much more a compendium of menus, with values and characteristics built in. A balanced diet should also be facilitated by the items offered – the starter to whet the appetite for evidence at the conception and inception of an application; a main event which studies in depth a particular application; and a dessert of how to make the resultant knowledge available and effective for future wiser actions.

The spread of contents is intended to:

- Promote the concept of evidence-based health informatics.
- Promote health informatics as a scientific discipline.
- Look at health IT evaluation from a broader perspective of evidence-based health informatics.
- Provide an overview on health IT evaluation methodology and its relation to the life cycle of information systems.
- Define open challenges and issues to be solved in the next years.
- Enable the wider health community to appreciate and appraise evidence about health informatics application.

4. Organization of the book

The book comprises free-standing scientific chapters, but with a considerable degree of cross-linking. It is divided into three parts:

Part I: The context and importance of evidence-based health informatics

The three chapters in this first part focus on the need for evidence in health informatics, and on the idea of evidence-based health informatics. These chapters present an overall framework for the subsequent chapters.

Part II: Methodological considerations of health IT evaluation

The 19 chapters in this part present methodological approaches for health IT evaluation from a broad range of perspectives, including chapters on theoretical foundation of evaluation, practice of health IT evaluation including selection of design and methods, quantitative and qualitative approaches to health IT evaluation, and systematic reviews and meta-narrative reviews, to name just a few topics. This part intends to give the reader a solid overview and theoretical and practical guidelines on all issues around health IT evaluation.

Part III: Ensuring the relevance and application of evidence

These six chapters discuss issues around the quality of evidence, including quality of evaluation studies, publishing and searching studies, and training of health IT evaluation.

All chapters comprise a list of selected recommended further readings that allow the reader to go into the depth of a given topic where necessary. In addition, all chapters present food for thought questions at the end that may guide the reader to think further about the topic, and to identify cross-links between chapters. This part especially makes the book useful for classroom usage.

5. Cross-cutting key themes to be considered throughout

Across the described three parts of the book, a number of cross-cutting themes emerge as important:

5.1. Application Objective

No health IT system, whether electronic records or a telemedicine application, should be seen as an end in itself. If it is to be designed, selected and implemented effectively, it is essential to know the core purpose in healthcare delivery and in health outcomes terms.

5.2. System, Setting and Clinical Contexts

Settings are all different. There are different economies and stages of development, and different cultural norms, not just across the globe, but within continents and even within countries. Health IT systems, especially those concerned with organization and delivery, need to be appropriate to their context. Then there are the different contexts of primary, secondary and tertiary care. There are self-contained delivery organizations, and those which necessarily network with others supporting the same patient. Different clinical domains and disease groups, and different health professions, each have different needs, and indeed these will be different between adjacent settings or countries. Examples speak clearly from the pages.

5.3. Subsidiary Effects

Even where the core objective of health IT is clear, there may be secondary objectives which have not been considered in the same depth, but whose own significance becomes surprisingly clear when change is proposed. For instance, the delivery of a dermatology service to a scattered population and with a shortage of specialists may well be boosted by teledermatology, as indeed may remote clinical education. But if subsidiary roles (maybe not even formalized) of a dermatologist are to provide second opinions and advice to a range of other disciplines and their patients, ranging maybe from pediatrics to oncology, as well as informal education to other professions, and the dermatologist post is removed from the hospital setting and is no longer available, these other functions will be lost and a different group of patients be disadvantaged.

5.4. Stakeholders

On many occasions a health IT approach is the result of a high level political or ‘modernization’ decision. The politician or senior policy maker thereby then becomes a vested interest stakeholder, determined to portray ‘success’. But there are many other stakeholders, including patients, clinicians, and staff in support departments. Evidence, evaluation, and grounded policy making are only possible if all the stakeholders involved in a topic are identified, and their needs and views assessed. Many of the

stakeholders, from politicians through informaticians to clinicians, all have their livelihoods at stake in the event of a serious problem with the applied use of a health IT system.

5.5. Safety and Health Outcomes

It should not be necessary to identify health outcomes, and the safety of patients and staff, as a composite theme. But so often these fundamental aspects of health care are overlooked, or assumed indirectly, in health IT in the quest for efficiency or savings. But concerns about patient effects and outcomes should be the prime currency, given that health outcome is the core purpose of any health system, or of any individual interaction or intervention. Efficiency, economics, and modernization have their value, but only in support of the health outcome goal.

5.6. Investment in Evidence, Evaluation and Education

Knowledge undiscovered in a system, and evidence locked in obscure (to the wider reader) scientific literature, are of limited value. As promoted by the International Medical Informatics Association (IMIA) in its 2013 Yearbook [1], and in the context of developing countries by the World Health Organisation with its Bellagio workshop [2], health informatics should be evidence-based. This requires investment in evaluation to produce the evidence, in knowledge bases to make the evidence more readily available, and in education of policy makers, informaticians and clinicians in how to set standards of evidence use. This too is a recurrent theme through many of the contributions, if only by the plea for more notice to be taken of each dimension a contribution covers.

6. The Evidence for Evidence

Above all, the intention in that this book speaks out loudly and clearly that Evidence-Based Health Informatics is the necessary basic standard, ethically and morally, but also to achieve successfully the intended aims of the domain, and of any specific application. Each chapter speaks eloquently from its own viewpoint on the advantages of building forward based on evidence, while many also show examples of the weaknesses and adverse results of a less rigorous or scientific approach.

We hope that this book meets the needs of readers, but more importantly that it thereby strengthens health systems' informed use of health IT to the benefit of local populations and patients.

References

- [1] International Medical Informatics Association, *IMIA Yearbook of Medical Informatics 2013: Evidence-Based Health Informatics*, Stuttgart: Schattauer, 2013.
- [2] WHO, *Call to Action on Global eHealth Evaluation - Consensus Statement of the WHO Global eHealth Evaluation Meeting*, Bellagio, http://www.ghdonline.org/uploads/The_Bellagio_eHealth_Evaluation_Call_to_Action-Release.docx, last access 11 February 2016.

This page intentionally left blank

Contents

Foreword <i>Reinhold Haux</i>	v
Preface <i>Elske Ammenwerth and Michael Rigby</i>	vii
Part I: The Context and Importance of Evidence-Based Health Informatics	
The Need for Evidence in Health Informatics <i>Michael Rigby and Elske Ammenwerth</i>	3
Evidence-Based Health Informatics and the Scientific Development of the Field <i>Jeremy C. Wyatt</i>	14
Health IT for Patient Safety and Improving the Safety of Health IT <i>Farah Magrabi, Mei-Sing Ong and Enrico Coiera</i>	25
Part II: Methodological Considerations of Health IT Evaluation	
Theoretical Basis of Health IT Evaluation <i>Jytte Brender McNair</i>	39
Understanding Stakeholder Interests and Perspectives in Evaluations of Health IT <i>Lisa Lee and Aziz Sheikh</i>	53
Domains of Health IT and Tailoring of Evaluation: Practicing Process Modeling for Multi-Stakeholder Benefits <i>Sara Marceglia, Simona Ferrante, Stefano Bonacina, Francesco Pincioli, Irene Lasorsa, Chiara Savino and Giuseppe Pozzi</i>	63
Evidence-Based Health Informatics Frameworks for Applied Use <i>Catherine K. Craven, Brad Doebbeling, Dominic Furniss, Richard J. Holden, Francis Lau and Laurie L. Novak</i>	77
Ensuring the Quality of Evidence: Using the Best Design to Answer Health IT Questions <i>Charlene R. Weir</i>	90
Mixed Methods: A Paradigm for Holistic Evaluation of Health IT <i>Philip J. Scott</i>	102
Evaluation of People, Social, and Organizational Issues – Sociotechnical Ethnographic Evaluation <i>Bonnie Kaplan</i>	114
From Usability Engineering to Evidence-Based Usability in Health IT <i>Romaric Marcilly, Linda Peute and Marie-Catherine Beuscart-Zephir</i>	126

Participatory Design, User Involvement and Health IT Evaluation <i>Andre Kushniruk and Christian Nøhr</i>	139
Clinical Simulation as an Evaluation Method in Health Informatics <i>Sanne Jensen</i>	152
Economic Evaluation of Health IT <i>Daniela Luzi, Fabrizio Pecoraro and Oscar Tamburis</i>	165
Health Care Performance Indicators for Health Information Systems <i>Hannele Hyppönen, Elettra Ronchi and Julia Adler-Milstein</i>	181
Evaluating the Impact of Health IT on Medication Safety <i>Hanna M. Seidling and David W. Bates</i>	195
Evaluation of Implementation of Health IT <i>Kathrin Cresswell</i>	206
Evolving Health IT Systems Evaluation: The Convergence of Health Informatics and HTA <i>Persephone Doupi</i>	220
Realizing the Potential of Patient Engagement: Designing IT to Support Health in Everyday Life <i>Laurie L. Novak, Kim M. Unertl and Richard J. Holden</i>	237
Ensuring Evidence-Based Safe and Effective mHealth Applications <i>Bárbara Vallespin, Joan Cornet and Anna Kotzeva</i>	248
Systematic Reviews and Meta-Analysis of Health IT <i>Christine Urquhart and Rosemary Currell</i>	262
Going Beyond Systematic Reviews: Realist and Meta-Narrative Reviews <i>Terese Otte-Trojel and Geoff Wong</i>	275
Part III: Ensuring the Relevance and Application of Evidence	
Quality of Health IT Evaluations <i>Pirkko Nykänen and Johanna Kaipio</i>	291
Publishing Health IT Evaluation Studies <i>Elske Ammenwerth and Nicolet F. de Keizer</i>	304
Finding, Appraising and Interpreting the Evidence of Health IT <i>Andrew Georgiou</i>	312
Evaluation of Health IT in Low-Income Countries <i>Tom Oluoch and Nicolet F. de Keizer</i>	324
Learning, Training and Teaching of Health Informatics and Its Evidence for Informaticians and Clinical Practice <i>Evelyn Hovenga and Heather Grain</i>	336

Forward Outlook: The Need for Evidence and for Action in Health Informatics <i>Michael Rigby, Elske Ammenwerth and Jan Talmon</i>	355
--	-----

Appendix

Appendix: Resources on Health IT Evaluation	367
Author Index	369

This page intentionally left blank

Part I: The Context and Importance of Evidence-Based Health Informatics

This page intentionally left blank

The Need for Evidence in Health Informatics

Michael RIGBY^{a,1} and Elske AMMENWERTH^b

^a*Keele University, Keele, Staffordshire, ST5 5BG, U.K.*

^b*UMIT – University for Health Sciences, Medical Informatics, Hall in Tirol, Austria*

Abstract. While the use of health IT applications has increased rapidly over past decades, this does not compare strongly with other business sectors. Both reluctance to invest in, and lack of demand to use IT systems may in part be due to lack of robust evidence as to proven benefits. At the same time, the health IT sector has lagged behind other health technology areas in working to, and being expected to work to, robust evidence standards showing benefit and also avoidance of harm. Exacerbating this, limited availability of evidence has perpetuated this misplaced comfort in use of aspiration and expectation rather than evidence in driving investment in health IT applications. Reference back to the core principles drawn from influential thinkers shows the essential centrality of the need for evidence of safety and effectiveness, and for its use relevantly related to context.

Keywords. Health IT, health informatics, evidence, decision making, effectiveness, safety.

1. Introduction – the Exponential Growth of Information in Health and Society

It is self-evidently true that information is key to health care – information about the patient; information about treatment options; information about the ongoing care (and prevention) processes; and if we are to achieve improvement through critical learning, information on outcomes. But each of these information components within healthcare is also expanding exponentially, at individual patient level and at societal level, as care becomes more accessible; as patients as consumers become more knowledgeable and have increasing expectations; as treatments become more sophisticated and fine-tuned; as diagnostic technologies become more advanced and in themselves information rich (including digitisation of images and videos as well as documents); and as healthcare delivery becomes increasingly closely managed and coordinated among specialised health care providers.

It is also self-evident that all sectors of industry, commerce and civil society are increasingly using information technology to collect, store, and process information, but then go much further not only to create added value and deeper knowledge from that information. They have moved more fundamentally to a new paradigm of activity as a result of the fast, reliable and generally low cost of these processes. This is

¹ Corresponding author: Prof. Dr. Michael Rigby, Emeritus Professor of Health Information Strategy, School of Social Science and Public Policy, Keele University, UK, c/o Lavender Hill, 6 Carrighill Lower, Calverstown, Kilcullen, Co. Kildare, Ireland, m.j.rigby@keele.ac.uk.

conspicuous across the full range of societal activity, from computerised ‘fly by wire’ aircraft to social interaction of teenagers through social media; from computer-aided design to consumer on-line shopping; and from teleconferencing to electronic news media.

In that transition over past decades the question of ‘can we computerise that process?’ has been inverted to ‘how do we optimise the business (or social) process?’. It was half a century ago, in 1964, that Marshall McLuhan wrote *‘The Medium is the Message’* – not a slogan, though it could well have been, but an essay in his book on media and their effects on man and society. [1] The core concept was that the medium (in our case computers and related data management) change the societal expectation and processes such that the new medium defines both the service and consumer behaviour. A good example is in the hospitality and leisure industries – for many people the process of booking a holiday or a business trip is now progressed by checking web sites in real time to see what is available, where and when – a process totally different from previous processes of looking at brochures and then instructing an agent to make a booking within defined parameters, then paying by cheque. Yet the hotel itself, the stay, and the leisure and business activities are essentially unchanged. The medium (real time comparator sites and hotel web pages) creates the way the consumer thinks and acts, and in particular opens up new horizons of choice and of optimising decisions such as availing of offers, as well as speed, ease, and personal control of selection.

This has led many commercial sectors to totally redesign business processes, including (but by no means restricted to) the services they offer, and how they interface with the public. For instance, civil aviation has reshaped from a model based on travel agents and check-in desks to a model based on consumer searching and booking flights, checking in and selecting the seat of their choice, and with many other added value options such as choosing in-flight meals. The older methods of booking and checking in are still available for customers not comfortable with digital options, or for those with special situations such as cancelled flight and missed connections.

2. The Cause and Effects of Innate Conservatism

By contrast to the general commercial world, or to general consumer behaviours, in healthcare there have been few major changes in core service approaches and processes. There has not been the same handover to the consumer or user of core interface processes as has occurred in banking and insurance, in civil aviation, or in retail purchasing. Overall, healthcare has continued with its traditional processes, and seen information and communication technology (ICT) as an enabler of those.

To a large degree this is because of the special nature of health services, and in particular their special fixed assets of hospitals, diagnostic facilities, and the sensitive and complex role of the health sector, and the highly specialist staff. But many of the features claimed to be unique are not in fact so. Civil aviation systems are clearly life critical. Banking is highly personal.

Two strong underlying factors are the lack of clear strategic investment decisions, and lack of evidence, and these interlink. Those making health IT investments have a dearth of reliable and robust evidence available and accessible to them, and often have to rely on material from elsewhere and earlier systems, interlaced with vendor or industry sector promises, and a general feeling that investment in modern systems

ought to be worthwhile. Even a decision to invest, inevitably in the face of competition for other reasoned claims on development funds, can be difficult to justify, and any subsequent call for tenders may be less than perfectly constructed in terms of obtaining the most appropriate solution.

But this lack of evidence is in many ways caused by the reluctance to evaluate systems after implementation [2], and by likely publication bias against disappointing results. Policy makers may be reluctant to have less than optimal results broadcast, and vendors have an interest in protecting their products, and indeed the sector, against directly or indirectly adverse publicity. So, with the pressure on resources, it is natural not to seek to invest in, or enable, thorough rigorous analysis. However, this is both selfish and short-sighted, as it is a roadblock to policy makers making future optimally informed investment decisions.

Thus where evidence to support rational and wise strategic and investment decisions is needed, it is missing because of reluctance by others. This has all the makings of a downward spiral, if investment decisions are hampered by lack of scientifically grounded knowledge. And in turn, this leaves the field susceptible to unmoderated influence from the hopes of advocates and promises of suppliers, which however well-intentioned are unlikely to be as grounded as validated evidence.

3. Scientific Evidence and Health Informatics – addressing the aversion

This is an anomalous situation for the health sector, which in all other respects is firmly grounded in evidence, and in not making patient-related intervention changes without rigorous independent appraisal of the evidence. This applies, for instance, to pharmaceuticals, to changes to treatment regime, to prosthetic devices, or to patient-specific forms of health technology.

Yet all health IT systems affect patients. Some applications, such as decision support systems, do this in a very direct way; others such as computerised physician order entry or electronic prescribing do so by being a key part of the clinical process; but even scheduling systems and recall systems have patient effects through being tools which are depended upon to organise care, and which if malfunctioning or incorrectly operated will deprive patients of intended clinical interventions. Through such errors harm can be caused to patients, even up to death, as has been documented [3, 4, and e.g. 5, 6].

3.1. The Inappropriately Low Expectations

It is now an anomaly that health IT systems do not have to submit to the standards of science, evidence, and probity expected of all other health sciences and technologies. Not only are there the risks of direct harm, but even safe but inefficient or ineffective systems are detrimental to patients by diverting resources, or by adding to the burdens of clinicians.

Given that health IT is safety critical, directly and indirectly affecting patients as indicated above, this low expectation is difficult to justify but is being perpetrated widely. For instance, the International Society for Quality in Health Care (ISQua), with a proclaimed mission of “Inspiring, promoting and supporting continuous improvement in the quality and safety of healthcare worldwide”, held in summer 2015 what was entitled their “ISQua Education’s Global Debate for 2015” on the subject “Health

Information Technology is already improving healthcare safety and current regulation around it is sufficient" [7]. For a serious global body to think that health IT can be considered homogenously, for all its safety aspects to be linked to regulation, and for the four debaters to be drawn from just the USA, UK and Australia, would indicate how low are expectations of a true evidence-based approach to considering health IT – even when addressing the key issue of patient safety.

In the modern healthcare setting the Cochrane Collaboration is seen as the must-go-to source of robust evidence. However, as Urquhart and Currell show elsewhere in this volume² the evidence there is very sparse. Health Informatics primarily falls into the Effective Practice and Organisation of Care (EPOC) category, and there are very few robust systematic studies. This is of concern, given the role, application spread, and global ubiquity of health IT.

3.2. Risks to Patients and Practitioners

The introduction of change is desirable if this is known to be a positive move, and in simpler settings the decision maker and the user will each be able to assess the problem, the proposed solution, and the intended benefit, and be able to assess reasonably confidently that risks will be controlled and benefits achieved. However, as computing became more powerful this made health informatics more challenging. In 1995 François Grémy was one of the first to point out that the computer in clinical systems was now becoming a 'black box' whose contents and thus whose functioning the clinician could not know in detail, and thus whose effects (s)he could not know. Grémy therefore argued the need for evaluation of this new construct of informatics applications, and for this to be by class of complexity of application which would require not just health technology assessment skills but also human and psychological sciences, and social science [8].

Such an approach once technology becomes too advanced for the individual practitioner, or policy maker, to see in totality is important for maintaining the Precautionary Principle, which is European Commission policy, and assumed as a default position elsewhere, namely that change should not be made until it can be assured that it will not have harmful effects [9,10]. First and foremost, this is to protect the patient against adverse outcomes of new technologies, with patient safety always being a high priority in any health system. But secondly, it has to be remembered that the causing of harm by using a system, even one provided by their employing organisation, could be seen as a breach of their duty of care by a health professional and thus render them liable to disciplinary action, even up to the point of losing their licence to practice.

3.3. International Moves to Promote Evidence

To recognise this need to move to an evidence culture, the European Federation for Medical Informatics (EFMI) set up an Evaluation Working Group, and the International Medical Informatics Association (IMIA) a Working Group on Technology Assessment & Quality Development. In this context, in order to stimulate

² C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

further action, an expert European workshop was convened in Innsbruck in 2004, funded by the European Science Foundation as this was seen as the development of a new application of scientific study [11]. This workshop created an action plan, much of which has been achieved. Meanwhile, in the USA the Agency for Health Research and Quality (AHRQ) also has addressed the issue [12].

As this momentum developed, in 2013 IMIA took Evidence-based Health Informatics as the theme for its Yearbook of Medical Informatics [13]. This incorporated many papers on the theme of creating appropriate evidence, as well as a scene-setting paper on the decade of work to move towards a more evidence-based culture in the sector and to promote the concept and principles of Evidence-based Health Informatics [13].

The World Health Organisation has established a Global Observatory for E-Health, which seeks to promote effective use of information technologies [14]. However, this operates primarily at the level of advising on national systems, undertaking useful periodic global surveys, and producing updated E-Health Atlases and collations of national policies. This is important at the national policy and infrastructure level, but is not an application evidence source.

In 2011 the WHO joined the Rockefeller Foundation in convening a workshop in Bellagio, Italy with the title “eHealth, Evaluation, Evidence”, whose purpose was based on the credo “To improve health and reduce health inequalities, rigorous evaluation of eHealth is necessary to generate evidence and promote the appropriate integration and use of technologies.” [15]. This workshop built on an earlier initiative with Archbishop Desmond Tutu to initiate an eHealth Call to Action, and this time initiated a Call to Action on Global eHealth Evaluation. This meeting agreed nine Principles, including “5. Evidence is needed to demonstrate costs and benefits of eHealth implementations, and maximize eHealth’s beneficial impact on health system performance and population health” and “9. Improving the eHealth evidence base requires more than increased numbers of studies but also improved quality of eHealth research studies.” These are welcome principles, but the Nine Recommendations for action seem not to have had significant overall impact or follow-through.

On a more practical note, the IMIA and the European Federation of Medical Informatics (EFMI) have linked through their relevant working groups to sponsor an Inventory of Health IT Evaluation Studies and Systematic Reviews [16]. This repository was created to help researchers to identify studies that have been conducted in defined settings, and now contains approaching 2,000 references to published evaluation studies and reviews of evaluation studies of health information systems. However, it is voluntarily maintained and depends on studies being published.

4. The Limited but increasing Volume of Activity

Evaluation studies of health IT have been conducted and have been published since the early 1970s, and thus since the emergence of medical informatics as an individual discipline, but the numbers were small. Only in the mid 1990s did the number of published health IT evaluation studies start to rise steadily, with around 1% of all medical informatics papers in the year 2000 being evaluation papers [16].

Systematic reviews represent the building of the evidence base of a scientific field. In health informatics, as a recent analysis shows, a larger number of systematic reviews only started to appear after 2005 [17]. This 10-years lag behind in systematic reviews

can be explained by the time needed to build up reviews on published studies. At the moment, around 30 systematic reviews related to health IT are being published annually, with a clearly increasing trend in number. This leads to a slowly, but steadily growing evidence base of health informatics.

5. Getting Decision-makers to Expect and Use Evidence

While one serious problem is that the healthcare sector is accepting of the fact that health IT investment decisions are made based on a lower standard of evidence than rightly is applied in all other areas of healthcare, including pharmaceuticals and medical devices, the corollary is that many decisions are made on inadequate evidence. Often this is at a political level, or at a national level by generic policy makers rather than informatics experts – who themselves have inadequate data to draw on. Two examples show the span of depth given to such decisions.

At the one extreme, it has already been reported that the huge policy decision to create a universal electronic health record system for the NHS in England was made by Tony Blair as Prime Minister in a meeting lasting under two hours, and devoid of health informaticians or sound evidence [18, 19]. At the other extreme, many countries have set up mechanisms which may have taken years to collect and assimilate evidence, which then runs the risk of being out of date. Evidence needs to be constantly refreshed, and so also do policies. The intransigence of informatics innovators to application of new evidence, resulting in their opposing updated versions of their vision, has been documented [20], and is even more likely at institutional level.

Thus creation of, availability and use of evidence for decision-making in health IT have both a pull and push effect, both of which are weak [21, 22]. There should be an expectation from policy makers at all levels to be able to get comprehensive, robust evidence on health IT matters as they would for any other type of health decision, and to be willing to invest in its creation through policy or research channels. It thus needs stronger demand to stimulate the process, and the funding, of evaluation and of evidence publication. At the same time the health informatics community needs to raise its game, to be much more rigorous in generating an evidence culture and processes, to enable supply of good evidence and the establishment of its position as a serious partner alongside the other health sciences and related production sectors. This was also recognised by the WHO-Rockefeller Bellagio event, whose very relevant concluding recommendation was to: “Create a multi-stakeholder web-based platform for constructive sharing, publication and learning from successes and failures. Include a registry of eHealth evaluation studies and results, and a repository of evidence-based eHealth best principles and practices”[16]. Neither the push nor the pull have so far been strong enough to see this implemented, though in global terms the cost would be small and the potential payback large.

6. Returning to Visionaries and distilling Core Principles

While this contribution extols the need for robust evidence, and an evidence culture, there is equally a need for moderation and focus. Because health IT has many aspects, each of which has many stakeholders, and many dimensions from safety to cost-benefit, and then each must be seen in both the national health system and the local context,

there is a real risk of moving eventually toward an overload of evidence and issues. This can lead to the situation of 'paralysis by analysis', and the perfect becoming the enemy of the good – which is a stage too far beyond the current lack of knowledge in many cases about what is good. It may therefore be sound to return to core principles, and the insight of key thought leaders.

6.1. *Hippocrates*

Hippocrates first and foremost gave us the dictum *non nocere* – do no harm. That should be a key tenet – yet too often it is conveniently passed over. However, Hippocrates did not just bring ethical principles to healthcare – he brought the principle of systematic record keeping which underpins current thinking and delivery in healthcare, and of subsequent analysis to create new knowledge [23]. While clearly Hippocrates was not talking of computers, he was instilling the importance of recording full evidence as the source of robust and reliable knowledge to inform future actions. Within this, he emphasised the importance of patient outcomes as the prime consideration and currency. This indicates that when implementing health IT full records should be kept of the actual effects, particularly on outcomes, so as to form the basis of analysis and shared learning.

Despite the fact that, as explained earlier, we do know that health IT can do harm, and frequently we choose not to enquire too deeply, either before implementation as to the effects of the application approach, or after implementation concerning the system in operation, that 'blind eye' approach is not defensible. In effect policy-makers and sections of the supplier industry are acting unethically by Hippocrates' standards by not protecting against possibly (or actually) causing harm.

6.2. *George Boole*

Our second key thought leader is George Boole, a largely self-educated man who moved from being a teacher in Lincolnshire in 1849 to be founding Professor of Mathematics at the newly-established and thus somewhat independent-minded Queen's College Cork, now University College Cork. Boole is often held up as the founder of computing, which indirectly means that he was instrumental in medical informatics. Of course, he had no concept of computers, and incidentally might well have views on our cause since he died young as a result of his caring wife insisting on a non-evidence based treatment for a severe winter chill.

However, Boole's underlying mission was to systematise thought [24]. One core concept included the differentiation between 'OR' and 'AND'. In the ever increasing complexity of evidence, including the evidence related to health IT applications, this central differentiation between effects that are mutually exclusive (including the opportunity cost of commitment of resources of investment and time), and those which may be accumulative (particularly unintended as well as intended outcomes) is important. As decision-making and underpinning analysis continue the tendency to increasing complexity, refining them back by simple rule will help clarify the options, and frame the consequences in format closer to summated net effects, thereby increasing clarity and accountability.

6.3. Archie Cochrane

Whereas Boole's advanced thinking has been simplified to its core, our next visionary, Archie Cochrane, has had his simple ideas turned into an industry, with increasing complexity but arguably with a deviation from his core clarity. While the gold standard developed by the Cochrane Collaboration is the double blind RCT – which is so hard to achieve in health IT implementations, and which in clinical fields is remote from the reality of comorbidity and local treatment contexts – Cochrane's starting points were quite different. Cochrane's first peer reviewed scientific paper was not set in purposefully constructed controlled trial settings, but in prisoner of war camps in wartime Germany [25] – not exactly ideal conditions, but yielding evidence from observation, and the first of four studies from those settings. By 1951 he was publishing epidemiological analysis from the applied Medical Research Council Unit in South Wales that he was to make so effective, starting with [26]. The later, reflective main opus of Cochrane gives us the core concepts in its title – *Effectiveness and Efficiency* [27]. These objectives, rather than an elusive complex method, and recognition of the challenges of very real worlds, should be taken as Cochrane's core insights for us.

6.4. Evidence in the Real World Context

Thus from these three thought leaders we understand the importance of evidence based approaches; contemporaneous recording of all aspects a situation and interventions; looking at the real world and at context; the importance of patient outcomes as the most important currency; systematizing our thinking to be most effective; and above all of looking at avoiding doing harm while looking for optimal effectiveness and efficiency. In the modern field of using health IT to harness a new science in the service of health and healthcare, these key principles point to the importance of Evidence-based Health Informatics (EBHI) as the essential route to take.

Yet this should be followed in a balanced and reflective way, not as the unthinking applications of a formula or rule set. From the paradigm of Evidence-based Medicine, both Sackett as a core early protagonist, and Greenhalgh as a modern informed commentator, have emphasized that the evidence is a tool to be applied informedly [28, 29]. Context, and application, are vital and are key essentials or professionalism.

7. The Motivation for this Book

This dearth of good evidence in the face of the need for it has provided the impetus for this book. It is clear that the communities of policy makers, informaticians, system suppliers, and healthcare and other users, are poorly served by the shortage of effective objective evidence in health informatics, by the limited volume of activity and publishing and by the lack of readily available evidence sources.

It is important for all in the health sector to realise and recognise the importance of evidence, what aspects it should cover, how it should be obtained, and thus how to assess it. Health informaticians, and users of health IT systems, should be no exceptions to this. The inclusion of all domains of healthcare, and the many aspects to be considered, are intended to give a comprehensive overview and source of understanding.

8. Discussion and Conclusion

There are increasing expectations that healthcare, and activities to support the health of the population, will be based on good science, safely applied in an equitable and efficient way. Health IT is one of the newest sciences, is in parallel with the major use of ICTs in most sectors of societal endeavour both commercial and social, and is assumed to be harmless. However, more informed assessment shows the falseness of these assumptions. Health IT can be well applied and strongly beneficial, but applications can also be badly conceived or applied; resources can be wasted; staff and patients can be disadvantaged; and actual harm and death can result.

The use of robust evidence drawn impartially from evaluation and from objective observational studies, and informedly applied in the local context, is the essential methodology for policy makers of all types and levels. This book seeks to put forward the types and sources of evidence applicable to each type of situation, how to create and to source that evidence, and the dimensions to be taken into account in making health IT decisions in any situation. We hope that it will result in better decisions, and thus in better health for populations.

Recommended further readings

1. M. Rigby, E. Ammenwerth, M.-C. Beuscart-Zephir, J. Brender, H Hyppönen, S. Melia, P. Nykänen, J. Talmon, N. de Keizer, Evidence Based Health Informatics: 10 Years of Efforts to Promote the Principle. *Yearb Med Inform* **8**(1) (2013), 34-46.
2. J. Brender, *Handbook of Evaluation Methods for Health Informatics*, Academic Press, New York, 2006.
3. E. Ammenwerth, Evidence-Based Health Informatics: How do we know what we know? *Methods Inf Med* **54**(4) (2015), 298-307.
4. M. Berg, J. Aarts, J. van der Lei, ICT in Health Care: Sociotechnical Approaches (Editorial), *Methods Inf Med* **42** (2003), 297-301.

Food for thought

1. Are we confident, from scientific evidence, that our planned approach, application or implementation will do no harm?
2. Are health IT implementations monitored to assess their effects, not least on patient outcomes?
3. Have the real health IT investment alternatives been identified, and their anticipated cumulative effects (within the organisation, and more widely), been assessed based on sound analysis?
4. Subsequently, have these predictions of outcomes been verified, and can they be improved?

5. Have the application impacts been assessed in terms of (a) is the health IT intervention effective?; and (b) is it efficient?; based on robust analysis in the real world setting and context?

References

- [1] M. McLuhan, *The Medium is the Message*; in: M. McLuhan. *Understanding Media: The Extensions of Man*, Signet, New York, 1964
- [2] M. Rigby, Evaluation: *16 Powerful Reasons Why Not to Do It - And 6 Over-Riding Imperatives*, in: V. Patel V, R. Rogers, R. Haux (eds.), *Medinfo 2001*, Proceedings of the 10th World Congress on Medical Informatics, IOS Press, Amsterdam, 2001, pp. 1198-1202.
- [3] E. Ammenwerth, N.T. Shaw, Bad health informatics can kill - is evaluation the answer? *Methods Inf Med* **44**(1) (2005), 1-3.
- [4] Web list "Bad Health Informatics can kill" (accessible via the Bad Health Informatics link at <http://iig.umit.at/efmi>), last access 11 February 2016.
- [5] Y.Y.I. Han, J.A. Carcillo, S.T. Venkataraman, R.S. Clark, R.S. Watson, T.C. Nguyen, H. Bayir, R.A. Orr, Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system, *Pediatrics* **116**(6) (2005), 1506-12.
- [6] F. Magrabi, M.S. Ong, W. Runciman, E. Coiera, Using FDA reports to inform a classification for health information technology safety problems, *J Am Med Inform Assoc* **19**(1) (2012), 45-53.
- [7] ISQua, *Education Global Debate on Health Information Technology*, <http://www.isqua.org/education/activities/debates/live-isqua-debate-2015>, last access 11 February 2016.
- [8] F. Grémy, M. Bonnin, *Evaluation of Automatic Health Information Systems – What and How?*, in: van Gennip EMSJ and Talmon J (Eds.), *Assessment and Evaluation of Information Technologies in Medicine*, IOS Press, Amsterdam, 1995.
- [9] European Commission. *Communication from the Commission on the Precautionary Principle*, COM (2000) 1, Brussels, 2000.
- [10] E. Fisher, J. Jones, R. von Schomberg (eds), *Implementing the Precautionary Principle: Perspectives and Prospects*, Edward Elgar, Cheltenham, UK and Northampton, MA, 2006.
- [11] E. Ammenwerth, J. Brender, P. Nykänen, H.-U.- Prokosch, M. Rigby, J. Talmon, et al., Visions and strategies to improve evaluation of health information systems: Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, *Int J Med Inform* **73**(6) (2004), 479-91.
- [12] C.M. Cusack, C. Byrne, J.M. Hook, J. McGowan, E.G. Poon, A. Zafar, *Health information technology evaluation toolkit: 2009 Update* (Prepared for the AHRQ National Resource Center for Health Information Technology under Contract No. 290-04-0016), AHRQ publication No. 09-0083-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2009.
- [13] M. Rigby, E. Ammenwerth, M.-C. Beuscart-Zephir, J. Brender, H. Hyppönen, S. Melia, P. Nykänen, J. Talmon, N. de Keizer, Evidence Based Health Informatics: 10 Years of Efforts to Promote the Principle, *Yearb Med Inform* **8**(1) (2013), 34-46.
- [14] WHO, Global Observatory for eHealth, *eHealth and innovation in women's and children's health: A baseline review*, <http://www.who.int/goe/en>, last access 11 February 2016.
- [15] Call to Action on Global eHealth Evaluation: Consensus Statement of the WHO Global eHealth Evaluation Meeting, Bellagio, September 2011, http://www.ghdonline.org/uploads/The_Bellagio_eHealth_Evaluation_Call_to_Action-Release.docx, last access 11 February 2016.
- [16] E. Ammenwerth, N. de Keizer, An inventory of evaluation studies of information technology in health care: Trends in evaluation research 1982 - 2002, *Methods Inf Med* **44** (2005), 44-56.
- [17] E. Ammenwerth, Evidence-Based Health Informatics: How do we know what we know? *Methods Inf Med* **54**(4) (2015), 298-307.
- [18] M. Rigby, Optimising Health Informatics Outcomes - Getting Good Evidence to Where it Matters; *Methods Inf Med* **54**(4) (2015), 295-7.
- [19] T. Ritter, Who was at Downing Street Meeting? *Computer Weekly*, 17 February 2008, <http://www.computerweekly.com/blogs/public-sector/2008/02/who-was-at-downing-street-npfi.html>, last access 11 February 2016.
- [20] S.H.F. Guist, M. Rigby, *The Rise and Fall of the Innovator*, in: K.C. Lun, P. Degoulet, T.E. Piemme, O. Rienhoff, *Medinfo 92*, Proceedings of the Seventh World Congress on Medical Informatics, Geneva Palexpo, Switzerland, 6-10 September 1992, Elsevier Science Publications BV, North-Holland, Amsterdam, 1992.

- [21] M. Rigby, Evaluation: *16 Powerful Reasons Why Not to Do It - And 6 Over-Riding Imperatives*, in V. Patel, R. Rogers, R. Haux (eds.), Medinfo 2001, Proceedings of the 10th World Congress on Medical Informatics, IOS Press, Amsterdam, 2001, pp. 1198-1202.
- [22] M. Rigby, Evaluation – the Cinderella Science of ICT in Health, in: Kulikowski C, Haux R (Eds.) *IMIA Yearbook of Medical Informatics* (2006); Schattauer, Stuttgart, pp. 114-120.
- [23] E. Tsionpanou, S.G. Marketos, Hippocrates: timeless still, *Journal of the Royal Society of Medicine* **106** (2013), 288-292.
- [24] G. Boole, *An Investigation of the Laws of Thought*, Walton & Maberly, London, 1854.
- [25] A.L. Cochrane, Tuberculosis among Prisoners of War in Germany, *Br Med J* **2** (4427) (1945), 656-8.
- [26] A.L. Cochrane, C.M. Fletcher, J.C. Gilson, P. Hugh-Jones, The Role of Periodic Examination in the Prevention of Coalworkers' Pneumoconiosis, *Br J Industrial Medicine* **8** (2) (1951), 53-61.
- [27] A.L. Cochrane, *Effectiveness and efficiency: random reflections on health services*, Oxford University Press, Oxford, 1973.
- [28] D.L. Sackett, W.M. Rosenberg, J.A. Gray, R.B. Haynes, W.S. Richardson, Evidence based medicine: what it is and what it isn't, *BMJ* **312**(7023) (1996), 71-2.
- [29] T. Greenhalgh, R. Snow, S. Ryan, S. Rees, H. Salisbury, Six 'biases' against patients and carers in evidence-based medicine, *BMC Med* **13** (2015), 200.

Evidence-based Health Informatics and the Scientific Development of the Field

Jeremy C. WYATT^{a,1}

^a*Leadership chair in eHealth Research, Leeds University, UK*

Abstract. We define and discuss the nature of Evidence-based Health Informatics (EBHI), the kind of evidence health informatics researchers must generate to make EBHI a reality, and how we should grade such evidence. We propose adding principle-based evaluation studies to the list of common evaluation study types, and outline how to carry out such studies to generate evidence that will prove useful for establishing EBHI. The main purpose of a principle-based evaluation study is to test the impact on system acceptability, usage or effectiveness of a generalizable system design principle, so we also explore when during the system design process such principles are needed, and which disciplines are most promising as sources of design principles. We conclude with some challenges for EBHI, a list of the benefits of adopting this approach, and a test to ensure that we are advancing in the direction of science, as opposed to pseudoscience.

Keywords. Health informatics, evaluation methods, evidence-based health, professional organisation and administration.

1. Introduction: What is evidence-based health informatics, and why does it matter?

Recently, a national body asked for advice on how to improve the quality of patient data captured by electronic health records. After a reminder that data quality has several dimensions [1], I suggested some strategies that might work and should be easy to implement: audits of data quality with weekly feedback to users, alternative screen layouts or data entry widgets, adding pop-up definitions of data items, or making certain data items obligatory. Fortunately, they did not ask for evidence about the relative impact of each strategy nor for which types of data or users each strategy is most appropriate. They would certainly have asked for such evidence if their question had been about which drugs work in a named disease, and there is copious high quality evidence about drug effectiveness. However, there is scarcely any good quality evidence about how to improve data quality [2], despite this being a common question and one which we in health informatics should be uniquely qualified and able to answer.

This is a major criticism of health informatics as a profession: we have not yet assembled a robust evidence base to answer basic questions about common clinical

¹ Corresponding author: Prof. Jeremy Wyatt, Leadership chair in eHealth, Research, Faculty of Medicine and Health, Worsley Building, University of Leeds, Leeds, LS2 9JT UK, j.c.wyatt@leeds.ac.uk. From 1-1-2016: Director, Wessex Institute of Health, University of Southampton, Alpha House Enterprise Road, Chilworth, SO16 7NS.

information management problems. Instead, we rely on experts, untested theories, poorly understood principles or low grade evidence. In short, the prevailing approach of health informatics is unscientific, so we cannot reliably predict the impact of the strategies we use to build or improve information systems.

To remedy this problem we need what can be called “Evidence-based Health Informatics (EBHI)”. This can be defined by analogy with Evidence-based Medicine [3], which means doctors using the results of well-designed research in intact humans (evidence) to guide their patient management decisions, rather than relying on advice from experts or reasoning from first principles like pathophysiology. This requires the medical profession to take responsibility for developing and curating this knowledge base [4], a task which is now undertaken by the Cochrane Collaboration².

By analogy, EBHI means that the people designing, developing and implementing health information systems should be able to rely on an explicit evidence base derived from rigorous studies on what makes systems clinically acceptable, safe and effective – not on basic science or experts alone (see Figure 1).

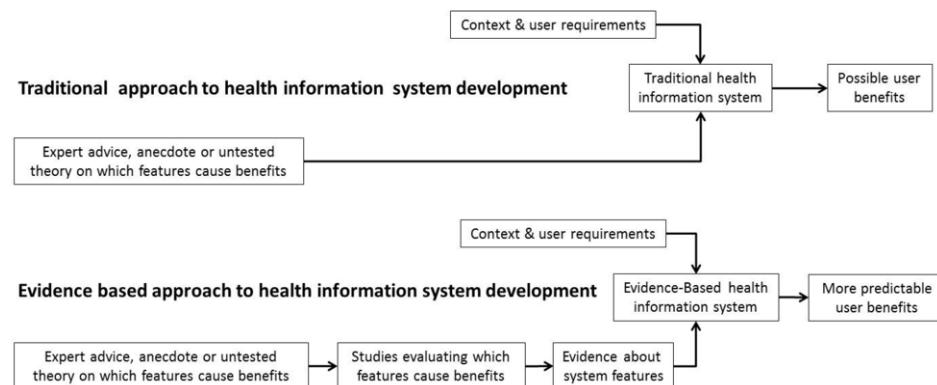


Figure 1. Comparison of traditional and evidence-based system development methods.

Once we in health informatics assemble this evidence base, this means that the design and implementation decisions taken by system developers will usually lead to predictably acceptable, safe, affordable and effective systems – which unfortunately is not the case at present [5]. The analogy for system development will be with cardiology or bridge building: with EBHI, system developers will become professionals relying on a proven body of knowledge (about test accuracy and drug effectiveness in the former case, or construction materials and how to use them in the latter), not craftsmen relying on a lifetime’s experience of trial and error [6]. This will slow the excessive pace of technical innovation in our field, with every new technological development being tested for its contribution to important patient or health system outcomes. Over time, this evidence-based approach will lead to a number of benefits (see Table 1).

² For more information on the Cochrane Collaboration, see: Urquhart et al., Systematic reviews and meta-analysis of health IT. in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

Table 1. Some benefits of Evidence-Based Health Informatics for various stakeholder groups.

Stakeholder	Benefit
Patient	Safer, more effective health information systems; faster, more efficient care
Clinician / health professional	Systems that are easier to learn and use, fit better with clinical workflows, are safe and effective, with no surprises. Lower professional liability premiums as a result.
System developers	A clear set of guidelines for use in system development
Funders of clinical information systems e.g. health insurers	Systems that cost less and have predictable benefits
Tax payers, the public	Systems that cost less and have predictable and optimised benefits
Professional indemnity organisations	More reliable, effective health systems, so fewer legal claims against health professionals
System purchasers	A clear set of criteria to use during system procurement
People working in health informatics	Clarity about what to teach students Clarity about what works, when consulted about this A strong core of knowledge to inform future development of the profession
Regulatory organisations e.g. Medicine & Healthcare products Regulatory Agency (MHRA, UK), Food and Drug Administration (FDA, US)	An evidence base of tested principles against which to check new health information technologies
Clinical guideline developers	Good evidence on which to base their recommendations to use - or avoid - clinical information systems

The obvious next question is, what kind of evidence will we need to enable EBHI, and where will it come from? This is addressed next.

2. How to practise evidence-based health informatics?

2.1 What kind of evidence will we need to realise EBHI?

Evidence comes from primary and secondary research studies, but the best research design varies for each research question [7]. If we focus on the most important question in health informatics – which system design and development methods lead to safe and effective systems – then we can develop an approximate hierarchy of evidence for EBHI, analogous to the hierarchy of evidence for health technologies – drugs etc. At the top of this hierarchy are the most reliable sources of evidence, including systematic reviews³ and randomized controlled trial⁴ and the evidence gets steadily less reliable as we descend the hierarchy. A draft evidence hierarchy to support EBHI is shown in Table 2.

This implies an addition to the type of evaluation studies that we conduct in health informatics, adding to the usual studies (designed to answer the question “does it work”) studies that ask “Will systems based on this generic design principle work better than other systems?”. This is explored in the next two sections.

³ See also: Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics 222, Stud Health Technol Inform, IOS Press, Amsterdam, 2016.

⁴ See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: *ibid.*

Table 2. Draft hierarchy of evidence to support EBHI, loosely based on [8].

Level	Type of evidence to support “What works?” questions
1a	Systematic reviews of well-designed impact studies designed to directly test a relevant design principle, with low heterogeneity
1b	Systematic reviews indirectly comparing well-designed impact studies that evaluate systems that demonstrate or lack a relevant design principle, with low heterogeneity
2	An individual randomised controlled study comparing the impact on real decisions or actions of a system designed according to a design principle or theory vs. a system not designed according to that principle
3a	Study comparing the safety or accuracy of a system based on the design principle against one not based on that principle, using real patient data
3b	Laboratory studies of simulated decisions or actions in response to a system based on the design principle vs. one not based on the principle, using real or simulated patient data
4	Untested theories or expert advice about what works in system design Anecdotes and case studies (“It worked for me”)

2.2 How will this development change our evaluation methods?

Evaluation can be defined as carrying out studies to generate information to guide future decisions [9, chapter 1]. However, while all studies conform to this generic definition, from my observations over 35 years there are at least five different motives for conducting studies. These motives, along with some typical questions addressed by each type of study, are listed in table 3.

Table 3. Types of evaluation study.

Study type	Motive for carrying out study	Typical questions
1. Formative evaluation	How to improve an information system?	Is it accurate? Is it safe? Will people use it? How to improve it?
2. Summative evaluation	Can the finished system solve a specific problem?	Does this system work? How much does it cost? Will people use it?
3. Defensive evaluation	Was the funders' money spent well without making the situation worse?	Has anything improved since the system was implemented?
4. Self-interested evaluation	Can this study help the evaluator build their own CV?	Will this study have an impact on my colleagues?
5. Principle-based evaluation	Can this generic principle contribute to system design and EBHI?	Does this general design principle make systems more usable, effective, safer, less expensive, or more maintainable?

While the first four types of evaluation are relatively well known, the next section explains what we mean by the fifth.

2.3 How to design and conduct “principle-based evaluation”?

Principle-based evaluation means designing and conducting studies to test a generic design principle that if true, can guide future system development or implementation, thus helping to build the EBHI evidence base. Figure 2 below shows the steps that principle-based evaluation requires.

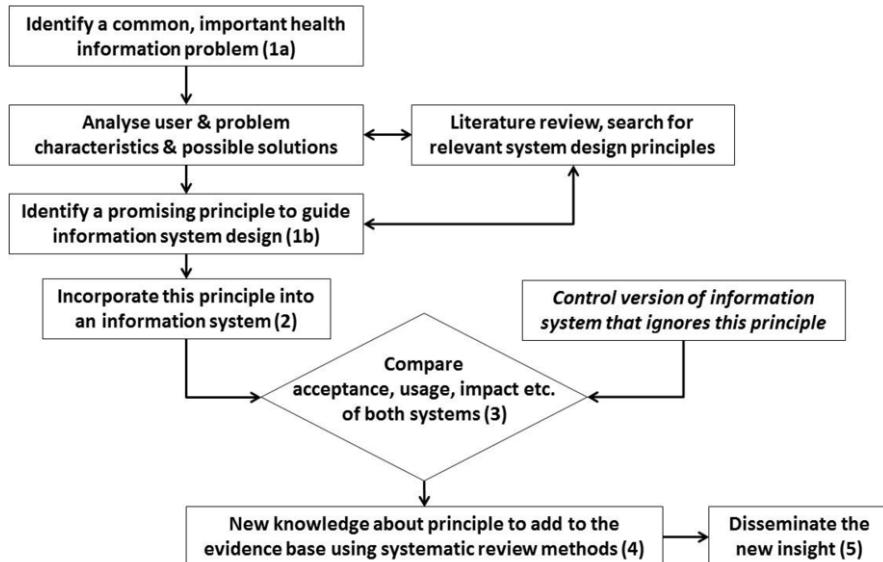


Figure 2. The process of designing and conducting principle-based evaluation studies (numbers refer to the text below).

If a researcher is planning a design principle-based evaluation study, they will need to carry out the following steps:

1. Working from a careful analysis of a common important problem (1a) in our domain (such as alert fatigue or poor data quality), identify a plausible generic system design principle or theory that may help resolve this (1b)
2. Use the selected principle to improve an existing information system, taking care that the only difference between the two systems is associated with application of the principle, not e.g. incorporating extra data or changing system usability (unless the principle concerns these specific actions)
3. Design one or more studies that rigorously test whether the design principle is supported or not, in terms of system acceptability, usability, accuracy, safety or impact on user decisions, actions or behaviours; or system maintainability or cost
4. If the study was small, integrate the results into the global evidence base of similar studies, using the well-established systematic review methods.
5. Accept and disseminate the results of their study, whatever these are – i.e. whether the principle makes sense to them or not. If the study was well designed, then its results should be respected.

There are some significant implications here for all evaluators. With the advent of EBHI, evaluators will need to think more clearly about their motives for carrying out a study and the consequences of this for their study design – particularly for the choice of controls. They will need to be clear about the differing aims of evaluation and their focus for each study. They will need to be familiar with a wide variety of evaluation methods, and how to identify and eliminate or control for biases and confounders [9, chapter 8]. They will also need to be aware of the obligation to publish their study

results (whether positive or negative), so that these are available to others aggregating evidence about what works and what does not.

2.4 Examples of studies and systematic reviews that contribute to EBHI

Some examples of studies that illustrate this approach and can potentially contribute to the health informatics evidence base are listed in table 4.

Table 4. Example studies and systematic reviews evaluating system design principles, in order of evidence grade.

Question	Type of study	Results	Source	Evidence grade (see Table 2) & comments
How to improve data quality?	Systematic review of 12 (mostly before-after) studies of various strategies in UK primary care	Most strategies appeared to have a positive effect, but study quality poor	Brouwer et.al. 2006 [2]	Evidence grade 1a. But systematic review was limited by poor study designs.
Does the use of psychological theory make a difference in behaviour change websites?	Systematic review and meta regression of 85 RCTs of theory based websites for health behaviour change	Use of theory to design website or recruit participants improved effectiveness by about one third of a standard deviation	Webb et.al. 2010 [10]	Evidence grade 1b. Use of theory may be confounded with better quality website design.
How much of a difference does tailoring and targeting make to text message impact?	Systematic review and meta regression of 19 RCTs of tailored SMS interventions for health behaviour change	Use of tailoring and targeting improves intervention effectiveness by 0.44 of a standard deviation	Head et.al. 2013 [11]	Evidence grade 1b. Use of tailoring may be confounded with better quality text design.
How to improve diagnostic accuracy?	RCT of a checklist	A well designed disease specific checklist improves accuracy by 10%	Adams et.al. 1986 [12]	Evidence grade 2. May reflect limited accuracy of junior doctors.
Can Fogg's principles of Persuasive computing improve websites for health-related decisions?	Online RCT of two websites to encourage 900 students to join NHS organ donation register	No – no difference (38% in both groups)	Nind et.al. 2009 [13]	Evidence grade 2. May only generalise to significant decisions such as organ donation.
Which kind of user interface speeds up data entry?	Experiment with 15 clinicians each entering 63 medical findings from 3 simulated cases using alternative prototype pen based user interfaces	Paged interface 5 seconds faster than scrolling. Complete list of codes 4 seconds faster than patient-specific list. Fixed position on screen 2 seconds faster than variable position.	Poon et.al. 1996 [14]	Evidence grade 3b. Limited to pen-based interfaces?
Can non-interruptive advice reduce	Within-subject experiment measuring	Prescribing alert in modal dialogue box twice as effective as	Scott et.al. 2011 [15]	Evidence grade 3b. Only tested one alert at a

errors?	prescribing errors in 20 junior doctors using case scenarios	same alert on ePrescribing system interface, but less acceptable.		time.
---------	--	---	--	-------

3. What kinds of design principle or theory to test?

So, what makes a good design principle to test? One aspect of this question is, *from which discipline or area* are promising design principles or theories likely to originate? To answer this question, consider a worked example: the design decisions faced by a team developing a typical eHealth system: an online forum to promote smoking cessation. Table 5 lists some of the fundamental decisions they need to take, together with possible disciplines or academic areas which could provide relevant design principles.

Table 5. Some design decisions made during the development of a sample information system, and possible origins of relevant design principles.

Design question / task	Discipline or area from which relevant design principles can originate
How to brand the website, and how to publicise it?	Marketing, public relations
What content to place on the website to attract smokers willing to quit?	Material to promote any of the techniques in the Behaviour Change Taxonomy [16]
How to encourage site visitors to enter, locate and retrieve information relevant to stopping smoking?	Search techniques; what makes risk / health information relevant Communication theory – common ground, etc. [17]
How to present information on the website in a manner that influences user decisions to quit?	Information design [18] Risk communication [19] Human decision making: heuristics and biases [20]
How to maximise the chances that a one-off user decision to stop smoking becomes a long term behaviour change?	Techniques drawn from the Behaviour Change taxonomy [16]

Another aspect of the big question is, what *kinds of design principle* are useful to test? Some properties of a candidate design principle that make a rigorous test valuable include that the design principle is:

- *Specific*: Sufficiently well formulated to be testable.
- *Actionable*: If proven, it would practically influence the design of health information systems.
- *Generic*: Can be applied across a range of information systems, user groups or contexts.
- *Credible*: The design principle appears well founded, so if proven is likely to be applied by others.
- *Enduring*: Such as theories about how people interact with and respond to information (eg. Risk perception), not theories about fleeting generations of technology (eg. High resolution vs. medium resolution virtual reality).
- *Novel or untested*: Not previously well tested for its impact on health information system design.

4. Some challenges arising from adopting the EBHI approach

Of course, the approach advocated will not solve every health informatics problem, such as use of the label “health informatics” by epidemiologists who can then attract funding intended for our discipline. As described earlier, we still need to conduct a wide range of evaluation studies more rigorously [9]. Also, we still need exploratory studies and experts to help us formulate plausible, generic design principles or theories for rigorous testing.

Another concern is that we should not over interpret the results of any individual study, as study results always vary randomly around the true effect size. So, we need to build evidence-based system design guidelines using systematic reviews⁵, not on individual studies, unless we see “mega-trials” in our discipline as we see in cardiology, which in the current climate of health informatics evaluation scepticism seems highly unlikely. While it is tempting to use the systematic review method to compare the effectiveness of systems that do and do not incorporate a design principle from separate studies, caution is needed – which is why we consider such reviews as grade 1b evidence in Table 2. Using meta regression to test a design principle is not rigorous – all it shows is that there is an **association** between the principle and the outcome, not **causation**. To show causality, we need a direct randomised head-to-head comparison of the effectiveness of systems that did and did not incorporate the design principle in a single study (grade 2 evidence)⁶, or ideally, a systematic review of head-to-head studies, which provides grade 1a evidence.

One dilemma is that while many design principles are generic (e.g. Schneiderman’s user interface design guidelines [21]), some other principles (e.g. how to format displays of clinical data or alerts) may be bound up in the context of the specific users, data items or the task they support. The concept of ecological user interface design supports this: for each work domain or environment we design a user interface that supports this, with all the relevant information formatted in the optimum way to support the task in hand [22]. Realist approaches to evaluation and realist synthesis may have a place here to uncover what works, when, for whom and why [23]⁷.

5. Conclusions

In my opinion, the advent of principle-based EBHI marks the beginning of an exciting and fundamentally new approach to our discipline that, over time, will yield the evidence we need to place our discipline on a firmer base. It will allow us to authoritatively answer core questions fundamental to our discipline, such as “How to improve data quality?” with which this contribution started. This will bring greater confidence to our discipline and assure its ability to deliver safe, effective clinical

⁵ See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁶ See also: C.R. Weir, Ensuring the Quality of Evidence: Using the best design to answer health IT questions, in: *ibid*.

⁷ See also: T. Otte Trojel et al., Going beyond systematic reviews: Realist and meta-narratives reviews, in: *ibid*.

information systems. However, it does mean more emphasis on rigorous study design and systematic reviews to identify and test potentially useful generic system design principles.

The benefits of a sound evidence base of system design principles or theories will include:

- The systems we produce will be reliably safe, efficient & predictable (like bridges).
- eHealth will evolve from an *intuitive craft* reliant on experts and apprenticeship into a *professional discipline*, making its decisions based on tested principles [6].
- There will be much less need for trial and error, or for re-invention of ad hoc systems that “seemed sensible at the time”.
- Aspirational drives to ‘modernise’ or ‘automate’, followed by searches for available systems, will be considered inappropriate; instead there will be a call to grasp the proven benefits of validated systems.
- There will be no need to evaluate **every** version of **every** app, website, serious game etc., as long as the original one was built using tested principles, and the users or context of use have not changed too much to render these principles invalid.

A final comment is that to avoid what Grémy called “The idolatry of technology” (personal communication, Francois Grémy, 1999), health informatics should focus on science rather than on computer artefacts [18]. However, whenever we talk about science, we must also be wary of pseudoscience [24]. Fortunately, pseudoscience can be distinguished from science by the fact that scientific theories can be tested and *disproved*, rather than confirmed [25]. So, health informatics professionals should avoid vague theories that cannot be tested, but also recognise that we will never know the limits of our new design principles until they fail us. However, meanwhile these design principles and theories will provide constructive new knowledge to inform future system design.

Recommended further readings

1. S.E. Straus, P. Glasziou, W.S. Richardson, R.B. Haynes, *Evidence-Based Medicine: How to Practice and Teach It*, 4th Edition, Churchill Livingston, Edinburgh, 2011.
2. C. Friedman, J. Wyatt, *Evaluation methods in biomedical informatics*, 2nd edition, Springer, New York, 2005.
3. C. Friedman, Where's the science in medical informatics? *J Am Med Inform Assoc* **2**(1) (1995), 65-7.
4. H.A. Heathfield, J. Wyatt, The road to professionalism in medical informatics: a proposal for debate, *Methods Inf Med* **34**(5) (1995), 426-33.
5. J. Wyatt, Medical informatics, artefacts or science? *Methods Inf Med* **35**(3) (1996), 197-200.

Food for thought

1. What are some disadvantages of the evidence-based approach to a scientific discipline?
2. Clinicians tend to consider clinical and cost effectiveness as the key evaluation criteria for a health technology. What alternative metrics might a computer scientist or a public health physician wish to consider, to help broaden the EBHI knowledge base?
3. How might a specific system design principle improve effectiveness while worsening system maintainability or widening health inequalities, for example? How do we manage those trade-offs?
4. Will health informatics as a discipline ever amass sufficient evidence-based design principles to allow us to develop and implement information systems with no need to carry out laboratory or field studies of safety and effectiveness?

References

- [1] J.C. Wyatt, F. Sullivan, ABC of Health Informatics 1: What is health information? *BMJ* **331** (2005), 566-8.
- [2] H.J. Brouwer, P.J.E. Bindels, H.C. Van Weert, Data quality improvement in general practice, *Fam Practice* **23** (2006), 529-36.
- [3] D.L. Sackett, W.M. Rosenberg, J.A. Gray, R.B. Haynes, W.S. Richardson, Evidence based medicine: what it is and what it isn't, *BMJ* **312**(7023) (1996), 71-2.
- [4] J. Wyatt. Use and sources of medical knowledge. *Lancet* **338** (1991), 1368-1373.
- [5] B. Middleton, M. Bloomrosen, M.A. Dente, B. Hashmat, R. Koppel, J.M. Overhage, T.H. Payne, S.T. Rosenbloom, C. Weaver, J. Zhang, American Medical Informatics Association: Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA, *J Am Med Inform Assoc* **20**(e1) (2013), e2-8.
- [6] H.A. Heathfield, J. Wyatt, The road to professionalism in medical informatics: a proposal for debate, *Methods Inf Med* **34**(5) (1995), 426-33.
- [7] D.L. Sackett, J.E. Wennberg, Choosing the best research design for each question, *BMJ* **315**(7123) (1997), 1636.
- [8] D.L. Sackett, Screening in family practice: prevention, levels of evidence, and the pitfalls of common sense, *J Fam Pract* **24**(3) (1987), 233-4.
- [9] C. Friedman, J. Wyatt, *Evaluation methods in biomedical informatics*, 2nd edition, Springer, New York, 2005.
- [10] T.L. Webb, J. Joseph, L. Yardley, S. Michie, Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy, *J Med Internet Res* **12**(1) (2010), e4.
- [11] K.J. Head, S.M. Noar, N.T. Iannarino, N.G. Harrington, Efficacy of text messaging-based interventions for health promotion: a meta-analysis, *Soc Sci Med* **97** (2013), 41-8.
- [12] I.D. Adams, M. Chan, P.C. Clifford, W.M. Cooke, V. Dallos, F.T. de Dombal, M.H. Edwards, D.M. Hancock, D.J. Hewett, N. McIntyre, et al., Computer aided diagnosis of acute abdominal pain: a multicentre study, *Br Med J (Clin Res Ed)* **293**(6550) (1986), 800-4.
- [13] T. Nind, J. Wyatt, I. Ricketts, P. McPate, J. Liu, *Website credibility and intervention effectiveness*, in Proceedings of the Persuasive Technology and Digital Behaviour Intervention Symposium: a symposium at the AISB 2009 Convention (6-9 April 2009) Heriot-Watt University, Edinburgh, Scotland, Society for the Study of Artificial Intelligence and Simulation of Behaviour, 2009, pp. 36-39.
- [14] A.D. Poon, L.M. Fagan, E.H. Shortliffe, The PEN-Ivory project: exploring user-interface design for the selection of items from large controlled vocabularies of medicine, *J Am Med Inform Assoc* **3**(2) (1996), 168-83.
- [15] G.P. Scott, P. Shah, J.C. Wyatt, B. Makubate, F.W. Cross, Making electronic prescribing alerts more effective: scenario-based experimental study in junior doctors, *J Am Med Inform Assoc* **18**(6) (2011), 789-98.

- [16] S. Michie, N. Hyder, A. Walia, R. West, Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation, *Addict Behav* **36**(4) (2011), 315-9.
- [17] E. Coiera, When conversation is better than computation, *J Am Med Inform Assoc* **7**(3) (2000), 277-86.
- [18] J.C. Wyatt, P. Wright, Medical Records 1: Design should help use of patient data, *Lancet* **352** (1998), 1375-8.
- [19] G. Gigerenzer, A. Edwards, Simple tools for understanding risks: from innumeracy to insight, *BMJ* **327**(7417) (2003), 741-4.
- [20] A. Tversky, D. Kahneman, The framing of decisions and the psychology of choice, *Science* **211** (4481) (1981), 453-8.
- [21] B. Schneiderman, C. Plaisant, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 5th edition, Pearson press 2009.
- [22] *Ecological interface design*, https://en.wikipedia.org/wiki/Ecological_interface_design, last access 11 February 2016.
- [23] R. Pawson, J. Greenhalgh, C. Brennan, E. Glidewell, Do reviews of healthcare interventions teach us how to improve healthcare systems? *Soc Sci Med* **114** (2014), 129-37.
- [24] S.O. Hansson, "Science and Pseudo-Science", *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/spr2015/entries/pseudo-science>, last access 11 February 2016.
- [25] J. Wyatt, Medical informatics, artefacts or science? *Methods Inf Med* **35**(3) (1996), 197-200.

Health IT for Patient Safety and Improving the Safety of Health IT

Farah MAGRABI^{a1}, Mei-Sing ONG^{a,b} and Enrico COIERA^a

^a*Centre for Health Informatics, Australian Institute of Health Innovation,*

Macquarie University, Sydney, Australia

^b*Computational Health Informatics Program, Boston Children's Hospital, Boston, MA,
USA*

Abstract. Alongside their benefits health IT applications can pose new risks to patient safety. Problems with IT have been linked to many different types of clinical errors including prescribing and administration of medications; as well as wrong-patient, wrong-site errors, and delays in procedures. There is also growing concern about the risks of data breach and cyber-security. IT-related clinical errors have their origins in processes undertaken to design, build, implement and use software systems in a broader sociotechnical context. Safety can be improved with greater standardization of clinical software and by improving the quality of processes at different points in the technology life cycle, spanning design, build, implementation and use in clinical settings. Oversight processes can be set up at a regional or national level to ensure that clinical software systems meet specific standards. Certification and regulation are two mechanisms to improve oversight. In the absence of clear standards, guidelines are useful to promote safe design and implementation practices. Processes to identify and mitigate hazards can be formalised via a safety management system. Minimizing new patient safety risks is critical to realizing the benefits of IT.

Keywords. Medical informatics, patient safety, medical errors.

1. Introduction

IT systems are integral to healthcare delivery and have a tremendous potential to bring about an overall improvement to patient safety. IT broadly includes all computer software used by health professionals and patients to support care [1]. At the same time, use of IT, just like any other technology, can introduce new, often unforeseen, errors that can affect care delivery and can lead to patient harm. It is now widely recognized that problems with IT and their use can pose risks to patient safety.

The objective of this contribution is to provide a motivation for evidence-based health informatics to improve patient safety, and to minimise the risks of harm associated with IT. The contribution begins with a broad-based review of the impact of IT on patient safety. We then turn our attention to the current evidence about patient harms. The next section examines the underlying causes of errors associated with IT and the final section looks at the types of safety strategies that need to be applied

¹ Corresponding author: Associate Professor Dr. Farah Magrabi, Centre for Health Informatics, Australian Institute of Health Innovation Level 6, 75 Talavera Road, Macquarie University, NSW 2109, Australia, farah.magrabi@mq.edu.au.

throughout the lifecycle of an IT system to improve safety. By understanding how problems with IT can give rise to clinical errors and having knowledge about their underlying causes, we can be better equipped to design, implement and use safer systems and to mitigate the risks of harm to patients.

2. Health IT can improve patient safety

Much of clinical care involves the gathering and synthesizing of information. In healthcare systems with increasing patient complexity and distribution of care, traditional paper-based information management is no longer adequate for supporting high patient care standards. Effective clinical decision-making requires careful assimilation of patient information from multiple fragmented sources, and the integration of vast amounts of new scientific evidence into practice. Reliable and efficient care can often only be achieved with the use of IT [2]. IT can substantially improve the safety of care by improving information collation, sharing, and access.

IT systems like electronic health records (EHR) facilitate access to patient information in a distributed manner. Using an EHR, patient information such as diagnoses, medications, and test results can be consolidated into a single system that can be accessed at any time, in different localities, and by different team members. Wireless technology coupled with portable handheld devices allows clinicians to retrieve the most up-to-date patient information while on the move. This has the potential to significantly improve information sharing across the continuum of care, enhancing patient safety and coordination of care [3].

A study that assessed the quality of diabetes care showed 51% of patients at EHR sites, as compared with 7% of patients at paper-based sites, received care for diabetes that met the recommended standards of care [4]. Following the implementation of computerized handoff system, the number of patients missed on resident rounds was reduced by half [5], and the rate of preventable adverse events was also reduced [6]. The advantage of electronically enhancing the availability of medical data was perhaps most evident during the recent storms in the US, notably Hurricane Katrina in 2005 [7], and the Joplin tornado in 2011 (Box 1) [8]. In both storms, many medical paper records were lost. Health care providers who were supported by decentralized EHR systems were able to continue the provision of care during and after the storm, while patients from paper-based sites were left stranded without adequate care.

Box 1: A Tale of Two Cities

The 2005 Hurricane Katrina

In 2005, Hurricane Katrina caused severe destruction in New Orleans. In the chaos that ensued after this disaster, displaced individuals many of whom had chronic health conditions left their medications and medical records behind. Responding clinicians were challenged by the need to care for these patients without any knowledge of their medical history. Standing in stark contrast were Veteran Affairs providers in the same city, who were able to maintain uninterrupted care supported by nationwide access to comprehensive EHR systems.

The 2011 Joplin tornado

In 2011, a devastating tornado struck Joplin, killing 134 people. A hospital in Joplin, St John's Regional Medical Center, was severely damaged, and medical paper records were lost. Three weeks before the storm, the hospital had completed its transfer to an EHR system. Six days after the tornado, the hospital staff returned to work in a new temporary mobile medical unit. Because the full patient records were available through the EHR, medical staff was able to continue deliver care and identify displaced individuals.

3. Evidence about IT-related patient harms is mounting

While IT promises to improve the provision of care, as discussed before, it is important to note the unanticipated negative consequences of such systems. The extent of patient harm associated with IT is, however, hard to quantify, due to the lack of empirical data [9]. The “hold harmless” clauses that protect software vendors from lawsuits effectively limit the freedom to publicly raise questions about software errors [10]. Thus, many problems with IT remain hidden, and unresolved. Based on error rates in other industries, the US Agency for Healthcare Research and Quality estimates that if EHRs are fully adopted, they could be linked to at least 60,000 adverse events a year [11].

While we currently cannot ascertain the actual rate of adverse events associated with IT, a growing body of evidence elucidates the pervasiveness of IT-related problems. The largest source of evidence comes from incident reports voluntarily submitted by software vendors and clinical workers to governing bodies, both at national and local levels [12-15]. The US Food and Drug Administration (FDA) maintains a medical device incident reporting system, known as the Manufacturer and User Facility Device Experience (MAUDE) database. In 2010, 260 IT-related incident reports were submitted to the database, 44 of which were linked to patient injuries, and 6 deaths were reported [13]. The Australian Incident Management System (AIMS) is yet another national initiative for the surveillance of patient safety issues. Between 2003 and 2005, 117 IT-related incidents were submitted to AIMS [12]. While no deaths were reported, 38% of the incidents were associated with adverse consequences caused by delay in treatment and care. Since neither system was designed specifically for the surveillance of IT-related adverse events, it is very likely that they were under-reported.

At a local level, the Pennsylvania Patient Safety Authority received 3,099 reports from Pennsylvania hospitals on EHR-related problems, between the years 2004 and 2012 [16]. More than 2,700 incidents involved near misses and 15 involved patient harm. The report showed a stark rise in the number of IT-related incidents over the years. Of the 3,099 incidents reported over an eight-year period, 1,142 were filed in 2011, more than double the number in 2010. With the increased adoption of IT incentivized by the Affordable Care Act, the problem will only worsen.

Flaws in software design and system glitches accounted for many of the reported incidents. For example, poorly designed user interface obscured clinical data, causing clinicians to prescribe the wrong medications, and to send the wrong patients for a procedure; computer-network delays resulted in delay in treatment; dangerous doses of medications were given to patients due to ambiguous drop-down menus; orientation markers on CT images were reversed, causing a surgeon to operate on the wrong side of patient’s head. These seemingly simple errors, when occurred in a healthcare setting, could potentially cascade into serious life-threatening events.

The transition between paper-based and EHR records represents a risky period, as physicians often use both systems in tandem [16]. At Children’s Hospital of Pittsburgh, mortality rates increased after the implementation of an EHR system in 2002 [17]: During the 18 months following the EHR implementation, mortality rate increased to 6.6% in the 5 months after the system was installed, from 2.8% in the 13 months before. A separate study on CPOE systems showed that the rate of computer-related pediatric errors was 10 errors per 1000 patient-days, and the rate of serious computer-related pediatric errors was 3.6 errors per 1000 patient-days [18].

The incidence of IT-related medication errors has been explored in several other studies [19-23]. A report on 4,416 incidents submitted to the Dutch central reporting system showed that 16% of incidents were linked to IT [20]. Incorrect selection of medication is the leading cause of medication errors, followed by failure to enter prescription data in the CPOE. Two patients died as a result, and 20 patients were seriously harmed. Similar types of errors were observed in an observational study in an Australian hospital [23]. Of the 1,164 prescribing errors observed, 43% were caused by selection errors, 32% were due to failure to complete prescription task, and 21% were a result of editing errors.

Another unintended consequence arising from the digitalization of the medical records is the risk of data breach. The number of medical data breaches has increased dramatically in recent years. As of July 2012, there were 464 data breaches reported to the U.S. Department of Health and Human Services (HHS) since August 2009, involving more than 20 million patients – the most common forms of data breach were thefts, unauthorized access or disclosure, and data loss [24]. In the same year, a bi-annual survey of 250 U.S. healthcare organizations showed that 27% of respondents had at least one security breach over the past year, compared to 19% in 2010 and 13% in 2008 [25]. The rise in data breach incidents was largely due to the proliferation of laptops and mobile devices. The number of cases where data were compromised as a result of a lost or stolen device had doubled. Concerns about data security has prompted the HHS to update the Health Insurance Portability and Accountability Act (HIPAA) in 2013, to expand security protections required of health care providers that contract or subcontract with business associates to handle medical information [26]. Providers can be penalized up to \$1.5 million if the business associates do not comply.

Cyber-security is also a growing concern. In June 2013, the FDA issued a safety communication, warning medical device manufacturers and hospitals of the risk of cyber-security [27]. While the actual number of incidents is difficult to assess, news reports on cyber-attacks proliferate. In a recent case, research computers at Kaiser Permanente were infected with malicious software for more than two and a half years before being discovered, affecting in excess of 5,000 patients [28]. In another high profile case, the infamous hacker group, Anonymous, allegedly launched a cyber-attack against Boston Children's Hospital [29]. Such events can bring down IT systems, causing disruptions in care delivery. With increased interconnectedness of health care information systems, the potential for large-scale events due to cyber-attacks is real.

4. IT-related harms have their origin in system design, implementation and use

Processes undertaken to design, build, implement and use IT provide the fundamental system safety against errors [30]. As we have seen in the previous sections, patients are harmed when design issues cause systems to fail or behave in unexpected ways.

4.1. System design

A clinical system may behave in unexpected ways when the *system design does not reflect how it will be used*. When designers have a poor understanding of clinical work they will often make wrong assumptions about how a system will be used, the tasks it must support and the clinical workflow in which those tasks need to be executed. As a consequence the designed system will result in clinical tasks being missed or executed

incorrectly. Incomplete or wrong assumptions about the clinical tasks that a system must support are one of the most important sources of error. For instance an order entry system that does not support discontinuation and modification of orders is likely to cause medication errors [13]. Errors are also generated when there is a mismatch of the system with the mental model of users. An example is an EHR that did not represent weight in the unit of measure used by clinicians e.g. displaying weight in pounds instead of kilograms [13].

Safe use is also influenced by the system user interface. *Inadequate or poorly designed user interfaces* increase cognitive load causing clinicians to make errors in using systems (use errors) [31]². IT use is hampered by poor usability when systems are hard to learn, and do not allow users to complete tasks in an efficient manner. Ease of use is also affected when users cannot easily re-establish proficiency after a period of not using the system. An interface that results in severe use errors can be hazardous to patients. Consider the case of a prescribing system that requires users to scroll through a drop down menu with an excessive number of options that are counter-intuitively arranged. As a result of using this system a patient received an excessive dose of a medication [13]. Risks to patients are also increased when systems do not facilitate recovery from use errors. For example, an order entry system that does not allow clinicians to modify or cancel an order for a chemotherapy protocol once it is entered into the system [12].

Another design related issue is a *mismatch between the system model and actual clinical workflow* which can lead to errors in task execution [32, 33]. For instance, a nurse cannot review medication lists at the time of administration because the system is not accessible at the patient's bedside. Errors are also generated when system functions and the display of information do not account for the sequence in which clinical tasks are carried out. For example, prescribing decision support is ineffective in an order entry system that does not require users to complete allergy information before medications are entered because allergies cannot be checked if that information is not known by the system prior to the entry of orders. Another example is an order entry system that does not separate pre- and postoperative orders resulting in a wrong procedure being undertaken based on a preoperative order.

Software defects introduced during development also cause IT to behave in unexpected ways. Such defects will remain if software is not adequately tested. For instance, an EHR that allocates test results to the wrong patient due to a programming flaw that is exposed when the system processes large volumes of test results.

4.2. System implementation

Beyond system design, IT safety is influenced by sociotechnical variables of the clinical setting in which systems are used [34]. For instance installation of an order entry system in a hospital with a poor safety culture or an inadequate IT network might lead to new errors. Introduction of new technology into an organization, or system implementation, may involve a changeover from a paper-based to an electronic system or from an existing electronic system to a new one. This period is characterized by a high degree of sociotechnical change which can pose safety risks when the transition to

² See also: R. Marcilly et al., From usability engineering to evidence-based usability in health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

new technology, changes to clinical workflows and, organizational policies and procedures are not effectively managed [35]. Creation of a hybrid paper and electronic records system due to partial system implementation has also been shown to create new opportunities for error [36]. Any changes to an IT system post-implementation such as updates to software or installation of new hardware can similarly be a threat [37]. Conversely, failure to update software in a timely manner can also pose a risk. For example, a new guideline may not be updated in an operational EHR.

Errors can arise from unexpected *interactions between system modules or with other systems* [12]. IT systems are usually composed of multiple modules and they seldom operate in isolation. For instance, an ambulatory care system will contain modules for record keeping, prescribing and ordering tests. The system could also be connected to a medical device such as a spirometer and other systems like a laboratory information system to download test results. Errors can arise from communication failures between system modules and other systems. For example, images from a full body x-ray of a child were lost when they were transferred from the x-ray machine to a PACS (picture archiving and communication system) [38]. And the x-ray needed to be repeated to acquire the missing images, re-exposing the child to high levels of radiation.

The supporting IT infrastructure including computer hardware, software, networks and data storage facilities are critical to safe implementation and operation. Analysis of US and Australian data indicates that technical failure is a major contributor to IT incidents [12, 13]. Ninety-six percent of the problems reported to the FDA were associated with technical failure [13]. Problems with the IT infrastructure that hosts software affect safety because poor availability of systems disrupts delivery of care to patients. For example, when their desktop computer or printer fails, a primary care physician cannot access the EHR in their consultation room or provide a prescription to the patient. Another example relates to a network problem in a hospital that caused a PACS to be inaccessible for 6 hours making it impossible to read or create records while the system was unavailable [38]. As a result procedures were cancelled and clinics were rescheduled. Failure of back up facilities and computer viruses can similarly disrupt care delivery.

4.3. System use

Safe IT use is a product of the system and the environment in which it is used. When system use is compromised by human factors which include environmental influences like the structural, cultural and policy related characteristics of an organization, risks to patients are increased [39].

The *knowledge and skills* of users are fundamental to safe use of IT³. Training programs are thus essential and need to be appropriately tailored to the needs of different clinical seniorities and roles to ensure safe operation of systems. For example, training for a prescribing system that will be used by physicians, pharmacists and nurses will need to be tailored to the needs of each group respectively. Equally when users are unaware of system limitations, errors of omission will be generated [40]. For instance, a clinician may inadvertently prescribe the wrong medication wrongly assuming that the system will alert them about any drug interactions [41, 42]. Errors

³ See also: E. Hovenga et al., Learning, training and teaching of health IT and its evidence for informaticians and clinical practice, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

can also be generated when *cognitive resources* devoted to using a system are inadequate. A clinician's workload plus environmental influences like distractions and interruptions can lead to errors [43]. For example, when interrupted by a phone call a physician wrote a prescription for the wrong patient because they returned to the wrong record at the end of the call [37].

Deficiencies in *organizational policies and procedures* for system use are another threat. As we have already discussed, training is critical to safe operation of IT. However the lack of a policy or a failure to enforce the requirement to complete training may result in untrained clinicians accessing systems. Thus an organization might create a procedure for new staff to complete mandatory training and then receive access to systems in a timely manner. Policies that govern system access directly impact safety as lack of access to systems or critical information can potentially delay care increasing risks to patients. For example, an attending physician was unable to access critical test results from a previous hospital admission because the results of an HIV test were only visible to the ordering physician due to privacy considerations [38].

Thus we have seen that the safety of IT is an emergent property of the broader sociotechnical system. As safety is an emergent system property it needs to be addressed throughout the lifecycle of IT systems including design, build, implementation and use [44]. All the possible interactions among system components are not predictable at design, especially when IT systems are used in context of a broader sociotechnical system⁴. In large complex systems, safety problems or *hazards* tend to emerge from unexpected interactions between system components and human users. There is potential for unsafe interactions when IT systems are integrated with local clinical workflows including other technology and the organizational structure. Therefore safety should also be addressed during and after the implementation of systems.

5. Safety management covers the IT lifecycle

Strategies to improve the safety of health IT can be formalised. The overall set of processes used to identify and mitigate hazards throughout the life cycle of a system is called a safety management system, and these have evolved in other high-risk industries like aviation.[45]. For example, England has a safety management program for health IT [46]. Such programs formalize and document hazard assessment and mitigation so that system safety can be independently verified. A range of hazard assessment techniques can be applied at different points in the system life cycle [44]. The documents that set out the evidence for how hazards have been identified and managed are called a *safety case* [47]. For instance, a manufacturer is required to create a safety case when deploying a new EHR. The safety case will be continuously updated with new hazards identified during deployment or when changes are made to the system.

Standardization via guidelines or mandatory standards, and operational oversight via certification, regulation or surveillance are the two main governance approaches that are relevant to improving the safety of health IT [30, 48]:

⁴ See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

Standards: Many international technical standards can be applied to clinical software and to the quality of processes at different points in the system life cycle, spanning design, build, implementation and use in a clinical setting. However, few standards directly address the safety of clinical systems [49]. England's safety management program has implemented two standards for managing clinical risks in the design, implementation and use of health IT [47, 50]. These standards are consistent with those for safety critical software (e.g. International Electrotechnical Commission IEC 61508) and medical devices (e.g. International Organisation for Standardisation ISO 14971), and were formally adopted as NHS standards in 2009.

Guidelines: In the absence of clear standards, looser guidelines can still offer a mechanism to promote safe design and implementation practices. For instance, a guideline can be used to provide recommendations for the safe display of patient information within an EHR based upon usability principles. The US National Institute of Standards and Technology (NIST) has published a guide to evaluate EHR usability [51]. The NIST guide proposes formative usability evaluation by experts and summative testing in the hands of users incorporating a risk-based approach to examining usability problems.⁵ Guidelines can similarly be applied to system implementation and use. The Australian guidelines for implementing medication systems in hospital are one example [52]. Another example is the Safety Assurance Factors for EHR Resilience (SAFER) guides sponsored by the US Office of the National Coordinator for Health IT [53]. Such guidelines are generally directed at manufacturers and healthcare organizations to assess the safety of clinical software systems as they are used in clinical setting.

Certification: Oversight process can be set up at a national or regional level to ensure that clinical software systems meet specific standards. Certification provides independent assurance that software is fit for purpose and that it meets specific requirements for functionality, interoperability and security. For instance, the manufacturer of a prescribing system may be required to show that their system provides certain core clinical functions, that it is secure and that it can be integrated with other information systems such as the EHR. Safety is addressed alongside interoperability in the Australian certification program but it is not explicitly addressed in the US and Canadian programs, though conformance with functionality, usability, interoperability, security and privacy requirements may lead to safer systems [49].

Regulation: Certification can be voluntary, and it requires regulation to compel manufacturers to comply with standards or performance targets [54]. Regulation ensures that manufacturers comply with legal requirements for software to be designed and built in a manner that its use does not compromise patient safety. For example, a manufacturer may need to submit a safety case that demonstrates that its equipment is safe for use in a clinical setting before it is allowed to deploy the system. Although standalone software has largely been outside the strict regulatory regimen applied to medical devices, current initiatives indicate a gradual move towards regulation. Existing regulatory regimens for medical devices such as the US FDA process and the CE mark in Europe provide a template for the regulation of clinical software. In Europe the safety of medical devices is regulated through a directive that focuses on manufacturing and pre-market testing leading to a declaration of conformity. In general,

⁵ See also: R. Marcilly et al., From Usability Engineering to Evidence-based Usability in health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

the level of oversight or regulatory control should be proportional to the degree of risk that an information system poses to patients [1].

Surveillance of emerging safety issues: Beyond the stages of design and implementation, effective surveillance mechanisms are required to track any emergent safety problems associated with routine use of IT. The monitoring of incidents is central to detecting emerging problems in mainstream patient safety programs which are now well-established in most developed nations [55]. While it is mandatory to report incidents associated with regulated software, the reporting of general patient safety incidents (including those involving most health IT) is voluntary. One large-scale program directed at monitoring and responding to IT incidents reported by healthcare organizations and manufacturers is part of the England's safety management program which has been in place since 2005 [46, 49]. Yet, as we have seen before, IT incidents are being reported amongst general patient safety incidents and alongside reports of medical device failure and hazards. One source of such reports is the US FDA's MAUDE. Although the FDA does not enforce its regulatory requirements with respect to IT, some manufacturers have voluntarily listed their systems and reported incidents [13]. To facilitate the reporting of such incidents the US AHRQ has developed a new standard called a "common format" and a software tool to support detection and management of IT-related hazards [56]. The Health IT Hazard Manager facilitates the characterisation and communication of hazards along with their actual and potential adverse effects to support learning within healthcare organisations, across organisations using the same software and, by manufacturers and policymakers [11].

6. Conclusion

IT systems can enhance patient safety by improving access to information and by providing decision support, but problems with IT can pose new risks. Minimizing these risks is critical to realizing the benefits of IT. The risks of data breach and cyber-crime are also important concerns. We have seen that safety is an emergent property of the broader sociotechnical system in which IT is used, and errors arise from processes to design, build, implement, and use IT. Thus a holistic system approach that addresses IT errors at different points the system lifecycle is needed. In addition to greater standardisation and oversight to ensure safe system design and build, appropriate implementation and use of IT is critical to bring about overall improvements to patient safety. The effectiveness of specific strategies to address risks is not known and further research is required to evaluate their impact for a more evidence-based approach to IT safety. There is also a need for greater transparency where manufacturer contracts are governed by commitment to patient safety; and a balanced risk avoidance and safety promoting culture across the clinical process and IT support spectrum. When IT is used and governed responsibly it can improve patient safety, but it is not a panacea for managing safety risks. Taking a 'blind eye' approach to the risks of health IT can only lead to new forms of avoidable patient harm.

Recommended further readings

1. E. Coiera, F. Magrabi, *Information system safety*, Guide to Health Informatics, CRC Press, Taylor & Francis Group, Boca Raton, FL, 2015, pp. 195-220.

2. Institute of Medicine, *Health IT and Patient Safety: Building Safer Systems for Better Care*, The National Academies Press, Washington, DC, 2012, <http://www.nap.edu/catalog/13269/health-it-and-patient-safety-building-safer-systems-for-better>, last access 11 February 2016.
3. N.G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. Massachusetts Institute of Technology, 2011, <https://mitpress.mit.edu/books/engineering-safer-world>, last access 11 February 2016.
4. W. Runciman, P. Hibbert, R. Thomson, T. van der Schaaf, H. Sherman, P. Lewalle, Towards an International Classification for Patient Safety: key concepts and terms, *Int J Qual Health Care* **21**(1) (2009), 18-26.
5. W.B. Runciman, J.A. Williamson, A. Deakin, K.A. Benveniste, K. Bannon, P.D. Hibbert, An integrated framework for safety, quality and risk management: an information and incident management system based on a universal patient safety classification, *Qual Saf Health Care* **15** Suppl 1 (2006), i82-90.

Food for thought

1. Discuss the safety benefits and risks of a hybrid paper and electronic records system in a hospital setting.
2. Examine the patient safety risks of implementing a personal health record on a large-scale e.g. for a hospital, healthcare system or nationally. Hint: Think about how individual incidents can harm or increase the risk of harm to numerous patients when an IT system is used at different scales.
3. Why is a highly usable EHR not necessarily safe?
4. Why is an EHR built on the best clinical principles not necessarily safe?
5. What are some ways to improve surveillance of IT-related safety issues? Discuss the role of automated techniques.
6. Discuss the advantages and disadvantages of regulating all clinical software systems as medical devices.

References

- [1] E. Coiera, *The guide to health informatics*, third edition, CRC Press, 2015.
- [2] D.W. Bates, A.A. Gawande, Improving safety with information technology, *N Engl J Med* **348**(25) (2003), 2526-34.
- [3] C.M. DesRoches, E.G. Campbell, S.R. Rao, K. Donelan, T.G. Ferris, A. Jha, R. Kaushal, D.E. Levy, S. Rosenbaum, A.E. Shields, D. Blumenthal, Electronic health records in ambulatory care - a national survey of physicians, *N Engl J Med* **359**(1) (2008), 50-60.
- [4] R.D. Cebul, T.E. Love, A.K. Jain, C.J. Hebert, Electronic health records and quality of diabetes care, *N Engl J Med* **365**(9) (2011), 825-33.
- [5] E.G. Van Eaton, K.D. Horvath, W.B. Lober, A.J. Rossini, C.A. Pellegrini, A randomized, controlled trial evaluating the impact of a computerized rounding and sign-out system on continuity of care and resident work hours, *J Am Coll Surg* **200**(4) (2005), 538-45.
- [6] L.A. Petersen, T.A. Brennan, A.C. O'Neil, E.F. Cook, T.H. Lee, Does housestaff discontinuity of care increase the risk for preventable adverse events? *Ann Intern Med* **121**(11) (1994), 866-72.
- [7] S.H. Brown, L.F. Fischetti, G. Graham, J. Bates, A.E. Lancaster, D. McDaniel, J. Gillon, M. Darbe, R.M. Kolodner, Use of electronic health records in disaster response: the experience of Department of Veterans Affairs after Hurricane Katrina, *Am J Public Health* **97** Suppl 1 (2007), S136-41.
- [8] M. Abir, F. Mostashari, P. Atwal, N. Lurie, Electronic health records critical in the aftermath of disasters, *Prehosp Disaster Med* **27**(6) (2012), 620-2.

- [9] *Health IT and Patient Safety: Building Safer Systems for Better Care*. Washington, DC: The National Academies Press: Institute of Medicine; 2012.
- [10] R. Koppel, D. Kreda, Health care information technology vendors' "hold harmless" clause: implications for patients and clinicians, *JAMA* **301**(12) (2009), 1276-8.
- [11] J.M. Walker, A. Hassol, B. Bradshaw, M.E. Rezaee. *Health IT Hazard Manager Beta-Test: Final Report*. (Prepared by Abt Associates and Geisinger Health System, under Contract No. HHS-A290200600011i, #14), AHRQ Publication No. 12-0058-EF, Rockville, MD: Agency for Health care Research and Quality. May 2012.
- [12] F. Magrabi, M.S. Ong, W. Runciman, E. Coiera, An analysis of computer-related patient safety incidents to inform the development of a classification, *J Am Med Inform Assoc* **17**(6) (2010), 663-70.
- [13] F. Magrabi, M.S. Ong, W. Runciman, E. Coiera, Using FDA reports to inform a classification for health information technology safety problems, *J Am Med Inform Assoc* **19**(1) (2012), 45-53.
- [14] R.B. Myers, S.L. Jones, D.F. Sittig, Review of Reported Clinical Information System Adverse Events in US Food and Drug Administration Databases, *Appl Clin Inform* **2** (2011), 63-74.
- [15] D. Warm, P. Edwards, Classifying Health Information Technology patient safety related incidents - an approach used in Wales, *Appl Clin Inform* **3**(2) (2012), 248-57.
- [16] E. Sparnon, W.M. Marella, The role of the electronic health record in patient safety events, *Pennsylvania Patient Safety Authority* **9**(4) (2012), 1276-8.
- [17] Y.Y. Han, J.A. Carillo, S.T. Venkataraman, R.S. Clark, R.S. Watson, T.C. Nguyen, H. Bayir, R.A. Orr, Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system, *Pediatrics* **116**(6) (2005), 1506-12.
- [18] K.E. Walsh, W.G. Adams, H. Bauchner, R.J. Vinci, J.B. Chessare, M.R. Cooper, P.M. Hebert, E.G. Schainker, C.P. Landrigan, Medication errors related to computerized order entry for children, *Pediatrics* **118**(5) (2006), 1872-9.
- [19] J.S. Ash, D.F. Sittig, E.G. Poon, K. Guappone, E. Campbell, R.H. Dykstra, The extent and importance of unintended consequences related to computerized provider order entry, *J Am Med Inform Assoc* **14**(4) (2007), 415-23.
- [20] K.C. Cheung, W. van der Veen, M.L. Bouvy, M. Wensing, P.M. van den Bemt, P.A. de Smet, Classification of medication incidents associated with information technology, *J Am Med Inform Assoc* **21**(e1) (2013), e63-70.
- [21] R. Koppel, J.P. Metlay, A. Cohen, B. Abaluck, A.R. Localio, S.E. Kimmel, B.L. Strom, Role of computerized physician order entry systems in facilitating medication errors, *JAMA* **293**(10) (2005), 1197-203.
- [22] P. Littlejohns, J.C. Wyatt, L. Garvican, Evaluating computerised health information systems: hard lessons still to be learnt, *BMJ* **326**(7394) (2003), 860-3.
- [23] J.I. Westbrook, M.T. Baysari, L. Li, R. Burke, K.L. Richardson, R.O. Day, The safety of electronic prescribing: manifestations, mechanisms, and rates of system-related errors associated with two commercial systems in hospitals, *J Am Med Inform Assoc* **20**(6) (2013), 1159-67.
- [24] U.S. Dept. Health & Human Services, *Breaches Affecting 500 or More Individuals*, <https://ocrportal.hhs.gov/ocr/breach>, last access 11 February 2016.
- [25] Ponemon Institute LLC, *Fourth annual benchmark study on patient privacy & data security*. March 2014. <https://www2.idexpertscorp.com/ponemon-report-on-patient-privacy-data-security-incidents>, last access 11 February 2016.
- [26] *Modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules Under the Health Information Technology for Economic and Clinical Health Act and the Genetic Information Nondiscrimination Act, Other Modifications to the HIPAA Rules, Final Rule*, 78 Fed. Reg. 5566 (Jan. 25, 2013).
- [27] US Food and Drug Administration, Cybersecurity for medical devices and hospital networks: FDA safety communication, <http://www.fda.gov/medicaldevices/safety/alertsandnotices/ucm356423.htm>, last access 11 February 2016.
- [28] *Computer virus at heart of Kaiser data breach*, in: Government Health IT, 2014, April 7, <http://www.govhealthit.com/news/computer-virus-heart-kaiser-data-breach>, last access 11 February 2016.
- [29] M.B. Farrell, P. Wen, *Hacker group Anonymous targets Children's Hospital*, in: The Boston Globe; 2014, April 24.
- [30] J.M. Walker, P. Carayon, N. Leveson, R.A. Paulus, J. Tooker, H. Chin, A. Bothe, Jr., W.F. Stewart, EHR safety: the way forward to safe and effective systems, *J Am Med Inform Assoc* **15**(3) (2008), 272-7.
- [31] J. Nielsen, *Usability engineering*, Morgan Kaufmann, San Francisco, 1993.
- [32] L.L. Novak, R.J. Holden, S.H. Anders, J.Y. Hong, B.T. Karsh, Using a sociotechnical framework to understand adaptations in health IT implementation, *Int J Med Inform* **82**(12) (2013), e331-44.

- [33] D.S. Debono, D. Greenfield, J.F. Travaglia, J.C. Long, D. Black, J. Johnson, J. Braithwaite, Nurses' workarounds in acute healthcare settings: a scoping review, *BMC Health Services Research* **13** (2013), 175.
- [34] D.F. Sittig, H. Singh, Defining health information technology-related errors: new developments since to err is human, *Arch Intern Med* **171**(14) (2011), 1281-4.
- [35] B.T. Karsh, Beyond usability: designing effective technology implementation systems to promote patient safety, *Qual Saf Health Care* **13**(5) (2004), 388-94.
- [36] Spotlight on Electronic Health Record Errors: Paper or Electronic Hybrid Workflows, Pennsylvania Patient Safety Advisory **10**(2) (2013), 55-8.
- [37] F. Magrabi, T. Liaw, D. Arachi, W.B. Runciman, E. Coiera, M.R. Kidd, Identifying patient safety problems associated with information technology in general practice: an analysis of incident reports, *BMJ Qual Saf*, doi:10.1136/bmjqqs-2015-004323.
- [38] F. Magrabi, M. Baker, I. Sinha, M.S. Ong, S. Harrison, M.R. Kidd, W.B. Runciman, E. Coiera, Clinical safety of England's national programme for IT: A retrospective analysis of all reported safety events 2005 to 2011, *Int J Med Inform* **84**(3) (2015), 198-206.
- [39] B.T. Karsh, R.J. Holden, S.J. Alper, C.K. Or, A human factors engineering paradigm for patient safety: designing to support the performance of the healthcare professional, *Qual Saf Health Care* **15** Suppl 1 (2006), i59-65.
- [40] J. Horsky, G.J. Kuperman, V.L. Patel, Comprehensive analysis of a medication dosing error related to CPOE, *J Am Med Inform Assoc* **12**(4) (2005), 377-82.
- [41] K. Goddard, A. Roudsari, J.C. Wyatt, Automation bias: a systematic review of frequency, effect mediators, and mitigators, *J Am Med Inform Assoc* **19**(1) (2012), 121-7.
- [42] R. Parasuraman, D.H. Manzey, Complacency and bias in human use of automation: an attentional integration, *Hum Factors* **52**(3) (2010), 381-410.
- [43] S.Y. Li, F. Magrabi, E. Coiera, A systematic review of the psychological literature on interruption and its patient safety implications, *J Am Med Inform Assoc* **19**(1) (2012), 6-12.
- [44] N.G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*, Massachusetts Institute of Technology, USA, 2011.
- [45] *Safety Management Manual (SMM)*, 3rd edition, Montreal, Canada, International Civil Aviation Organisation (ICAO), 2012.
- [46] M. Baker, I. Harrison, M. Gray, Safer IT in a Safer NHS: account of a partnership, *The British Healthcare Computing & Information Management* **23**(7) (2006), 11-14.
- [47] *ISB 0129 Clinical Risk Management: its Application in the Manufacture of Health IT Systems*, <http://systems.hscic.gov.uk/clinsafety/dscn>, last access 11 February 2016.
- [48] H. Singh, D.C. Classen, D.F. Sittig, Creating an oversight infrastructure for electronic health record-related patient safety hazards, *J Patient Saf* **7**(4) (2011), 169-74.
- [49] F. Magrabi, J. Aarts, C. Nohr, M. Baker, S. Harrison, S. Pelayo, J. Talmon, D.F. Sittig, E. Coiera, A comparative review of patient safety initiatives for national health information technology, *Int J Med Inform* **82**(5) (2013), e139-48.
- [50] *ISB 0160 Clinical Risk Management: its Application in the Deployment and Use of Health IT Systems*, <http://systems.hscic.gov.uk/clinsafety/dscn>, last access 11 February 2016.
- [51] *NIST Technical Evaluation, Testing, and Validation of the Usability of Electronic Health Records*, US National Institute of Standards and Technology (NISTIR 7804), 2012, http://www.nist.gov/manuscript-publication-search.cfm?pub_id=909701, last access 11 February 2016.
- [52] Australian Commission on Safety and Quality in Health Care. *Electronic Medication Management Systems - A Guide to Safe Implementation*, 2nd edition, ACSQHC, 2012, <http://www.safetyandquality.gov.au/our-work/medication-safety/electronic-medication-management-systems/development-of-the-guide-to-safe-electronic-medication-management-in-hospitals>, last access 11 February 2016.
- [53] D.F. Sittig, J.S. Ash, H. Singh, The SAFER guides: empowering organizations to improve the safety and effectiveness of electronic health records, *Am J Manag Care* **20**(5) (2014), 418-23.
- [54] E. Coiera, J. Westbrook, J. Wyatt, The safety and quality of decision support systems, *Methods Inf Med* **45** Suppl 1 (2006), 20-5.
- [55] J.C. Pham, S. Giani, J. Battles, P. Beard, J.R. Clarke, H. Coates, L. Donaldson, N. Eldridge, M. Fletcher, C.A. Goeschel, E. Heitmiller, J. Hensen, E. Kelley, J. Loeb, W. Runciman, S. Sheridan, A.W. Wu, P.J. Pronovost, Establishing a global learning community for incident-reporting systems, *Qual Saf Health Care* **19**(5) (2010), 446-51.
- [56] Agency for Healthcare Research and Quality. *Common formats*. Rockville, Maryland: US Department of Health and Human Services, <http://www.pso.ahrq.gov/common>, last access 11 February 2016.

Part II: Methodological Considerations of Health IT Evaluation

This page intentionally left blank

Theoretical Basis of Health IT Evaluation

Jytte BRENDER McNAIR^{a,1}

^a*Dept. of Health Science & Technology, Aalborg University*

Abstract. The focus of this contribution is on the theoretical principles and concepts behind evaluation of IT-based systems, discussing their presuppositions, implications and interrelationships; for instance in relation to a series of issues to consider: terminology for the concepts used as that is a reason for many disputes, bias as that is a common reason for less accuracy and trustworthiness in conclusions, culture as the tacit driver of everything we do and design, constructive evaluation as this has strict time and timing issues, preparing for meta-analyses as that is in the near future, and top-level issues in choice of methodology. Awareness in these respects will lead to avoidance of major pitfalls and perils at evaluation and thereby improve the validity and trustworthiness of an evaluation outcome, supporting the initiative towards evidence-based health informatics.

Keywords. Evaluation, systems theory, IT system, bias.

1. Introduction

“Evaluation is the act of measuring or exploring properties of a health information system (in planning, in development, in implementation, or in operation), the result of which informs a decision to be made concerning that system in a specific context. Evaluation of health information systems has to deal with the actors, the artifacts, and their interaction to best support the decisions to be made.” [1].

Many tend to believe that evaluation is something everyone can do. It is indeed a common day activity in one's life and it appears so easy. Nevertheless, there is a difference between providing somewhat random or subjective evidence and providing measures based on systematic judgements, and one has to know which of the two extremes to apply when. When you buy a new car, would you do it only based on your own test-drive? Probably not, you would likely read the professional associations' assessment from test-driving, crash tests and more, and then use your own assessment to add a judgement on whether the car really is suited for your practical purpose, because there are always compromises to be made.

‘Providing evidence’ requires a stringent approach adopting the principles and methods used in science, and evaluation of health IT applications is a scientific discipline. The purpose of this book is to contribute to healthcare through the concept of ‘evidence-based health informatics’. Therefore, in order to secure maximum impact,

¹ Corresponding author: Emeritus associate professor Jytte Brender, Dept. of Health Science & Technology, Aalborg University, Fredrik Bajer's Vej 7D, DK-9220 Aalborg East, Denmark, jytte@brender.dk or jbr@hst.aau.dk.

the focus of this contribution shall be on those basic principles and concepts of evaluation that have a major impact on the validity and trustworthiness of an evaluation outcome. Thus, the aim of this contribution is to provide a scientifically-minded reader with the theoretical background for starting to design an evaluation study, and to show the non-scientific reader the importance and content of a structured objective approach.

The theoretical foundation of any scientific discipline is the philosophy of science, where the dictionary definitions of ‘philosophy’ includes two relevant for our purpose, both from [2]: “the academic discipline concerned with making explicit the nature and significance of ordinary and scientific beliefs and investigating the intelligibility of concepts by means of rational argument concerning their presuppositions, implications and interrelationships”; and “the critical study of the basic principles and concepts of a discipline.”. Similarly, we see a theory as “a set of hypotheses related by logical or mathematical arguments to explain and predict a wide variety of connected phenomena in general terms.” [2]. It is such arguments that this contribution will outline, arranged under the following headings:

- 1) Grounding possibilities: Theoretical assumptions and methodological considerations founding an evaluation study; value norms and *raison d'être*.
- 2) Communicative interactions: Matching scope, practical assumptions and delimitations.
- 3) Identifying and balancing the risk of bias in health IT evaluation.
- 4) Decision-making preferences: Culture is the driver of our decision-making whether we know it or not, and whether we want it or not.
- 5) Time and timing of evaluation: The concept of a constructive evaluation as opposed to traditional (summative) evaluations.
- 6) The next stage – that is, the indeterministic nature of systems development demands sustainability through flexibility and fluidity, and the demand for evidence enforce a next stage of methodological approaches.
- 7) Selecting/choosing the appropriate and sufficient methodology.

2. Grounding Possibilities: Theoretical Assumptions and Methodological Considerations

Which theoretical assumptions and methodological considerations can and should be the basis for an evaluation study? And when are they relevant to consider?

There are a set of interlinked concepts, like methodology, perspective and culture, for which a deep understanding will support the initiation of a successful planning and/or accomplishment of an evaluation study. In the following, these terms will be discussed briefly, also showing how important it is to always make one's terminology and values explicit.

“The term ‘methodology’ signifies “the science of methods” (BIPM et al. 1993) from the Greek ‘logos’, which means “the science of”. In functional terms it relates to the knowledge of how to prepare and use methods. Expressed in structural terms a methodology consists of “a coherent set of methods covering all the sub-tasks necessary for a given undertaking”. In other words, a methodology is supposed to a) provide the answer to what to do next, when to do what and how to do it, and b) to describe the ideas behind such choices and the suppositions (for instance the philosophical background) behind them.” ([3], p. 14).

Designing an evaluation study is in some ways like peeling an onion, because one answer brings out a next level with a number of new questions. For instance, before one can even start thinking of which methodology to choose one has to have a goal in order for the methodology to make the starting point and the end point meet – that is implicit from the structural definition above. At one such ‘deeper’ level before deciding on the methodology, one has to make clear one’s theoretical assumptions, for instance, a high-level reductionistic² versus a holistic perspective.

Such decisions are drivers towards the actual planning of an evaluation study. “The concept of ‘perspective’ stands for hidden aspects and assumptions deeply buried in the design and application of methods, see for instance (Mathiassen & Munk-Madsen 1986; Arbnor & Bjerke 1997; Brender 1997). In a generalized version, the perspective is the implicit assumptions of (cause-effect relations within) the object of study. So, the perspective is synonymous with “that aggregation of (conscious or unconscious, epistemological) assumptions of how things relate in combination with imprinted attitudes guiding our decision making e.g. in a problem solving situation.”” ([3] p. 18).

Few method designers are aware that our cultural background (professional, religious and national) maintains a series of tacit assumptions affecting our way of doing and perceiving things; see for instance [4] and [5]; as well as a brief overview in [6]. Caused by the tacit nature (i.e. completely unknown to inexperienced users of a given method or methodology), some perspectives may contain pitfalls, where the perspective of a method conflicts with the actual purpose which the method is intended to be used for. This is why this concept is so important in a profession-oriented context, here evaluation of health IT applications. A couple of simple examples of the implications of culture will illustrate this:

- In some Asian and African cultures it is highly impolite to answer a question with a ‘No’. Then think of many traditional questionnaires (which are typical evaluation instruments) or radio-buttons in the screen interactions between a computer and an end-user (i.e. relevant in a usability test). It doesn’t matter that the application as such isn’t situated in either of the cultures mentioned, because with today’s intensive migration of labour forces these cultures will be present everywhere, and adaptation to a new local culture is not something that comes overnight. Hence, such culture may unintentionally impose a bias in the evaluation outcome, or even worse may unintentionally compete with the design principles behind the screen functionality and falsify the input from a user-computer dialogue.
- In some cultures, a manager is considered the ultimate decision-maker (actually a decision-taker), he/she is always right, and the accuracy of his/her information or the appropriateness of his/her decision-making is never questioned by his/her employees. Then obviously, interviews and questionnaires involving end-users on the floor have a built-in risk of bias.

Clearly an evaluator has to manoeuvre within such local organisational context and conditions.

² In a reductionistic perspective one can observe a system’s components individually and then combine all such observations to provide a true and complete explanation of the whole system under investigation or of that part of the system which those components represent. In a holistic perspective, component parts of a system are dynamically interrelated and should be viewed within that wholeness.

Evaluation³ from a user perspective is always about an IT system operating within a context, and that context is the user organisation (i.e. another ‘system’). There are two definitions of ‘system’ that we find beneficial, and together they capture the essence of the concept of a ‘system’ for the present purpose. A system is:

“all the components, attributes and relationships needed to accomplish an objective” [7], p. 483 – that is, it has purpose, structure, behaviour and interactions, internally and across its boundary; this implicitness led in turn to the following definition, which however, is still not perfect:

“An organisation in which all structural components and dynamics are interrelational, participating internally, and affecting conditions externally” [8], p. 480. An ‘organisation’ here is to be understood in a business or administrative sense.

Both definitions emphasise the intricate relationship and hence dynamics between a system’s components and their properties. One can definitely use a reductionistic approach, and most methods applicable for evaluation are somewhat reductionistic; however, one has to be aware of the implications of the assumptions that one implicitly takes for given methods. It is always relevant to take this into consideration.

At an early point in time, make the policy and values as evaluator explicit; policy may be derived from one’s values. They show your view on what evaluation really is – that is, its role in a larger context and interrelations with components within the systems under evaluation. In systems development, the different development methodologies and methodologies for project management have implicit and/or explicit built-in values. This is not any different for evaluation. Further, policy and values constitute a commitment towards external parties such as the user organisation and potential sponsors of the evaluation, but they also dictate the trajectory for achieving the goal, and hence the choice of methods to be involved.

Examples of value statements for health IT evaluations are:

- “User involvement is essential.”
- “The users’ capacity, skills and responsibilities will be respected.”
- “Any interaction with users will take place on their premises, professionally, linguistically and otherwise.”

Examples of policy statements are:

- “(Constructive) evaluation is a dynamic (non-deterministic) process that obeys the information need of the decision makers and not unnecessarily hampers or delays ongoing processes within the user organisation or the development organisation.”
- “User involvement will be designed as specified tasks in manageable bites, in compliance with their respective managerial and professional competence, and will be continuously adjusted according to their respective relevance for the evaluation.”

³ Cf. the definition of evaluation, our perspective is the user’s as opposed to the developers’ debugging.

- “Even if there are theoretical considerations behind the practical tools and prescriptions, the user organisation shall not be unnecessarily distracted by being presented with this. That is, you will speak with the user organisation in their language and on their contextual premises.”

Given the definition of ‘evaluation’, its outcome is going to be used within some context. Users speak their profession-oriented language, so one should not enforce one’s own terminology upon them. Thus, the implication of for instance the last statement includes conditions for interacting with the users, including reports from the evaluations and perhaps even choice of evaluation method(s).

3. Communicative Interaction: Matching Scope, Practical Assumptions and Delimitations

Note how heavily this entire paper is filled with definitions. Definitions are indeed difficult to make, but when they are finally right they are extremely helpful instruments. Make the applied terminology clear before starting an evaluation study and then again when you wish to publish your study. Examples are the central terms ‘evaluation’, ‘verification’, ‘validation’ and ‘assessment’, where this author has witnessed so many heated disputes among colleagues – because of different use of the terms. Moreover, different domain professionals use the same terms as evaluators do, but in other contexts and with (slightly) different meanings; for instance, the term ‘phase’.

It is not necessary to make the definitions oneself, but find in the literature those that resonate with the study’ purpose and need. Being explicit about the terminology prevents a lot of miscommunication and misinterpretations, and may help harmonise the domain in which you are operating, and last but not least it contributes to securing ‘evidence-based health informatics’.

Then, make practical assumptions and delimitations explicit: Even when one aims at performing a scientific evaluation study in order to secure an appropriate level of evidence, there are toes and heels that have to be cut before Cinderella’s glass shoe will fit: available funding and local conditions, publication strategy/restrictions, confidentiality and personal data security, ... The level of ambition and the evaluation set-up have to match the local realities, and that requires a communicative interaction with and within the organisational context to identify the local conditions, a necessity for aligning with the practical reality.

While taking all of these issues into account, make the scope explicit: What is the question that the evaluation is going to answer? This is a top-level decision, the target of the entire study, and that which the methodology has to fulfil.

In [9] you will find a lot more details and prescriptions on how to design an evaluation study in practice, like which aspects to consider in each part of the study design, implementation and reporting phases of the evaluation study, such as operationalisation (making practical) of methods from their abstract version.⁴

An example, who are the stakeholders, who are the beneficiaries and who are the victims? Victimhood needs not be in term of power, salary or esteem, but may be in

⁴ For further discussion of these issues, see part II “Methodological considerations of health IT evaluation” in E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

terms of overruled professional responsibility. For example, a new IT system will change the business processes both in structure and content, and thereby potentially interfere with the user's responsibility or even liability (example: decision-support systems, or CPOEs).

4. Balancing the Risk of Bias

The necessity for stringency of scientific work emphasises the need to control all potential sources of variation, any bias within the study. A bias is an inclination towards a systematic deviation of measurements from the 'true' value for the population under investigation – that is, biased data may still be factual and objective but will not be accurate.

Bias is hard to avoid in evaluation of health IT applications, like in any scientific study, just think of the rule in physics that the mere observation of an object inevitably will change the object; so it is important to recognize and to the extent feasible control it. The Hawthorne effect is a similar bias identified in evaluation studies, and in a generalised version it simply states that an organisation (the study object and/or individual components and processes within it) under observation will change; see for instance [3]. It is important to recognize and control the bias at risk, or balancing parameters in the approach so that the bias will have minimal impact on the validity and trustworthiness of the study outcome. The issue is that one needs to be aware of and manoeuvre with the risk of biases in the planning, while constantly remembering that the theoretical and practical impact of biases shall be balanced against the study purpose and the study's role in a (users') decision-making context.

Since this contribution is limited with respect to space, only an abstraction of the issue of biases will be provided here, and the reader is referred to [3] (pages 253-313), where biases are discussed in terms of a meta-framework for assessing evaluation studies. Threats to validity of health IT evaluation studies are discussed elsewhere.⁵

Examples of biases are:

- Selection skewness – that is, when matching structures or components during the recruitment process for controlled studies intended comparables may easily be incomparable (like apples and bananas) and will provide different outcomes, such as comparing a given system when applied in different medical specialties, or recruiting physicians at different professional competence levels (e.g. registrar versus chief physician) for evaluating a decision-support system.
- Skewed frame of reference, for instance, placebo effects, Hawthorne effects, carry-over effects, checklist effects, confounding factors, and more.
- Value-based (emotional inclination), for instance, 1) a developer's versus a user's assessment and therefore a developer should have no influence on an evaluation study before, during or after the event; and 2) technophile or technophobia-based judgements.

⁵ See C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

- Cognitive-based, for instance a range of judgemental observer effects that have been identified in cognitive science, such as judgement of probabilities, insight bias, as well as post-rationalisation, and many more.
- Culture-based (judging phenomena based on one's own culture): cultural decision-making preferences (see later in this paper) will make people behave differently (stubbornly so and significantly so) for instance in the manager-context mentioned briefly in an earlier section.
- One particular bias worth mentioning here is circular inference: "Circular inference arises when one develops a method, a framework, or a technique dedicated to a specific (population of) case(s) and applies it on the very same case(s) for verification purposes." [3], p. 265. So, when designing a method for one particular case/purpose one cannot reuse that particular case to assess the validity of the same evaluation method.

The essence is that a bias when present may render one or more variables unable to reflect objectively the necessary population characteristics. Moreover, since bias is not black and white but comes from a scale of grey nuances, it need not have a significant impact in practise. The 'art' of science (actually 'craftsmanship', i.e. the ability to juggle with methods and handling perspectives) is to know which bias matters and which not for a given setting, and the size of the impact of biases at risk, while reflecting both such awareness and the impact on the conclusions of the study.

5. Decision-Making Preferences: The Impact of Culture

The driver of our decision-making is our respective 'culture' whether we know it or not, and whether we want it or not. Therefore, a simple awareness of its presuppositions, implications and interrelations with (evaluation) methodologies and methods is relevant to briefly address.

"Our understanding of the concept of 'culture' may be expressed shortly this way: *"By cultural behavior, we mean the stability across generations of behavioral patterns acquired through social communication within a group, and valued by the group"* (Maturana 1987, cited and discussed in (Demeester 1995). Culture is the style of working in the field, or the mental, tacit (learned) behavioral pattern behind the style of working (Hampden-Turner & Trompenaars 1993 and 1997; Trompenaars & Hampden-Turner 1997). Thus, culture is guiding the preferences; culture is what comes before starting a discussion of strategy, ..., in a chain of causal events towards problem solving. When specifically talking of the interpretation of culture in an organizational context it means "the acquired preferences in problem solving", where problem solving should be understood in the broadest sense and not only as problem solving in a profession oriented perspective." ([3], p. 289). This is this author's perspective on the concept of culture.

Smaldino brings a thought-provoking example [11] (p. 251): "Perception is constrained in part by our biology, but culture also constrains even our basic perceptions of a situation (Nisbett & Miyamoto 2005; Smaldino & Richerson 2012). For example, (Masuda & Nisbett 2001) showed American and Japanese university students animated underwater scenes with a focal fish. In a recall task, Americans were much better identifying fish they had seen independent of background information, but

Japanese students were much better at remembering details of the background scenes.” Considering that our perception (visual and other) provides the input for our decision-making, such difference is noteworthy. Therefore, beware that different cultures ‘see’ different things when they observe the same object. Further, Smaldino explicitly concludes that cultural differences in patterns of perception and memory fit larger cultural differences in epistemology and styles of thinking that exist between East and West [11].

‘Culture’ is one example of a nation and/or profession-oriented perspective that is hidden in the methods and methodologies that we use. From the two examples (Asian/African culture, and managers as decision-takers) in the Section on “Grounding Possibilities: Theoretical Assumptions and Methodological Considerations” plus one specific bias, it is obvious that culture has an impact on the choice of methodology and of methods; for instance, if ‘No’ is not an option for a fraction of the target end users, then questionnaires with yes/no answers should be excluded as candidate tools. Further, this should be seen in the context of our – subconscious/tacit – culturally conditioned way of perceiving a situation and interpreting observations or designing solutions. Few method designers make the cultural assumptions explicit, and therefore the evaluators need to be aware.

6. Time and Timing of Evaluation

There is a huge difference between constructive (also called formative) and summative evaluation. Constructive evaluation comprises evaluation activities that are completely intertwined with the systems development activities throughout the project (or for a circumscribed period/phase), while summative evaluation is concerned with evaluation at an end point in a developmental path or phase. Constructive evaluation may for instance take place during rapid prototyping, but also at usability studies, and even at the requirements specification phase. So, naturally there is a time and timing issue, because the evaluations cannot and must not significantly delay the systems development. Characteristic is that the outcome of constructive evaluation studies provides substantial input for revision and/or continued systems development, i.e. the trajectory for the future work, rather than merely a verification of contractual fulfilment such as is often the case with summative evaluation. Summative evaluation is also often used to gain insight or measurement of properties without a pressing decision-making information need.

Figure 1 shows the Dynamic Assessment Methodology. It is not a waterfall systems development model, irrespective of its depiction as four sequential phases. The difference is reflected in the contents laid in the arrows. It comprises an example of rapid prototyping developments. It is a model of constructive evaluation in a systems development context that is defined at four phases, from the point of conception of an idea, over iterations of a solution while iterating with usability evaluation, and till impact assessment in a mature clinical setting.

Now zooming out: the implication for scientific evaluation activities (i.e. those that you want to publish in the scientific literature) is tremendous: a) There is no fixed single frame of reference for the evaluation that is valid throughout, which means that the traditional approach to a user requirements document has to differ accordingly, for instance through rapid prototyping. b) There is a risk of a circular inference bias, see above (or more in [3]). c) Evaluation methods have to be chosen accordingly;

applicable methods for different phases and types of user assessment are marked in [3] by means of icons.

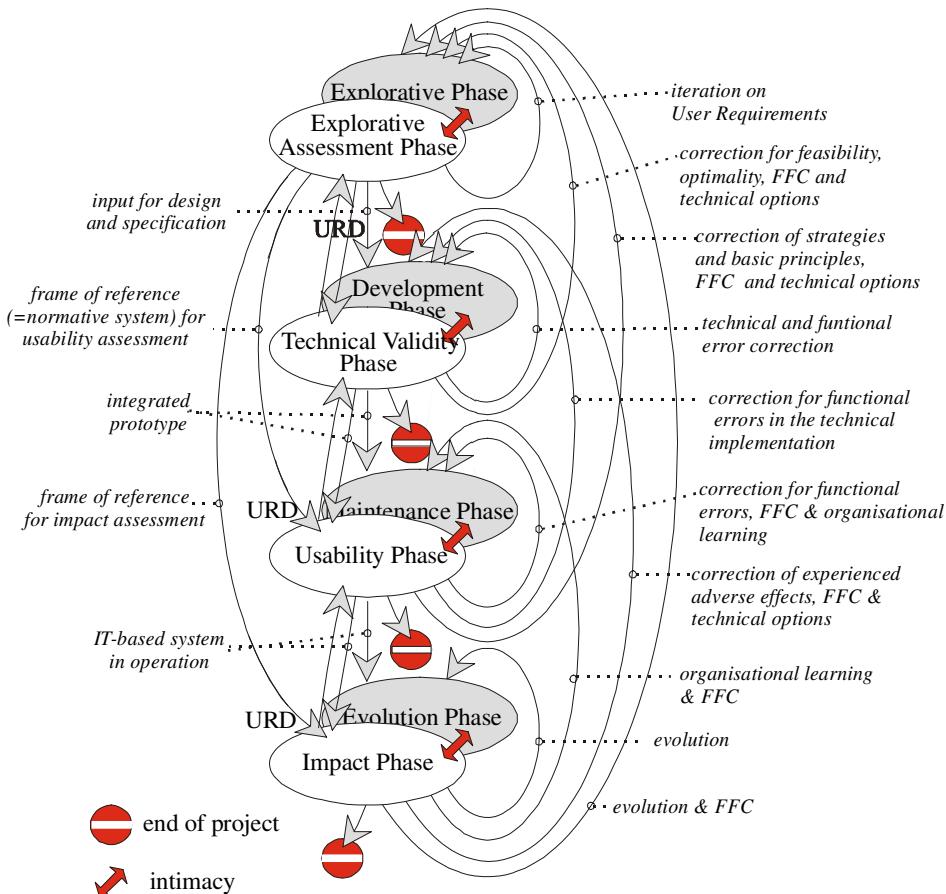


Figure 1. The Dynamic Assessment Methodology, complete with descriptions of feed-forward loops (providing frames of reference and preventive measures) and feed-back-loops (initiating corrective activities) and indicating the contents of this information flow. The shaded ellipses illustrate the technical or development activities, whereas the white ellipses illustrate the corresponding constructive assessment activities in a four-phased structure. The thick arrows indicate unspecified interaction between the technical development and assessment activities (co-ordination and collaboration). (URD = User Requirements Document; FFC = Four Founding Capacities, i.e. characteristics regarding the capability and capacity of accommodating changes), reproduced (modified) from ([3]).

The difficult aspects of constructive evaluation are the indeterminism of systems development and the consequential demand for creativity and innovation in order to continuously comply with the project's information need without delay. When publishing a study on constructive evaluation, one has to watch out for a circular

inference bias (see this above): a method or theory may indeed be highly successfully applied, but one cannot judge the method's various kinds of validity until assessed independently in another setting. Such aspects of validity include for instance: construct validity (does it really measure that which we believe or intend), internal validity (degree of compliance between the perceived meaning and the reality, i.e. with minimal bias), external validity (generalizability to other contexts of investigation), empirical validity (accuracy towards the true value of a measurement), rational validity (coverage or representativeness of characteristics), reliability (consistent outcome), and more; see e.g. [10] – or even Wikipedia – for more detail. When one is aware of this problem, the solution is to phrase the conclusion accordingly when relevant, for instance making a potential risk explicit and/or phrasing the certainty regarding the conclusion with caution (i.e. with weaker words).

7. The Demand for Evidence Enforce a Next Stage of Methodological Approaches

Systems development is indeterministic in nature (i.e. one cannot plan everything in detail, because things change or demand a new decision) and consequentially so is evaluation or at least that of constructive evaluation. Even at this point in the theoretical considerations, it is still important to keep an eye on whether – given the thoughts and considerations so far – the information need (that the evaluation is going to feed, cf. the definition of 'evaluation' in the introduction to this contribution) is likely to be appropriately fulfilled – that is, objectives fulfilment.

In an indeterministic context change is inevitable, so it is important to design the evaluation methodology while taking its sustainability into account. It means that the scientific evaluator needs the competence and experience to be able to incorporate sustainability into the evaluation methodology through fluidity and flexibility, creativity and innovation based on scientific premises.

At this point, a brief helicopter view will inform us whether all of the ends may converge into a coherent sustainable wholeness that in the end will fulfil the information need. They will; after all, evaluation of health IT systems have taken place for decades, meaning that a knowledge base of evaluation studies has accumulated that will pave the way for further studies and will support the validity of future evaluation studies.

Systematic reviews (in the literal sense of this concept) are already practised in the domain of evaluation in health informatics, as for instance seen from [12]. The next stage in evidence-based health informatics is the emergence of meta-analyses (in the Cochrane sense) of concrete cases of evaluation of health IT applications. Meta-analyses and systematic reviews provide answers to different questions, or answers with different levels of certainty attached.⁶ The former has a quantitative nature while the latter has a qualitative (or quasi-quantitative) nature; see also [12]. Key to meta-analyses is comparability among studies as well as degrees of errors and bias, which again points at the importance in the reporting of specific details of the evaluation study rather than raising new methods; see the Section 'Recommended further

⁶ See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

readings'. Errors and bias are really not easy to compensate for even if approaches are available, not least because of the nature of evaluation studies in health informatics as compared to those of clinical trials in medicine in general, or social sciences; nevertheless there may be food for thought and partial help in for instance [13].

8. Selecting the Appropriate and Sufficient Methodology

The key issue in selecting an appropriate and sufficient methodology while taking into account all of the above issues is match-making the scope with the methodology and applicable methods while building a plan based on strategic, tactical and operational issues, as seen in [9]. However, inherent within methodology is the supposition regarding the overall approach: is the evaluation study likely to be a desk-top exercise, a laboratory experiment or an application in a real case scenario? When it comes to evaluation of health IT applications aiming at providing evidence-based facts, then real-life evaluation studies are the most relevant. This is the issue to be discussed in the following.

Evaluation researchers need to demonstrate the validity of new evaluation approaches, methods, or theories, and users need information for their decision-making; thus together there is a potential for a fruitful partnership, which may be achieved through action-case research, and actually, some kinds of evaluation research have no alternative to action-case research. By 'action-case research' is meant the intentional trial application of a researcher's creation (a theory or method) in a real-life case with the purpose of verifying the validity of that creation in real practice. With this definition, 'action-case research' is not the same as the traditional definition of a 'case study' in the sense of illuminating the rationality and implication of a set of decisions, as defined by Yin [14]. We see 'action-case research' as being case-based action research.

McKay and Marshall define action research this way, "Action research is, quite literally, a coming together of action and research, or rephrased, of practice and theory." [15] (p. 219), that is, that such research is accomplished through action, or in other words utilising the research in a practical application. They discuss a set of approaches to action research, all involving informed action and reflection, but with varying degrees of control. The approaches range from, at one extreme, exerting full control – that is, the research interests have precedence in the decision-making regarding the evaluation issues; to an intermediate form where the real-world situation shapes the research interests and questions; to the other extreme, 'consultancy masquerading as action research', in which the real world have precedence over the research interests. "Apart from PhD projects, the majority of IT-systems development and implementation projects are for real-world usage, and the real-world is not an environment where a researcher can try out his methods or methodologies without consequence. Rather, the researcher has to accept the conditions of the real world. ... the user organisation is ... responsible for any decision that will impact their future practice. As an example, the health informatics applications may have huge implications for individual patients and/or for a hospital's economy, and hence may also have liabilities. That is why decision-making in such cases is a serious issue that the researcher can intervene with only within certain limits." [16], p. 51.

Action-case research needs a theoretical foundation and a methodology to guide the real-world problem situation, for instance a model for decision-making. The

research themes and the decision-making as well as the problem-solving are serious factors that must go hand in hand within the methodological design. In such a case, McKay and Marshall's recommendation of "Dual imperatives of Action Research" are taken appropriately into account. Further, in [16] the particular biases at risk in action-case research are discussed: Tacit knowledge and post-rationalisation, intention to treat, insight bias, circular inference, hypothesis fixation, as well as local minima. Apart from the last, they are all outlined above and discussed in more detail in [3]. The bias 'local minima' is a risk in large development or implementation projects where there is a succession of decision-making points; a non-optimal situation may arise when the basis for decision-making in a given situation points at a solution that constitutes a local minimum – that is, the decision appears optimal within the specific context, but may not be in a larger perspective.

With the above understanding it should be possible to start designing an evaluation methodology according to the guidelines in [9].

9. Discussion

The definition of philosophy states that "...making explicit the nature and significance of ordinary and scientific beliefs and investigating the intelligibility of concepts by means of rational argument concerning their presuppositions, implications and interrelationships", and this is what this contribution addressed. The theoretical considerations beneath stringent (evidence-based) evaluation of health IT applications belong under strategic factors in the framework comprising strategy, tactics and operations that are known from military operations and ISO9000. This contribution has dealt with only the strategic aspects at an early point of planning an evaluation study. The same issues have to be revisited during and after implementation - that is, at a follow-up; for example, (at least some of the) biases may be verified by measurement during the implementation of the evaluation study. The tactical aspects relate to making the study real in terms of choice of methodology, methods and action plans, while the operational aspects are concerned with the practical implementation of the evaluation study.

This contribution has mainly dealt with the presuppositions and implications, while the interrelations of the theoretical issues discussed have not yet been addressed. A framework was applied as a template for the entire contribution, namely that provided in [8]. This framework comprises seven sequential functions, each having a specific role in the wholeness of a system and each has an emergent property as output that serves as input for the subsequent function. This framework reflects a system in itself, and the dynamics, the interrelationships between the issues at hand, are handled through successive traversals of the framework – that is, *an iterative and incremental progression* of the issues dealt with; each of the issues discussed in this contribution are dependent on the solution of its predecessor issues.

Recommended further readings

1. J. Brender, *Handbook of Evaluation Methods for Health Informatics*, Academic Press, New York, 2006.
2. J. Talmon, E. Ammenwerth, J. Brender, N. de Keizer, P. Nykänen, M. Rigby, STARE-HI - Statement on reporting of evaluation studies in Health Informatics, *Int J Med Inform* **78**(1) (2009), 1–9.
3. J. Brender, J. Talmon, N. de Keizer, P. Nykänen, M. Rigby, E. Ammenwerth, STARE-HI – Statement on Reporting of Evaluation Studies in Health Informatics, Explanation and Elaboration, *Appl Clin Inform* **4** (2013), 331–358.

Food for thought

1. Why is the systems development model in Figure 1 not a waterfall model, even if it includes four primary and sequential phases? Explain the difference in nature and the implication for the practical evaluation activities.
2. What is the best approach to avoid or circumvent a circular inference bias and provide indicators of internal validity of one's study?
3. Explain the 'local minima' problem at constructive evaluation, described in the Section 'Balancing the Risk of Bias'. How does this concept relate to the other concepts discussed in the above? (e.g. bias, rapid prototyping, time and timing, ...), and what are the implications?

References

- [1] E. Ammenwerth, J. Brender, P. Nykänen, H.-U. Prokosch, M. Rigby, J. Talmon, Visions and strategies to improve evaluation of health information systems - Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, *Int J Med Inform* **73**(6) (2004), 479–491.
- [2] *Collins English Dictionary*, Third Updated Edition, HarperCollins Publishers, Glasgow, 1995.
- [3] J. Brender, *Handbook of Evaluation Methods for Health Informatics*, Academic Press, New York, 2006.
- [4] F. Trompenaars, C. Hampden-Turner, *Riding the waves of culture – understanding cultural diversity in business*, 2nd edition, Nicholas Brealey Publishing Ltd, London, 1997.
- [5] S.C. Schneider, J.-L. Barsoux, *Managing across cultures*, FT Prentice Hall, Harlow, 2003.
- [6] Á. Cabrera, E.F. Cabrera, S. Barajas, The key role of organizational culture in a multi-system view of technology-driven change, *Int J Inf Manag* **21** (2001), 245–261.
- [7] Y.Y. Haimes, C. Schneiter, Covey's seven habits and the systems approach: A Comparative Approach, *IEEE Trans Syst, Man Cybern* **26**(4) (1996), 483–487.
- [8] L. Dennis, J. Brender McNair, L.H. Kauffman, eds., *The Mereon Matrix: Unity, Perspective and Paradox*, Elsevier, London, 2013.
- [9] P. Nykänen, J. Brender, J. Talmon, N.F. de Keizer, M. Rigby, E. Ammenwerth, Guideline for good evaluation practices in health informatics (GEP-HI), *Int J Med Inform* **80**(12) (2011), 815–827.
- [10] J.C. Alderson, C. Clapham, D. Wall, *Language test construction and evaluation*, University Press, Cambridge, 1995.
- [11] P.E. Smaldino, The cultural evolution of emergent group-level traits, *Behav Brain Sci* **37**(3) (2014), 243–254.
- [12] M. Rigby, E. Ammenwerth, M.-C. Beuscart-Zephir, J. Brender, H. Hyppönen, et al., Evidence based health informatics: 10 years of efforts to promote the principle. Joint contribution of IMIA WG EVAL and EFMI WG EVAL, *Yearb Med Inform* **8**(1) (2013), 34–46.

- [13] J.E. Hunter, F.L. Schmidt, *Methods of Meta-Analysis, Correcting Error and Bias in Research Findings*, second edition, Sage Publications, Inc, Thousand Oaks, 2004.
- [14] R.K. Yin, *Case Study Research, Design and Methods*, 3rd edition, Sage Publications, Inc, Thousand Oaks, 2003.
- [15] J. McKay, P. Marshall, *Action Research. A Guide to Process and Procedure*, in: D. Remenyi, editor, Proceedings of the European Conference on Research Methodology for Business and Management Studies 2002. MCIL, Reading, 2002. pp. 219-227.
- [16] J. Brender, *Research Methods for IT/IS Systems' Evaluation, A Critical Appraisal*, in: D. Remenyi, editor, Proceedings of the 2nd European Conference on Information Management and Evaluation, Royal Holloway University of London, 11-12th of September 2008, Academic Publishing Limited, Reading, 2008. pp. 51-58.

Understanding Stakeholder Interests and Perspectives in Evaluations of Health IT

Lisa LEE^{1,a} and Aziz SHEIKH^a

^a*Centre for Medical Informatics, Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Medical School, United Kingdom*

Abstract. Appropriately identifying and representing stakeholders' interests and viewpoints in evaluations of health information technology (health IT) is a critical part of ensuring continued progress and innovation in eHealth. This contribution therefore seeks to clarify the principles of stakeholder analysis in an eHealth context. We describe this with reference to a mixed methods national evaluation of ePrescribing systems in English hospitals. We use this evaluation to exemplify the engagement and analytical tools required to ensure a detailed understanding of the issues, challenges and lessons learnt across stakeholder groups. We conclude that this type of approach may support the robustness of evaluations of health IT as well as their longer term impact on innovation in the field.

Keywords. Evaluation, health information technology, stakeholders.

1. Introduction

Stakeholder analysis, which includes identifying stakeholders and their interests and perspectives, is essential to ensuring a robust health IT evaluation in what are often the unpredictable political contexts in which health IT programmes occur. It is not uncommon for such programmes to repeatedly encounter delays and resistance before any anticipated positive outcomes can be measured [1] making it difficult to produce outcome-based evidence. More specifically health IT projects are often upstream interventions with relatively diffuse effects, which are difficult to measure [2].

Yet underlying the questions of measurement, analysis and application of health IT evaluations, we find a more fundamental, albeit complex, set of issues in relation to how we define stakeholder boundaries of participation, how individual and collective views can be brought together systematically and meaningfully to ensure a robust evaluation, and how this knowledge can be translated and applied to support optimal use of health IT.

Stakeholders in such evaluations may be broadly defined as those involved directly and indirectly in the production and use of health IT at every level. Applying this lens allows us therefore to formulate a simplified analytical framework with two major groups of actors: *producers* – seen as those involved in creating the appropriate contexts and products for the deployment of health IT (e.g., policy makers, software developers); and *users*, who can be seen as those making direct use of health IT (i.e.

¹ Corresponding author: Dr. Lisa Lee, Centre for Medical Informatics, Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Medical School, Teviot Place, Edinburgh EH8 9AG, United Kingdom, Lisa.Lee@ed.ac.uk.

end-users), as well as those who derive value or who are expected to benefit from its use (e.g. healthcare organisations, patients, and so on).

The multiple perspectives these stakeholders represent [3] and the degree of influence they may exert [4] result in complex stakeholder structures [1]. Added to this, the relational quality in terms of the purpose, values, needs and interests of stakeholders at key stages in the development, adoption and use of health IT can make the position of individual stakeholder groups both complex and fluid. This, in turn, can make defining the problems and solutions to system adoptions and use [5] challenging, and can risk leading to unsatisfactory recommendations for best practice.

The crux of the issue in this type of work therefore is ensuring that the evaluation of multiple stakeholders involves three steps – collating evidence from different stakeholders; analysing and interpreting this information, which by necessity will include comparing and contrasting evidence, and responding appropriately to this by striking an appropriate balance between maximising benefits of health ITs and minimising adverse effects for as many stakeholders as possible.

To explore this in more detail we have organised the contribution into three core areas:

- (1) The principles of stakeholder analysis and the range of stakeholders in health IT initiatives, such as purchasers, vendors, professionals, patients, and data warehousing and analytics firms.
- (2) Methods of stakeholder analysis, and how tensions may result from the complex relationships between actors, divergences in their goals and viewpoints, and how these may be addressed in practice.
- (3) Critical walk-through of a national evaluation [6] of hospital ePrescribing systems in England.

2. Principles of stakeholder analysis

It is now well established that the adoption and use of technology involves multiple social processes and unexpected consequences [7-9] on working practices resulting in workarounds or ad-hoc local usage policies and practices that shape new technologies beyond the point of production or market availability. This complex and multifaceted feature of eHealth innovation [4] has led to calls for a more holistic approach [10] to the evaluation and deployment of eHealth technologies in order to improve stakeholder engagement, participatory design and the interconnectedness of all those involved [11]. As such, stakeholder analysis is seen to help support a good 'fit' [10] between the technology and the environment in which it is used, by facilitating incremental improvements to the system over time as use may be optimised [12].

There are a number of key principles of stakeholder analysis which need to be considered at the outset [13] such as: What is the purpose of the analysis? At what stage is it occurring? What aspect(s) are being focused on? What resources are available to carry out the analysis? What is the timeframe?

Clearly decisions on these key aspects of the analysis will impact upon the results. For instance, an analysis occurring over extended timeframes [8] [14] will ensure that stakeholders are accounted for from the point of design right through to primary and secondary uses of the technology, yet may be unable to provide the level of detail required to understand a specific aspect of the deployment. These parameters need to be therefore determined at the point of inception of the evaluation and according to the

evidence required. Notwithstanding these dimensions, a stakeholder analysis needs to be clear about the actors involved and the perspectives they represent. Thus understanding stakeholders' areas of influence, their expectations and goals, are vital since these are seen as determinants in the outcome of health IT adoption [15]. To aid in this process, we suggest that a typology of actors, such as the one presented below, may provide a useful starting point to explore and map stakeholders and their perspectives.

The simplified schema of stakeholders presented in Figure 1 illustrates a number of key points. Firstly, it demonstrates that there may be an overlap between areas of influence and priority, even within a single stakeholder entity. By way of example, we may consider an organisation responsible for the delivery of healthcare nationally, such as National Health Service (NHS) England. Within the producer-user schema, the NHS may be seen as both the producer of an appropriate context of use (through for instance localised policies), as well as being the user of health IT systems, since it is involved in the procurement of the technologies it seeks to deploy.

Such overlaps as well as the distinct goals and expectations of individual stakeholders may result in multiple perspectives and agendas being held within or on behalf of a single organisation. This is perhaps what typified the introduction of Electronic Health Records as part of the National Programme for IT in England, where problems emerged from centrally negotiated contracts on behalf of individual hospitals [1] and therefore ultimately end-users.

In practice, this may translate into tensions between stakeholders and divergences of expectations with potentially disastrous implications for the engagement of end-users [8] and the success of the health IT implementation as a whole. In this respect it is important to ensure there is a detailed breakdown of individual user groups. For instance, even within a single health IT system, there will be divergences and conflicts of viewpoints resulting from the functionalities used within the system and individual professional tasks, so that the perspectives of each professional group may vary as each is may be affected differently by use of the system.

Addressing these tensions is of course an important aspect of the stakeholder analysis. They may be used to both flag up alarm points or areas where additional resources and support may be required to ensure successful system adoption, or where further evaluation and monitoring may be required to assess whether the tensions and conflicts are temporal or likely to be recurring long standing issues.

Stakeholder perspectives therefore need to be considered within a framework in which it becomes possible to disentangle the complex and fluid relationships between actors, the changing nature of the relationships and the environments in which health IT systems are deployed over time [16] as well as the evolving technologies and innovation shifts that occur [17]. In this respect, it is helpful to consider within a health IT evaluation how stakeholders' presence may be mapped and therefore selected over the lifecycle [8] of the technology from project initiation right through to deployment and beyond (which includes system optimisation and secondary data use).

In short a stakeholder analysis needs to reflect the 'social multidimensionality' [18] in which technological appropriation takes place within different institutional contexts. It is these changing contexts of use and interests that for many stakeholders bring about contradictions between the organisational culture to which they may belong and the parameters and resources provided by other stakeholders to which they have to conform, even if reluctantly [18].

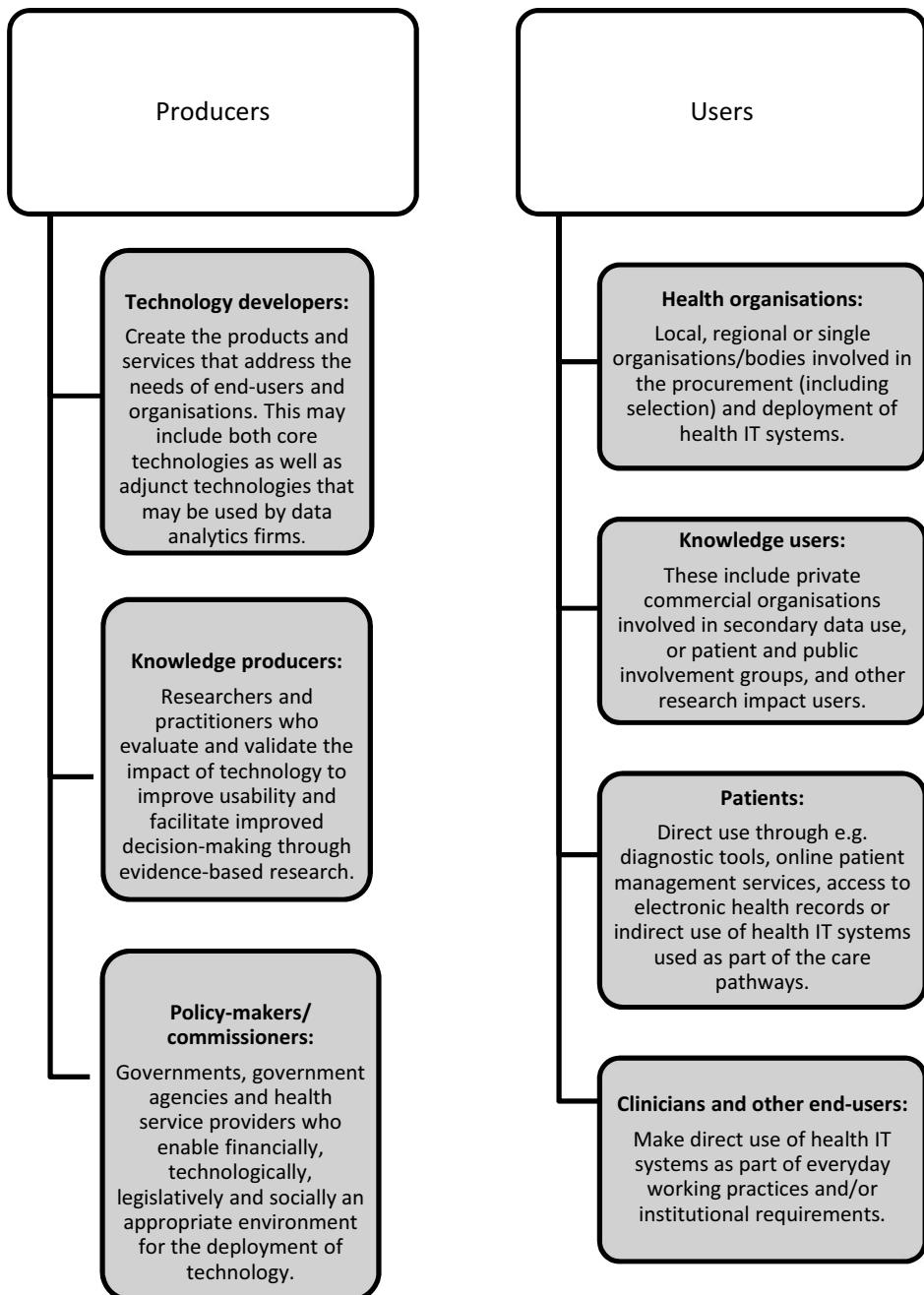


Figure 1. Simplified typology of stakeholders.

The implications for those involved in applying stakeholder analyses for the evaluation of health IT are clear from a methodological point of view: there is a need for flexible, consistent and sufficiently broad ranging evaluation tools that enable these multiple, changing and conflicting views to be both evaluated and brought together. Below, we consider in more detail how this may be achieved in practice.

3. Methods of stakeholder analysis

Methodologies in the evaluation of health IT systems have come under ever closer scrutiny [1, 4, 9-10, 19] and have led to calls to address their shortcomings [19] through more holistic models [10] that enable socio-technical factors [12, 20]² and multiple perspectives to be concurrently evaluated [5]. While quantitative measurements remain a central aspect of health IT evaluations, user-centred bottom-up approaches which can usefully be combined with top-down quantitative approaches offer the flexibility required to include the divergent perspectives of different stakeholders, and ensure a fuller understanding [14] in terms of which individual areas may result in positive, negative or neutral outcomes for instance in terms of levels of implementation and adoption [21].

In other words, there is a need for different perspectives to be explored to understand the impact of an intervention, by reflecting how each stakeholder is affected, why, and what variables need to be changed or adapted in order to improve outcomes. The richness and detail of the qualitative data become especially significant in the era of big data, or when anonymous automated reporting is available within a health IT system, as they provide the necessary contextual evidence while remaining cost-effective [7].

A review of key strategies for the evaluation of eHealth undertaken to date [10] clearly shows the multiple axioms along which health IT evaluations have been designed to capture a wide variety of stakeholder perspectives and views. As suggested earlier on in this contribution, a number of considerations need to be made both at theoretical and empirical levels to align the design of the evaluation with its intended outcomes, as this will support the robustness of the stakeholder analysis.

The evaluation of health IT by means of stakeholder analysis will need to consider first the perspectives that are being captured and analyzed, including whether the analysis is user-centered, multi-faceted and/or multidisciplinary. Weight will also need to be given to contextual factors and frameworks, including legislative, commercial, economic, or socio-technical. The timescales of the evaluation will also be a determinant of the outcome of the stakeholder analysis as views, perspectives and interests may change over time. Therefore whether the analysis is continuous, iterative or phased, will constitute an additional methodological consideration. Finally, special attention needs to be given to ensuring on which aspects of a health IT deployment or adoption are stakeholders' perspectives being sought and what benchmarks are being used to define their perspectives, including whether stakeholders views and interests relate to structure, process, outcomes, procedures, performance or a combination of these.

² See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

With all this in mind, it will become easier to define the appropriate methodological evaluation approaches, such as quantitative clinical trials, qualitative case studies or in larger studies, mixed methods that are able to offer a combination of approaches.

4. Example: Evaluation of ePrescribing in England

The UK's National Institute for Health Research (NIHR) funded evaluation of ePrescribing in England [6] provides a good illustration of the principles and methodological considerations of stakeholder analysis discussed so far, including (1) appropriate mapping of stakeholders and their changing interests and viewpoints over time, and (2) methodological approaches that ensure the ability to capture and triangulate stakeholder perspectives, and to engage with the stakeholder-base as part of the research process. Below we provide a critical walk through how each of these areas has been addressed in this national evaluation of ePrescribing.

4.1. *ePrescribing stakeholders*

The national ePrescribing evaluation highlights not only the multiple actors involved in large scale health IT deployments but also how their expectations and interests can be brought together in an attempt to find resolutions to any conflicts and divergences.

The stakeholder-base involved in the implementation and adoption of ePrescribing systems that offer varying degrees of functionality in the supply, administration, recording and ePrescribing of medication [22] is wide-ranging and includes both producers and users of technology, as discussed earlier. The evaluation has therefore sought to capture the perspectives of: physicians, nurses, pharmacists, and other healthcare professionals; health IT suppliers; patients and carers; policy-makers; hospital managers, IT and finance teams. It has done so by collecting qualitative interview data from key stakeholders, including end-users in case study hospitals before as well as three to six months post-implementation, and again once the system might be considered embedded. This has enabled a longer-term perspective on the introduction and use of ePrescribing systems in English hospitals to be taken in order to take into account evolving situations and to assess how changes that happen over time may impact on stakeholder perspectives [8].

Many eHealth implementations are tainted, especially in immature digital markets [12] with unrealistic and wide-ranging expectations that have adverse effects on engagement [12,14,23-26] which may provide falsely negative stakeholder perspectives for instance if the system is considered having few benefits, when problems may in fact be the result of lack of readiness.

The collation of detailed case studies of hospitals deploying ePrescribing systems with different functionalities, at different stages of deployment and adoption, and in different geographical regions, has allowed therefore for cross-comparisons and disconfirming searches to help understand divergences and similarities between sites. This strategic selection of case studies has provided an opportunity to balance stakeholder perspectives and conflicting views when developing recommendations for best practice. This type of approach further allows the narrative behind the introduction and adoption of ePrescribing to be meaningfully applied throughout the lifecycle of the

system [8] and to establish when behaviour, events or technical issues are transient, and where results suggest a longer term effect. This longer-term perspective is seen as especially critical in the context of an immature product, such as ePrescribing, which will be shaped by its users and the context within which it is being used.³

4.2. Mixed methods for robustness of stakeholder analysis

While the ability to capture and contrast stakeholder perspectives over time was in the ePrescribing evaluation achieved by means of qualitative case studies, robustness of the stakeholder analysis has been enhanced through the use of mixed methods⁴ which provide complementary stakeholder perspectives at key stages [27] along the system development and care pathways (supplier, NHS organisation, patients).

This mixed method approach has provided measurements in various forms of the anticipated benefits of ePrescribing, by looking both qualitatively and quantitatively at safety and error rates [28], efficiency and cost benefit [2] and communication [12, 25, 29].

It is important to note also how the perspectives of patients – a key yet often neglected stakeholder group – has been facilitated through the inclusion of a Patient and Public Involvement Group throughout the evaluation to influence and challenge perspectives individually and collectively at each stage of the research. Importantly, these research strategies and tools, as well as the findings being generated from them, have been used to engage and inform stakeholders via an online toolkit www.ePrescribingtoolkit.com [30]. This provides not only engagement but also an alignment of the goals of stakeholders by supporting and promoting successful implementation strategies that draw on evidence-based research.

Findings from the stakeholder analysis can thereby remain both reflective and outwardly engaging towards ePrescribing stakeholder communities, whether they be commercial players, policy-makers, health organisations or clinicians as well as patients themselves, and may help unpick the complex relationships between stakeholders [13] at critical stages in the health IT systems' adoption [27]. The toolkit alongside various closed and open stakeholder events organised as part of the evaluation [31] have moved the analysis beyond identifying its stakeholder-base and a description of their perspectives, to an active form of participation in the research as stakeholders are both subjects and users of the research, thus allowing knowledge derived from the analysis to be applied meaningfully.

5. Conclusions

It is worth remembering that while we advocate the use of a stakeholder analysis that enables as many perspectives as possible to be considered over extended timeframes and at different stages of health IT deployments, practical considerations such as costs,

³ For further discussions of evaluating health IT for medication safety, see: H. Seidling et al., Evaluating the impact of health IT for medication safety, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁴ See also: P.J. Scott et al., Mixed methods: a paradigm for holistic evaluation of health information system, in: *ibid.*

resources, expertise and so on, do need to be taken into account and compromises will need to be made where necessary.

In the context of the ePrescribing evaluation, this included the strategic selection of case study sites which hold particular known characteristics to enable a good balance between reliability and efficiency of data collection, attaching local researchers to individual sites as well as collaborative modalities of data collection at each site, for instance when ward pharmacists collected quantitative data on error rates as part of daily ward rounds.

While the complexity of the stakeholder relationships and the reconciliation of their perspectives to help foster technological usability, innovation and participation may be addressed through integrated methodological approaches [19, 32-33] and transdisciplinary collaboration [5,9] stakeholder analyses are far from straightforward [19]. A number of steps may help address the difficulties encountered.

Firstly, ensuring the timing of the analysis is appropriate enables the evolving nature of health IT [12] and its diffuse effects [2] to be considered. Secondly, when wide ranging issues from usability and design, staff training, increased time required to perform clinical duties, or the impact of eHealth systems on face-to-face interactions between patients and Health Care Professionals are flagged up during the analysis, it is vital to support appropriate utilisation of this knowledge [34] to address the translational gap in its application [35]. A stakeholder analysis which is being used as part of a health IT evaluation needs to consider fully therefore how best to manage findings [36] to allow stakeholders appropriately to plan, implement and make optimal use of this knowledge when expertise of eHealth system implementations and adoption is limited [30, 37]. This will help address challenges posed by conflicting stakeholder perspectives, such as when interventions are viewed positively by patients, but are found to be ineffective or not cost-effective in the analysis. Finally, it is worth noting that variations globally in how healthcare technologies may be adopted are significant for the applicability of a stakeholder analysis. Indeed local norms may affect the usefulness of a stakeholder analysis [13]. As such it is important to be mindful both of the feasibility and usefulness of seeking multiple stakeholder perspectives in particular geographical settings globally with distinct organisational cultures.

The points we have made throughout this contribution should be a stark reminder to both policy makers and researchers in the field that health IT evaluations do need appropriate time, methodological approaches, resources and expertise if they are to fulfil their objective.

Acknowledgement

The ePrescribing Programme is funded by the National Institute for Health Research's Programme Grants for Applied Research.

Recommended further reading

1. R. Brugha, Z. Varvasovszky, Stakeholder analysis: a review, *Health Policy and Planning* **15**(3) (2000), 239-246.

Food for thought

1. What are the key ways in which the complex stakeholders' perspectives can be evaluated?
2. How can knowledge transfer be used to help balance stakeholder perspectives in the evaluation?
3. What issues might arise in the evaluation of health IT systems across organisational cultures or geographical settings?

References

- [1] P. Klöcker, R. Bernnat, D.J. Veit, Stakeholder behavior in national eHealth implementation programs, *Health Policy and Technology* **4**(2) (2015), 113-120.
- [2] R. Lilford, A. Girling, A. Sheikh, J. Coleman, P. Chilton, S. Burn, D. Jenkinson, L. Blake, K. Hemming, Protocol for evaluation of the cost-effectiveness of ePrescribing systems and candidate prototype for other related health information technologies, *BMC Health Services Research* **14** (2014), 314.
- [3] D. Ahern, J. Kreslake, J. Phalen, What Is eHealth? Perspectives on the Evolution of eHealth Research, *J Med Internet Res* **8**(1) (2006):e4.
- [4] L. Catwell, A. Sheikh, Evaluating eHealth Interventions: The Need for Continuous Systemic Evaluation, *PLoS Med* **6**(8) (2009): e1000126.
- [5] L. Catwell, A. Sheikh, Information Technology (IT) users must be allowed to decide on the future direction of major national IT initiatives. But the task of redistributing power equally among stakeholders will not be an easy one, *Informatics in Primary Care* **17**(1) (2009), 1-4.
- [6] ePrescribing Programme, *NIHR ePrescribing Programme: Overview*, <http://www.cphs.mvm.ed.ac.uk/projects/eprescribing>, last access 11 February 2016.
- [7] I. Savage, T. Cornford, E. Klecun, N. Barber, S. Clifford, B. Franklin, Medication errors with electronic prescribing (eP): Two views of the same picture, *BMC Health Serv Res* **10** (2010), 135.
- [8] K. Cresswell, D. Bates, Sheikh, A. Ten key considerations for the successful implementation and adoption large-scale health information technology, *J Am Med Inform Assoc* **20**(e1) (2013), e9-13.
- [9] C. Pagliari, Design and evaluation in eHealth: Challenges and implications for an interdisciplinary field, *J Med Internet Res* **9**(2) (2007).
- [10] J. Van Gemert-Pijnen, N. Nijland, M. Van Limburg, H. Ossebaard, S. Kelders, G. Eysenbach, S. Erwin, A Holistic Framework to Improve the Uptake and Impact of eHealth Technologies, *J Med Internet Res* **13**(4) (2011).
- [11] M. Yusof, J. Kuljis, A. Papazafeiropoulou, L. Stergioulas, An evaluation framework for Health Information Systems: human, organization and technology-fit factors, *Int J Med Inform* **77**(6) (2008), 386-398.
- [12] K. Cresswell, D. Bates, R. Williams, Z. Morrison, A. Slee, J. Coleman, A. Sheikh, Evaluation of medium-term consequences of implementing commercial computerized physician order entry and clinical decision support ePrescribing systems in two 'early adopter' hospitals, *J Am Med Inform Assoc* **21**(e2) (2014), e194-202.
- [13] R. Brugha, Z. Varvasovszky, Stakeholder analysis: a review, *Health Policy and Planning*, **15**(3) (2000), 239-246.
- [14] T. Greenhalgh, J. Russell, R. Ashcroft, P. Parsons, Why National eHealth Programs Need Dead Philosophers: Wittgensteinian Reflections on Policymakers' Reluctance to Learn from History, *Milbank Quarterly* **89**(4) (2011), 533-563.
- [15] T. Ingebrigtsen, A. Georgiou, R. Clay-Williams et al., The impact of clinical leadership on health information technology adoption: Systematic review, *Int J Med Inform* **83** (2014), 393-405.
- [16] S. Bullas, J. Bryant, Complexity and its Implications for Health Systems Implementation, in *Information Technology in Health Care*, IOS Press Amsterdam, 2007, pp. 37-44.
- [17] M. Grisot, P. Vassilakopoulou, Infrastructures in healthcare: The interplay between generativity and standardization, *International Journal of Medical Informatics* **82**(5) (2013), 170-9.
- [18] A. Boonstra, M. Van Offenbeek, Towards consistent modes of e-health implementation: structural analysis of a telecare programme's limited success, *Info Systems Journal* **20**(6) (2010), 537-561.

- [19] A. Ekeland, A. Bowes, S. Flottorp, Methodologies for assessing telemedicine: A systematic review of reviews, *Int J Med Inf* **81**(1) (2011), 1-11.
- [20] L. Novak, R. A. S. Holden, J. Hong, B. Karsh, Using a sociotechnical framework to understand adaptations in health IT implementation, *Int J Med Inform* **82**(12) (2013), e331-44.
- [21] A. Jha, D. Doolan, D. Grandt, T. Scott, D. Bates, The use of health information technology in seven nations, *Int J Med Inform* **77** (2008), 848-854.
- [22] S. Goundrey-Smith, *Principles of Electronic Prescribing*, Springer-Verlag, London, 2012.
- [23] A. Takian, D. Petrakaki, T. Cornford, A. Sheikh, N. Barber, Building a house on shifting sand: methodological considerations when evaluating the implementation and adoption of national electronic health record systems, *BMC Health Services Research* **12** (2014), 105.
- [24] K. Cresswell, A. Slee, J. Coleman, R. Williams, D. Bates, A. Sheikh, Qualitative Analysis of Round-Table Discussions on the Business Case and Procurement Challenges for Hospital Electronic EPrescribing Systems, *PLoS ONE* **8**(11) (2013), e79394.
- [25] A. Sheikh, T. Cornford, N. Barber et al., Implementation and adoption of nationwide electronic health records in secondary care in England: final qualitative results from a prospective national evaluation in 'early adopter' hospitals, *BMJ* **343** (2011), d6054.
- [26] A. Robertson, K. Cresswell, A. Takian et al., Implementation and adoption of nationwide electronic health records in secondary care in England: qualitative analysis of interim results from a prospective national evaluation, *BMJ* **341** (2010), c4564.
- [27] K. Cresswell, J. Coleman, A. Slee et al., A toolkit to support the implementation of electronic ePrescribing systems into UK hospitals: preliminary recommendations, *J R Soc Med* **107**(1) (2014), 8-13.
- [28] S. Thomas, S. McDowell, J. Hodson, U. Nwulu, R. Howard, A. Avery, A. Slee, J. Coleman, Developing consensus on hospital ePrescribing indicators of potential harms amenable to decision support, *Br J Clin Pharmacol* **76**(5) (2013), 797-809.
- [29] J. Hendy, N. Fulop, B. Reeves, A. Hutchings, S. Collin, Implementing the NHS information technology programme: qualitative study of progress in acute trusts, *BMJ* **334**(7608) (2007), 1360-1364.
- [30] L. Lee, K. Cresswell, A. Slee et al., Using stakeholder perspectives to develop an ePrescribing toolkit for NHS Hospitals: a questionnaire study, *JRSM Open* **5**(10) (2014), 1-9.
- [31] NIHR ePrescribing Programme, Past Conferences and Events, <http://www.cphs.mvm.ed.ac.uk/projects/eprescribing/pastevents.html>, last access 11 February 2016.
- [32] D. Carr, A. Howells, M. Chang, H. N, E. A. N, An integrated approach to stakeholder engagement, *Healthc Q* **12**, (2009), 62-70.
- [33] K. Unertl, B. Johnson, C. Gadd, N. Lorenzi, Bridging Organizational Divides in Health Care: An Ecological View of Health Information Exchange, *JMIR Medical Informatics* **1**(1) (2013).
- [34] H. Länsisalmi, M. Kivimäki, M. Elovainio, Is underutilization of knowledge, skills, and abilities a major barrier to innovation? *Psychol Rep* **94** (2004), 739-50.
- [35] A. MacFarlane, P. Clerkin, E. Murray et al., The e-health implementation toolkit: qualitative evaluation across four European countries, *Implement Sci* **6**(122) (2011).
- [36] J. Wyatt, Management of Explicit and Tacit Knowledge, *J R Soc Med* **94** (2001), 6-9.
- [37] E. Murray, C. May, F. Mair, Development and formative evaluation of the e-Health Implementation Toolkit (e-HIT), *BMC Med Inform Decis Mak* **10**(61) (2010).

Domains of Health IT and Tailoring of Evaluation: Practicing Process Modeling for Multi-Stakeholder Benefits

Sara MARCEGLIA^{a,1}, Simona FERRANTE^b, Stefano BONACINA^c,

Francesco PINCIROLI^b, Irene LASORSA^a, Chiara SAVINO^b, Giuseppe POZZI^b

^a*Dipartimento di Ingegneria e Architettura, Università degli Studi di Trieste, Trieste, Italy*

^b*Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy*

^c*Health Informatics Centre, Department of learning, informatics, management and ethics, Karolinska Institutet, Stockholm, Sweden*

Abstract. This contribution focuses on the heterogeneity and complexity of health information technology services and systems in a multi-stakeholder environment. We propose the perspective of process modeling as a method to break out complexity, represent heterogeneity, and provide tailored evaluation and optimization of health IT systems and services. Two case studies are presented to show how process modeling is needed to fully understand the information flow, thus identifying requirements and specifications for information system re-engineering and interoperability; detect process weaknesses thus designing corrective measures; define metrics as a mean to evaluate and ensure system quality; and optimize the use of resources.

Keywords. Process assessment (health care), complexity, flexibility, electronic prescribing.

1. Complexity of healthcare and its impact on health-IT design and evaluation

Our journey for being and staying healthy is complex [1], is life-long, involves multiple actors to cover our different needs, and, as with many other aspects of our life, is now supported by technology. Or this is what we expect.

However, the concept of complexity should be defined more precisely: heterogeneous action sub-domains, dynamic evolution of knowledge, learning curves, indeterminacy, uncertainty, exceptions, transparency, and data protection are some of the features contributing to the concept of complexity that deserve some consideration [2].

In healthcare, two sub-domains of actions, the clinical sub-domain (devoted to patient care) and the administrative domain (devoted to the economic and financial aspects of care) share the same information regarding the patient, but need different views that focus on the specific data. For instance, in prescribing pharmacological treatments to patients, the active component, beneficial effects, side effects, adverse

¹ Corresponding author: Associate Professor Sara Marceglia, Dipartimento di Ingegneria e Architettura, Università degli Studi di Trieste, via A. Valerio 10, 34127 Trieste, Italy, smarceglia@units.it.

events, and drug-drug interaction are clinically relevant to identify the prescribed drug, whereas the costs and reimbursement levels are relevant from the administrative viewpoint [3].

The dynamic evolution of medical knowledge implies that any health IT system aimed at supporting medical decision-making not only has to deal with the available evidence-based medicine, but also has to be ready to dynamically and flexibly include new relevant evidence that may arise, personal experiences [4] and learning curves (i.e., learning by practice) [5].

In addition, uncertainty and indeterminacy [6] mainly arise from (1) the patient's compliance and response to treatment that depend on the ability of patients to follow the instructions, their level of engagement, and health literacy; and (2) the ability of the patient to report the outcomes or complications of a treatment or therapy thus affecting the interpretation and judgment of the healthcare professionals responsible for it [7]. Also, the patient-centric approach in the design and development of health IT systems requires a level of personalization that may introduce "exceptions" and "deviations" from available clinical guidelines and recommendations, thus introducing another layer of "complexity".

Finally, the need of cooperation among different actors or roles within the patient/citizen care pathway [8] is translated into the need for a clear definition of roles, profiles, data views, and actions allowed, that could be summarized with the term "transparency".

All the features that we discussed so far show that the intrinsic complexity of the whole healthcare domain cannot be avoided and needs to be uncovered when designing, developing, and evaluating effective health IT systems and services.

2. Heterogeneity of Health IT: Multiple systems and multiple actors

The complexity of the healthcare domain is reflected in health IT systems that provide the technological support to the whole healthcare journey which is not limited to the periods when we are "patients", but it spans across our whole life, with different needs. We therefore need to distinguish between the "citizen" who is not a patient until she/he receives a diagnosis, and the "patient" who suffers from a disease (with possible comorbidities).

The classical healthcare pathway, that starts from prevention, until the patient receives a diagnosis, and then a treatment (or rehabilitation), can be seen from the two perspectives of the "patient" and the "citizen". The citizen is the main actor in the prevention phase, but still contributes to the healthcare journey when involved as caregiver. Similarly, the patient is more active in the last two phases but also participates in the prevention phase, either to keep her/his pathology under control, or to avoid comorbidities.

In this promising scenario, the patient and the citizen, despite being the main characters, are only two of the actors involved. Table 1 shows an attempt to represent the available health IT systems and services according to their main user profiles and the phase of the personal healthcare pathway. Whereas patients and families/caregivers have tools for all phases, the citizen is not considered as an active user profile in the treatment or rehabilitation phase. Other stakeholders are healthcare professionals and providers, the payers (public/private/insurances), and also students and researchers. Of course, for students and researchers, the tools are not specific for a phase of the

healthcare pathway but cover all of them. Even though not exhaustive, Table 1 depicts a heterogeneous environment, in which patient's and citizen's health depends on the intervention of different stakeholders who mainly need to collect relevant information regarding the patient's/citizen's conditions to take the right decisions.

Table 1. Examples of health IT systems classified according to the main final user (rows) and the phase of the healthcare processes (columns).

		Prevention	Diagnosis	Treatment and Rehabilitation
Citizens	Diet monitoring Exercise monitoring Educational tools Personal Health Record (PHR)	Communication with healthcare professionals Unsupervised symptom checkers e-services for checking symptoms Health information websites/apps		
Patients	Telemonitoring & Sensors Environmental monitoring Educational tools Personal Health Record (PHR)	e-services for checking symptoms Telediagnosis tools Portals for ranking/finding physicians	Drug tracking systems Telerehabilitation systems Patient portals	
Family/Caregivers	Activity trackers Educational tools	Communication with healthcare professionals	Drug tracking systems Telerehabilitation systems Community support tools Family Health Records	
Healthcare Professionals and hospitals/care centres	Risk assessment tools Screening and Telescreening Decision Support Systems Electronic Health Record (EHR)	Electronic Health Record (EHR) Supervised Symptom checkers Decision Support Systems Domain Ontologies and Knowledge representation systems Hospital Information System Reference databanks Biosignal/Bioimage Databanks	Telecare systems Computer Interpretable Guidelines (CIGs) and Recommendations Electronic Health Record (EHR) Social care records and supporting systems Reference databanks	
Private/Public/insurance Payers	Insurance-provided PHRs Risk assessment	Health Information Systems	Telecare systems	
Medical Students		Visual knowledge tools Bioimage databanks Online reference systems Virtual environments		
Researchers		Clinical report Forms (CRF) Shared Databanks Multicentre research platforms Reference databanks Crowdsourcing tools		

The number of different actors involved, the various levels of digitalization in different healthcare organizations, the processes implemented within these organizations, as well as their privacy and security policies and issues, contribute to create a scattered and heterogeneous scenario, in which information systems manage heterogeneous, and often redundant, information, with poor inter-communication,

caused by a lack of interoperability. This argues the need for appropriate tools able to break out such complexity thus providing tailored and effective evaluations and to follow the patient or citizen in a longitudinal, life-long, integrated perspective [9]. A process modeling approach provides such tools.

3. Breaking out complexity and representing heterogeneity to provide tailored evaluation: the process modeling perspective

Being an abstract representation of a process under examination, a model provides a clear representation of the actors, the roles, their tasks, their actions and resources, and tracks the information flow and the core phases throughout the process [10]. Hence, a model provides a clear and “transparent” view, in which all the complex features of the process are represented and analysed.

Process modeling can be used both in the design phase of health IT systems, especially when the model is represented using standard languages for software design (e.g., the Unified Modeling Language, UML), and in the evaluation and re-engineering phase. In fact, the reliable, shared, transparent, and multi-level description of the process underlying the health IT system facilitates (1) the understanding of how a system works and how it can be integrated with other systems operating within the same process, and (2) the application of standard-based solutions. This ultimately supports interoperability and integration among different health IT systems [8] by:

- *Representing the impact of the single IT system* on the process itself, thus providing an evaluation of the benefits introduced with the use of the system and its limitations
- *Comparing different IT systems*, to establish which system better fits specific needs, in a given setting with specific constraints, in order to choose the most appropriate solution
- *Simulating the use of the IT system in another setting*, by changing the local constraints and evaluating its possible impacts and effectiveness in different environments.

The most important clinical benefits of the application of process modeling to health ITs are creating shared protocols based on clinical guidelines and local practices and monitoring the adherence to them; facilitating the communication among different actors and roles all contributing to patient’s care; highlighting process weaknesses and suggesting the applicable corrective measures; providing a clear view on the use and optimization of resources; fully understanding the information flow; and identifying requirements and specifications for information system re-engineering to promote interoperability [11].

Finally, models are usually represented in a graphical way that facilitates their sharing among the different actors involved, even though not expert in technologies and modeling. This implies that their multiple viewpoints can be involved in the evaluation (or design) of a system/service, thus enlarging the evaluation scenario and including the heterogeneous expertise, needs, and aims.

4. Basic principles of process modeling

The integration between the skills and knowledge of domain experts and analysts is essential to model a healthcare process (Figure 1). Domain experts provide the experience on the field, and are aware of the existence of clinical guidelines and evidence-based practices related to the specific process. On the other hand, analysts are able to translate the experiences and knowledge of domain experts in a model, and from it to extract requirements and evaluation criteria.

4.1. Basic steps

The general approach to process modeling for health ITs is composed of three phases (Figure 1) - the analysis of the environmental context, the conceptual modeling, and the logical modeling. They are described as follows:

1. *The analysis of the environmental context* includes the identification and analysis of the available sources of information (also evidence-based knowledge, international guidelines, and recommendations) and the understanding of the local domain in terms of local practices, and specific clinical pathways already in use locally. Focus groups and interviews with the medical staff or of the patient and caregivers highlight the personal experience of the actors involved in the process. This phase includes the analysis of the flow of information that is managed by the health IT system or service and its interaction (or need of interaction) with other existing systems, which is crucial to understand whether the data/information transmitted through the system under study is effectively used and received, and helps identifying the possible flaws. As part of this phase there is also the selection of the formal modeling notation (language, as, for example, the Unified Modeling Language – UML [12] or the Business Process Modeling Notation – BPMN [13]).
2. *The conceptual modeling* includes a pre-modeling and a modeling phase. During pre-modeling it is important to provide a high-level process description (process phases) in terms of functional aspects (main activities of the process, objects and data items managed), organizational aspects (agents, roles, tasks, skills, availabilities, authorizations required to enact the process), actors' responsibilities on the main activities, and business aspects (goals to be achieved). Outcomes have to be identified at this time, too. They can be either clinical outcomes (for example to evaluate a telemonitoring system) or functional or efficiency outcomes (for example to evaluate a booking system). The modeling phase produces the conceptual model of the process according to the formal notation adopted and comprises: the schema of the process with its variables, the specification of the expected exceptions and transactions, the access control model, and the representation of the information flow with external information systems. Thanks to the definition of the model, it is at this stage possible to define appropriate metrics, either to evaluate the process itself or to monitor the health IT service/tool that supports/implements the process. Domain experts act as feedback during the whole phase, to validate the model under construction. The validation, in fact, should not only verify whether the model is “syntactically” correct (internal consistency and usability

- for system requirements definition), but also if it is “semantically” correct (validation of the information flow in the simplest activities of the process, and verification of the expectations of all the actors involved).
3. *The logical modeling* is the final phase in which the model is implemented either in an executable modeling language, or as a full system (or system modules). In this phase it is important to design the possible solutions to the critical issues identified, or to highlight the requirements for system re-engineering [11].

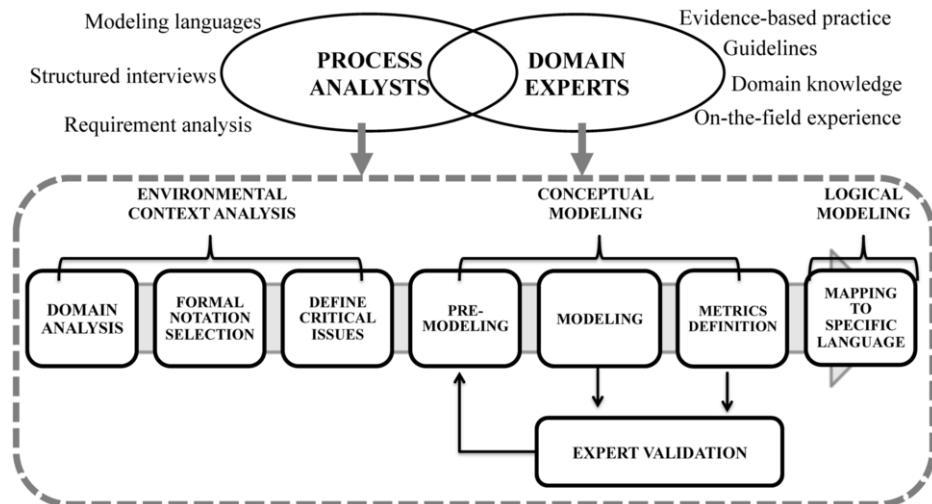


Figure 1. Basic modeling principles.

4.2. Metrics and process evaluation

A Metric is defined as “a quantitative measure of the degree to which a system, component, or process possesses a given attribute” [14] and it is based upon two or more measures. Metrics for the evaluation of health IT cannot be directly derived from the model itself. Models do not provide a direct means for cost-effectiveness or cost-benefit analysis. However, the model can be the basis for identifying the outcome variables to be introduced into e-management techniques as metrics for evaluation.

Moreover, as aforementioned, in order to deal with the overall heterogeneity of the healthcare processes and its stakeholders, not only metrics related to the economic factors must be considered. Health Technology Assessment (HTA) can represent a valuable approach for evaluating the single health IT system and for comparing it with different systems, since the HTA is a multidisciplinary and multidimensional approach for analyzing all the areas of interest (e.g. epidemiological, economic, social, ethical,

legal, organizational, and political implication) [15]². Nevertheless, HTA is not always supported by structured techniques for the evaluation and prioritization (i.e., multiple-criteria decision analysis, such as the Analytic Hierarchy Process - AHP [16]).

Besides the prioritization and evaluation of the overall health IT systems, the processes behind it need to be monitored identifying the best metrics, and the proper time when they need to be measured for process evaluation. Indeed, the major regulations and directives for hospital accreditation and certification (e.g., Joint Commission International, ISO 9001:2008) require to define and model processes, and to identify the most appropriate performance measures (where performance measurement is defined as “a system for assessing performance of development interventions against stated goals” [17]) that can be organized in metrics. Nevertheless, they do not always specify *how* to define and collect performance measures and, consequently, metrics.

On the other hand, the application of process improvement techniques is rapidly growing in the healthcare context, and approaches originally linked to manufacturing areas (i.e., Lean Management) are being recently extended also to hospitals. Lean Management techniques suggest metrics for evaluating a process and its wastes (the “Lean Key Performance Metrics”), such as On-Time Delivery, Customer Lead Time Reduction, Inventory Turns, and Overall Efficiency Percentage Gains. Nevertheless, some of these metrics lack a unique definition, especially in the healthcare context.

Another technique for identifying metrics of interest that overcomes the limitations described above, is the Goal Question Metrics (GQM) [18]. The GQM allows selecting metrics with a top-down and goal-oriented approach, and it can be exploited for gathering the measurement data and driving decision-making and improvements, providing a support for the identification of the metrics starting from the definition of goals. The definition of the goals during the initial conceptual modeling of the process facilitates the implementation of the GQM. Specifically, the GQM can be divided into three levels: Goal (Conceptual level, defines the main purposes of a work to be measured); Question (Operational level, defines a set of questions useful for achieving the goals); Metric (Quantitative level, defines a set of metrics for answering the questions in a measurable way). The GQM is a versatile approach and can be considered a useful technique for defining metrics of health IT.

5. Case studies

This section presents two case studies, chosen to instantiate the considerations discussed in the previous parts of this contribution. The first case study looks at the use of process modeling to represent e-prescribing systems, providing a model-based evaluation framework able to identify the aims and needs of different systems, and to identify the gaps that require re-engineering [3].

The second case study shows the evaluation of the system for managing the pathway for cancer patients. Process modeling in this case was able to highlight an information loss in the ambulatory setting that does not impact the clinical outcome of the patients but the treatment reimbursement (administrative perspective, unpublished

² See also: P. Doupi, Evolving Health IT systems evaluation: the convergence of health informatics and HTA, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

research), and led to the development of a new module of the hospital information system able to manage the information loss during the day care process.

5.1. Model-based representation of e-prescribing systems

E-prescribing is a complex process that differs from the simplistic idea of “transmitting a digital prescription to the pharmacy” [19]. It is a closed-loop process that starts from the decision of which treatment to prescribe to the patient and ends with the patient’s clinical outcomes, with adverse events and clinical outcomes as feedback variables [19]. Heterogeneous e-prescribing systems are available worldwide, with different aims, in different contexts, and processing different information. This varied scenario claims for a reliable framework for the representation of different e-prescribing systems and for the evaluation of the benefits associated to their introduction [3].

For this reason, in cooperation with the Italian Government and in the framework of the Italian digitalization program for the Public Administrations entities (DigitPA), the modeling approach was applied to the e-prescribing process in order to (1) understand the possible benefits gained by the introduction of an e-prescribing system, and (2) compare existing e-prescribing solutions in terms of benefits for the healthcare system.

The modeling steps reported in Figure 1 were followed to create the e-prescribing model described in [3]: after the environmental context phase, which included also the direct interaction with the Italian Ministry of Health, the conceptual modeling phase provided the high-level meta-model (activity diagram), the identification of the interacting systems and tools, the definition of the expected outputs from each activity, and the definition of the evaluation outcomes in terms of expected benefits (Figure 2). For details about the model, see [3].

The evaluation framework was based on the verification of the correct implementation of specific functions that were called “verification actions”. In each phase of the process, the model defines these “verification actions” that guarantee a specific benefit, with a fine granularity. For example, during the first phase of the process (i.e., Assign phase, when the treatment is prescribed to the patient, Figure 2A), if the e-prescribing system verifies the existence of a coded diagnosis in the prescription document, we can expect two kinds of benefits: (1) that the drug is assigned with a valid clinical reason (quality of care dimension, increased patient’s safety), and (2) that the relationship drug-diagnosis is tracked and can be used for drug surveillance (efficiency of care dimension). In a similar way, if the e-prescribing system guarantees the verification of drug-drug interactions, we can expect a decreased risk of undesired adverse drug events (ADEs) or altered outcomes due to the interaction of the prescribed drug with others already in use (quality of care dimension). For the full description of each verification action and its expected benefit, see [3].

Aimed at providing a framework for evaluating and comparing e-prescribing systems, the modeling effort ended at this point. So, the presence/absence of model elements (i.e., verification actions) was used to compare the potential benefits introduced by three e-prescribing systems, namely the Lombardy Region (Italy), the Andalusia Region (Spain), and the Italian Government. The analysis, that is fully reported in [3], showed that all systems lack the connection between the first three phases (from “assign” to “delivery”) and the administration phase, when the patient is at home, thus suggesting that the available e-prescribing systems fail in integrating the

patient as an active user and lose important information on drug administration and effect monitoring.

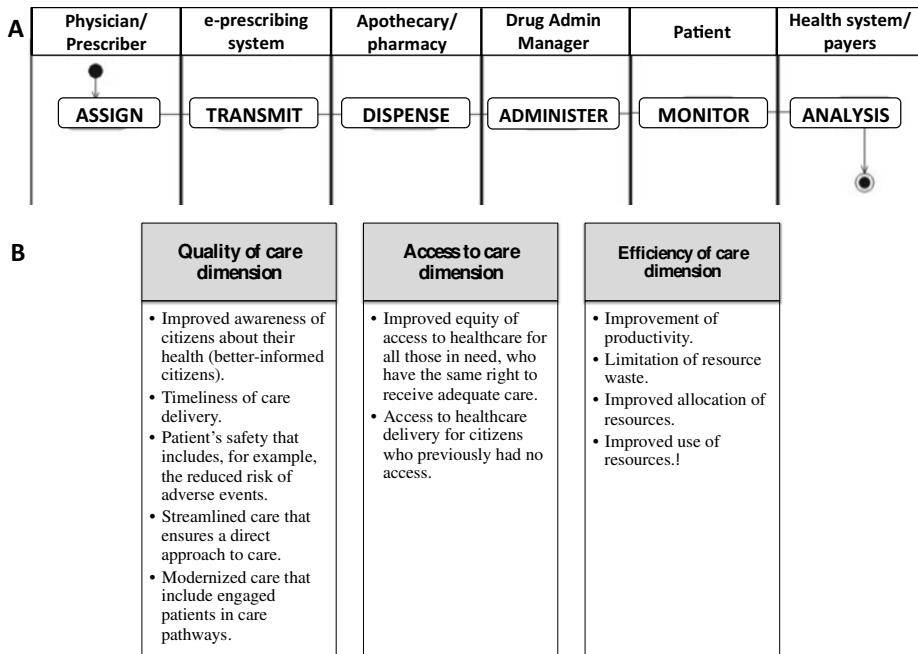


Figure 2. (A) High-level meta-model representing the main process phases.
(B) Expected benefits dimensions from the adoption of e-prescribing (adapted from [3]).

However, the model as it is also represents the basis for metrics definition, as a mean to ensure system quality. As an example, we can consider the “verification actions” identified by the model for the first process phase (“assign phase”). Each of them is associated with a set of benefits that, in terms of GQM (as explained above), which we can consider as the goals identified by the model that require metrics definition.

Table 2 presents a proposal of numeric metrics that can be used to evaluate the quality of both the e-prescribing system under examination and the e-prescribing process itself. For instance, in the case of “drug-drug interactions check”, the benefit measure can be the number of reported ADEs before and after the adoption of the e-prescribing system under examination; in the case of “summary of product characteristics and diagnosis” the adopted metrics can be the number (or %) of prescriptions with reported diagnosis/drug pair in accordance with indications that not only show whether the e-prescribing system is able to support the assignment of drugs according to guidelines but also helps identifying cases of drug misuse.

Table 2. E-prescribing assign phase: verification actions, their associated benefits and metrics.

VERIFICATION ACTIONS IN THE ASSIGN PHASE	BENEFITS FOR QUALITY OF CARE	BENEFITS FOR ACCESS TO CARE	BENEFITS FOR EFFICIENCY OF CARE	POSSIBLE METRICS
Valid patient (patient validation)	Identity error avoided	Ensures patient's existence within the National Healthcare System	Avoided time waste due to erroneous patient's identification	Number (or %) of prescriptions with incorrect, missed or unknown patient ID
Valid exemptions rights		Ensures that the patient has the right of an exemption	Possibility to analyze the relationship between a prescribed drug and a certain exemption, thus preventing possible frauds.	Number (or %) of prescriptions with: <ul style="list-style-type: none"> - Invalid exemption code - Invalid patient ID/exemption code pair - Invalid exemption code /drug code pair
Filled out diagnosis	Ensures that the prescription is the result of a new/previous diagnosis		Possibility to track the relationship between the diagnosis and a specific drug	Number (or %) of prescriptions with: <ul style="list-style-type: none"> - Diagnosis reported - Correctly coded diagnosis reported
Valid drug			Ensures that the drug is included in the official national nomenclature Avoided time waste due to non-existent drug	Number (or %) of prescriptions with valid drug code % of generic drug prescribed vs branded drugs
Drug-drug interaction check	Decreased risk of interactions with drugs already in use by the patient		Possibility to have a more efficient alerting system of drug-drug interactions and ADEs reporting	Number (or %) of prescriptions avoiding drug-drug interactions Number of reported ADEs Number of new ADEs identified
Coherence between summary of product characteristics and diagnosis	Decreased risk of incorrect drug assignment			Number (or %) of prescriptions with reported diagnosis/drug pair in accordance with indications
Valid GP identification			Ensures that the GP is recognized by the healthcare system as having the right to prescribe	Number (or %) of prescriptions with unknown or missed GP ID

In conclusion, in this first case study, the analysis through a process modeling approach was able to (1) highlight what is still missing in existing systems (new tools for the safe and monitored drug administration at home connected to the e-prescribing systems) and (2) evaluating e-prescribing systems and processes by associating metrics to the modelled “verification actions” that represent the goals of e-prescribing in terms of benefits for the healthcare system.

5.2. Health IT systems for supporting the chemotherapy care pathway for cancer patients

Chemotherapy (CHT) is a crucial component of protocol-based care for cancer patients [20]. The process of prescription, preparation (dose calculation), and administration of

CHT is complex, because of the high toxicity of drug, and impacts on patient safety [21]. Errors may be caused by an inappropriate therapy prescribed or delivered, the presence of drug-drug interactions, or an incorrect dosage. Errors may also impact on cancer therapy costs that have been increasing dramatically over the last few years, and, consequently, on the economic sustainability of patient's management for healthcare institutions [22]. The patient's pathway within the hospital (the European Institute of Oncology (EIO), Milan, Italy) is supported by different health IT systems usually included in the hospital information system. However, the development of reliable health IT systems, capable to ensure proper management on the process and to prevent errors, is heavily grounded on the understanding of the underlying process.

The modeling effort, in this case study, aimed to (1) to describe the care pathway involving cancer patients receiving chemotherapies or supportive therapies at the Day Care division for the EIO, and (2) to highlight the critical aspects of the care pathway and, at the same time, to provide possible solutions for them.

The modeling methodology is summarized in Figure 3.

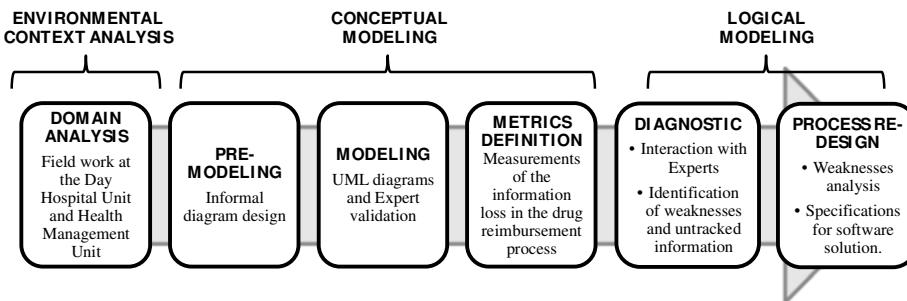


Figure 3. Modeling steps for the evaluation of the chemotherapy care pathways for cancer patients.

The first phase, the Analysis of Environmental Context, mainly consisted of field work, and lasted three months during which clinical and administrative practices, locally applied, were observed in presence (with attention to the clinical records pathway and information technology used). During the Conceptual Modeling, a high-level meta-model (pre-modeling) was used to identify the main activities, their sub-activities, and exceptions. The pre-model was then designed and validated during meetings with medical experts and administrative staff, during which misalignments with the proposed guidelines were also evaluated. The modeling phase involved the creation of UML structural and behavioural diagrams that were again validated both syntactically by the analyst experts and semantically by domain experts. The logical modeling phase, in this case, involved the "Critical Issues Identification", consisting of an analysis of each activity, represented on diagrams, that allowed highlighting process inefficiencies and their causes. From these, solutions that may allow a process reengineering, able to adapt the new models to the ongoing processes, were identified.

The model-based analysis identified the drugs reimbursement flow (called "File F flow") as one of the most critical processes in the patient's pathway. The main observed critical issues were associated with untracked information within the pathway (Figure 4). In fact, the pathway starts with the patient who has a prescription for chemotherapy and is admitted to the ambulatory process. The prescription is used to categorize the patient for the admission regimen and to define the level of reimbursement associated to the patient, and registered in the "file F". However, after

the first tests done by the nurse and the medical exam done by the oncologist specialist, the patient can be assigned to a therapy different from the one prescribed but more suitable for his/her current condition. This can affect the reimbursement and, in turn, the “file F flow”. The conceptual modeling of the process currently implemented highlighted other information loss in the “file F” tracking: the difference between the admission regimen and the prescription, the loss of paper-based documentations for reimbursement request, and the loss of drug information for reimbursement after patient’s visit. The process, in fact, didn’t track the decision-making during the patient’s visit (due to the lack of an appropriate information system) and the documentation running in the patient’s pathway was not updated after establishing the patient’s condition. This produced the lack of association between Reimbursement Rules (in the patient’s electronic health record) and Administered Therapy (in a different paper-based record), and, in turn, no drug reimbursement for the hospital.

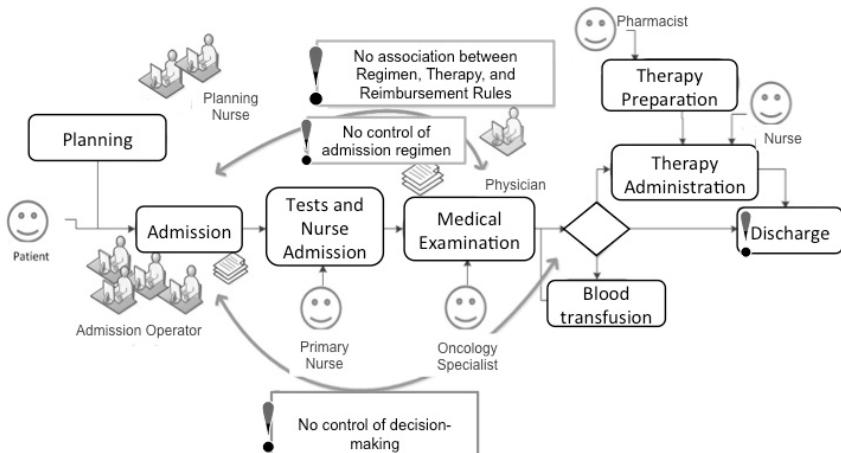


Figure 4. Pre-modeling: high-level UML-like activity diagram representing the care pathway with critical issues identified. Smileys represent process actors; sketches represent the interaction of actors with the available information systems.

Based on these considerations, a new model was created as a solution for the critical issues. The new model provided the technical specifications for the creation of a new module of the hospital information system that allows monitoring and controlling process variables, promoting operator coordination and integration, and optimizing the collaboration between operating units.

The goal of modeling in this case study was to limit the information loss for the drug reimbursement process. Also in this case, it is possible to identify metrics to evaluate whether or not the proposed solution satisfied such expectation. They can be, for example:

- the number of inconsistencies between the expected reimbursements and the obtained reimbursement;
 - the number of incomplete “file F” for the patients treated.

In conclusion, in this case study, the modeling effort was able to provide a full representation of the complex process of the chemotherapy care pathway for cancer patients (with a translational value for hospitals other than the EIO), allowed the identification of the main critical issues underlying process inefficiencies and the

creation of a feasible solution to the identified critical issues, and grounded the definition of metrics to evaluate the efficiency of the proposed solution.

6. Conclusions

This contribution presented the heterogeneity and complexity of health IT services and systems that are a consequence of the heterogeneity and complexity characterizing the healthcare domain. We proposed the perspective of process modeling as a method to break out complexity and represent heterogeneity and to provide tailored evaluation and optimization of health IT systems and services. Process modeling not only provides a way to effectively represent the requirements of a system or service. By also supporting the identification of goals and benefits, it allows the definition of quantitative metrics able to show whether a system is suitable for a specific context, also in terms of economic revenue/savings [23].

Recommended further readings

1. L. Baresi, F. Casati, L. Castano, M. Fugini, P. Grefen, I. Mirbel, B. Pernici, G. Pozzi, *Workflow Design Methodology*, in: Database Support for Workflow Management: The WIDE Project. Springer US, New York, 1999. pp.47-94.
2. S. Marceglia, L. Mazzola, S. Bonacina, P. Tarquini, P. Donzelli, F. Pincioli, A Comprehensive E-prescribing Model to Allow Representing, Comparing, and Analyzing Available Systems, *Methods of Information in Medicine* **52** (2013), 199–219.
3. S. Miksch, D. Riaño, A. ten Teije, Editors, *Knowledge Representation for Health Care*, Springer, Berlin Heidelberg, 2012.
4. C. Combi, E. Keravnou-Papailiou, Y. Shahar, *Temporal Information Systems in Medicine*. Springer US, Boston, MA, 2010.

Food for thought

1. To what extent is modeling needed for defining metrics?
2. Is the personalization of care only a matter of exceptions to a generic model, or is it a specialization of a general model? How do we map personalization when modeling is done for requirement analysis?
3. Is process modeling able to represent the local environment without losing its generalizability? How can we ensure that models are portable in different environments?
4. Does a model designed for a specific context have a predictive value in establishing metrics for specific contexts for which the model has not been specialized yet?

References

- [1] M.R. McGee-Lennon, *Requirements engineering for home care technology*, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2008, April 5–10, 2008, Florence, Italy, ACM, New York, 2008. pp. 1439–1442.
- [2] E.H. Shortliffe, G.O. Barnett, *Biomedical Data: Their Acquisition, Storage, and Use*, in: Biomedical Informatics, Springer-Verlag, London, 2014, pp. 39–66.
- [3] S. Marceglia, L. Mazzola, S. Bonacina, P. Tarquini, P. Donzelli, F. Pinciroli, A comprehensive e-prescribing model to allow representing, comparing, and analyzing available systems, *Methods of Information in Medicine* **52** (2013), 199–219.
- [4] H.S. Barrows, R.M. Tamblyn, *Problem-based learning: an approach to medical education*, Springer, New York, 1980.
- [5] R.J. Scalese, V.T. Obeso, S.B. Issenberg, Simulation Technology for Skills Training and Competency Assessment in Medical Education, *Journal of General Internal Medicine* **23** (2008), 46–49.
- [6] R. Lenz, M. Reichert, IT support for healthcare processes – premises, challenges, perspectives, *Data & Knowledge Engineering* **61** (2007), 39–58.
- [7] G. Kotronoulas, N. Kearney, R. Maguire, A. Harrow, D. Di Domenico, S. Croy, S. MacGillivray, What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials, *Journal of clinical oncology* **32** (2014), 1480–1501.
- [8] M. Ozkaynak, P. Brennan, An observation tool for studying patient-oriented workflow in hospital emergency departments, *Methods of Information in Medicine* **52** (2013), 503–513.
- [9] B. Séroussi, M.C. Jaulent, C.U. Lehmann, Health Information Technology Challenges to Support Patient-Centered Care Coordination, *Yearbook of Medical Informatics* **10** (2015), 8–10.
- [10] P. Lawrence, Ed., *Workflow handbook 1997*, Wiley, Chichester, 1997.
- [11] S. Mohapatra, *Business process reengineering: automation decision points in process reengineering*, Springer, New York, 2013.
- [12] H.-E. Eriksson, *Business modeling with UML: business patterns at work*, John Wiley & Sons, New York, 2000.
- [13] F. Ruiz, F. Garcia, L. Calahorra, C. Llorente, L. Gonçalves, C. Daniel, B. Blobel, Business process modeling in healthcare, *Studies in health technology and informatics* **179** (2012), 75–87.
- [14] G. Pozzi, S. Morasca, A. Antonini, A.M. Ferreira, Software Measures for Business Processes, in: ADBIS 2011 - Research Communications, *CEUR Workshop Proceedings*, Aachen, Germany, 2011. pp. 11–22.
- [15] A. Kazanjian, C.J. Green, Beyond effectiveness: the evaluation of information systems using a comprehensive health technology assessment framework, *Computers in biology and medicine* **32** (2002), 165–177.
- [16] L. Pecchia, F. Crispino, S.P. Morgan, A *Software Tool to Support the Health Technology Assessment (HTA) and the User Need Elicitation of Medical Devices via the Analytic Hierarchy Process (AHP)*, in: The International Conference on Health Informatics, Springer International Publishing, Switzerland, 2014, pp. 292–295.
- [17] DAC Network on Development Evaluation, *Glossary of key terms in evaluation and results based management*, 2002, <http://www.oecd.org/development/peer-reviews/2754804.pdf>, last access 11 February 2016.
- [18] R. van Solingen, V. Basili, G. Caldiera, H.D. Rombach, *Goal Question Metric (GQM) Approach*, in: Encyclopedia of Software Engineering, John Wiley & Sons Inc., Hoboken, NJ, 2002.
- [19] D.S. Bell, S. Cretin, R.S. Marken, A.B. Landman, A conceptual framework for evaluating outpatient electronic prescribing systems based on their functional capabilities, *Journal of the American Medical Informatics Association* **11** (2004), 60–70.
- [20] E. Frei III, K. Antman, *Principles of dose, schedule, and combination chemotherapy*, in: Holland-Frei Cancer Medicine, BC Decker Inc, Hamilton, ON, 2000, pp. 556–568.
- [21] M.N. Neuss, M. Polovich, K. McNiff, P. Esper, T.R. Gilmore, K.B. LeFebvre, L. Schulmeister, J.O. Jacobson, 2013 Updated American Society of Clinical Oncology/Oncology Nursing Society Chemotherapy Administration Safety Standards Including Standards for the Safe Administration and Management of Oral Chemotherapy, *Oncology Nursing Forum* **40** (2013), 225–233.
- [22] J. Phillips, S. Beam, A. Brinker, C. Holquist, P. Honig, L.Y. Lee, C. Pamer, Retrospective analysis of mortalities associated with medication errors, *American journal of health-system pharmacy* **58** (2001), 1835–1841.
- [23] KLAS, *Ultrasound 2014: Saving Time and Money with Workflow Automation*, 2014.

Evidence-based Health Informatics Frameworks for Applied Use

Catherine K. CRAVEN^{a,1}, Brad DOEBBELING^b, Dominic FURNISS^c,
Richard J. HOLDEN^d, Francis LAU^e, Laurie L. NOVAK^f

^a*Informatics Institute, University of Missouri, Columbia, USA*

^b*School for the Science of Health Care Delivery, Arizona State University, Phoenix,
USA*

^c*UCL Interaction Centre, University College London, London, UK*

^d*School of Informatics and Computing, Indiana University-Purdue University
Indianapolis, Indianapolis, USA*

^e*School of Health Information Science, University of Victoria, Victoria, Canada*

^f*Department of Biomedical Informatics, Vanderbilt University, Nashville, USA*

Abstract. Health Informatics frameworks have been created surrounding the implementation, optimization, adoption, use and evaluation of health information technology including electronic health record systems and medical devices. In this contribution, established health informatics frameworks are presented. Important considerations for each framework are its purpose, component parts, rigor of development, the level of testing and validation it has undergone, and its limitations. In order to understand how to use a framework effectively, it's often necessary to seek additional explanation via literature, documentation, and discussions with the developers.

Keywords. Medical informatics, frameworks, models theoretical, evaluation studies as a topic, qualitative research, implementation science.

1. Introduction

Academic disciplines create frameworks that characterize, describe, guide, analyze, and evaluate phenomena and processes. For example, the field of management, according to a 2015 Harvard Business Review article, has created 81 frameworks for management strategy between 1958 and 2013 [1]. Some of the more familiar examples include Gap Analysis (1965), SWOT (Strengths, Weaknesses, Opportunities, Threats) Analysis (1969), and Disruptive Innovation (1999) [1]. Nursing has created many frameworks such as the Nursing Process Model (1961) [2], Modeling & Role Modeling (1983) [3], and Nursing as Informed Caring for the Well-Being of Others (1993) [4]. In health informatics, frameworks have been created surrounding the implementation, optimization, adoption, use and evaluation of health information technology including electronic health record systems (EHR) and medical devices.

A common question is what exactly is a framework? Is it the same thing as a theory, a theoretical or conceptual model, a theoretical framework, or something

¹ Corresponding author: Catherine K. Craven, PhD, MLS, MA, PO Box 7702, Columbia, MO, 65205, catherine.craven@gmail.com.

distinct? At first glance across terms, definitions, and disciplines, U.S. Supreme Court Justice Potter Stewart's famous 1964 words on "obscenity" seems to apply: It's difficult to pinpoint the definition, but "I know it when I see it" [5]. What is agreed upon is that these varying terms are often used interchangeably, a practice which "has created confusion among scholars and practitioners [6]" [7, 8].

In his 2015 article "Making sense of implementation theories, models, and frameworks" [8], Per Nilsen provided a selective review of key theories, models, and frameworks used in implementation science. Implementation science encompasses and applies to health information technology, but it is defined more broadly than in health informatics, as the "scientific study of methods to promote the systematic uptake of research findings and other EBPs (evidence-based practices) into routine practice to improve the quality and effectiveness of health services and care" [8]. Certainly, these are the goals of evidence-based informatics research and evaluation studies, and informatics-based, evidence-driven health IT implementations.

Nilsen acknowledges that the terms theory, model, and framework are often used interchangeably, and explains that theories tend to be viewed, across disciplines, in terms of high-, mid-, and low-levels – "an abstraction continuum." However, he attempts to delineate the terms, and defines theories as "a set of analytical principles or statements designed to structure our observations, understanding, and explanation of the world ... usually comprising "definitions of variables, a domain where the theory applies, a set of relationships between the variables, and specific predictions" [8]. Nilsen says that models often involve a deliberate simplification of a phenomena or its aspects, have value "without having completely accurate representations of reality," and can be described as "theories with a more narrowly defined scope of explanation; a model is descriptive while a theory is explanatory as well as descriptive" with some predictive capacity [8].

Frameworks, however, do not provide explanations, but "describe empirical phenomena by fitting them into categories. Frameworks usually denote a structure, overview, outline, system, or plan, consisting of various descriptive categories, e.g. concepts, constructs, or variables, and the relations between them that are presumed to account for a phenomena" [8]. Nilsen adds that models and frameworks in implementation science do not specify the mechanisms of change. "They are typically more like checklists of factors relevant to various aspects of implementation, frameworks often have a descriptive purpose by pointing to factors believed or found to influence implementation outcomes." [8]. Nilsen' proposes three overarching aims of all theoretical approaches in implementation science, and five categories of approaches to achieve these aims [8] (See table 1).

In the next section, we present two health informatics frameworks for discussion purposes. The first, DiCoT-CL, is used for guiding implementation, evaluation, and use-optimization of medical devices, and the sociotechnical systems in which they are used. According to Nielsen's five categories, DiCoT-CL is a process model or framework, of the action sub-type. The second framework, the Clinical Adoption Framework, is an evaluation framework used to evaluate health IT adoption, particularly electronic health record system (EHR) adoption, in healthcare organizations from a sociotechnical perspective. For each of these frameworks, its purpose, component parts and development, testing and validation, limitations, and a basic explanation for how the framework is employed are discussed.

Table 1. Nilsen's Three Overarching Aims of All Theoretical Approaches and Five Categories of Theories, Models, and Frameworks Used In Implementation Science.

Category	Description
AIM 1	Describing and/or guiding the process of translating research into practice (process models)
Process Models	Specify steps (stages, phases) in the process of translating research into practice, including the implementation and use of research. The aim of process models is to describe and/or guide the process of translating research into practice. An action model is a type of process model that provides practical guidance in the planning and execution of implementation endeavors and/or implementation strategies to facilitate implementation. Note that the terms “model” and “framework” are both used, but the former appears to be the most common
AIM 2	Understanding and/or explaining what influences implementation outcomes (determinant frameworks, classic theories, implementation theories)
Determinant Frameworks	Specify types (also known as classes or domains) of determinants and individual determinants, which act as barriers and enablers (independent variables) that influence implementation outcomes (dependent variables). Some frameworks also specify relationships between some types of determinants. The overarching aim is to understand and/or explain influences on implementation outcomes, e.g. predicting outcomes or interpreting outcomes retrospectively
Classic Theories	Theories that originate from fields external to implementation science, e.g. psychology, sociology and organizational theory, which can be applied to provide understanding and/or explanation of aspects of implementation
Implementation Theories	Theories that have been developed by implementation researchers (from scratch or by adapting existing theories and concepts) to provide understanding and/or explanation of aspects of implementation
AIM 3	Evaluating implementation (evaluation frameworks)
Evaluation Frameworks	Specify aspects of implementation that could be evaluated to determine implementation success

2. Health Informatics Framework Example 1: DiCoT Concentric Layers Framework (DiCoT-CL)

2.1. Purpose of the DiCoT Concentric Layers Framework (DiCoT-CL)

Evaluating medical devices and a health IT in context is challenging. Technology is influenced by and influences the workflows, social settings, organizational contexts it is embedded within; also artefacts and equipment around it can impact its effectiveness and use. Further, it can be influenced by training, procurement, policy and technical configuration decisions that happen far away from its actual use. The DiCoT Concentric Layers framework (DiCoT-CL) [9] is a framework for investigating these issues.

DiCoT was a precursor to the DiCoT-CL framework. DiCoT (Distributed Cognition for Teamwork) [10, 11] facilitates the use of Distributed Cognition for analyzing sociotechnical systems. Distributed Cognition [12] focuses on the transformation and propagation of information in sociotechnical systems. The DiCoT

Concentric Layers framework (DiCoT-CL) [9] builds on DiCoT by focusing on how technology is coupled to different layers of sociotechnical context. DiCoT and DiCoT-CL help analysts investigate the underlying information architecture of a sociotechnical system, within which a technology is embedded.

DiCoT and DiCoT-CL have four proposed outputs: an understanding of the basic mechanics of the system, opportunity for deeper conceptual insight into the system, recognition of incremental design considerations, and more revolutionary design considerations [13]. For DiCoT-CL, Furniss et al. [9] argue that further insights can be gained by looking within and between the concentric layers of the sociotechnical system. Also, the framework can help provide micro-level insight (e.g. specific issues at the interface) and macro-level insight (e.g. problems with the way the device was configured when it was purchased months or years previously). The ultimate purpose of DiCoT-CL is to identify issues and make recommendations for improving the technology and the sociotechnical system it is embedded within.

2.2. Component Parts and Development of the DiCoT-CL Framework

The beginning of DiCoT stems from Furniss' master's thesis in 2004: *Codifying Distributed Cognition: A case study of emergency medical dispatch*. The output of this research was an analysis of the London Ambulance Service control room using Distributed Cognition [11] and the DiCoT method [10]. Distributed Cognition is promising for the design and evaluation of technology in practice. However, it has not been adopted as widely as one might expect. Some believe that this is due in part to a lack of an off-the-shelf method and analytical support. DiCoT helps to fill this gap. Furniss and Blandford (who supervised the earlier thesis) have continued work on DiCoT together and separately with master's and doctoral students, and involving external research teams. Of particular note is Rajkomar's 2014 PhD thesis summarized in [14]. He proposed further details on how tasks are distributed over time and how this impacts distributed cognition [15]. DiCoT has also been applied in intensive care [16] and medical equipment library design [17].

The critical breakthrough for creating DiCoT came from combining the theoretical literature on Distributed Cognition with the methodological structure and advice from Contextual Design [18]. The idea of analyzing the sociotechnical system through creating interdependent models of the context came from Contextual Design, but the models were adapted to suit the themes that occurred in Distributed Cognition.

DiCoT has five main models: the information flow model, the artefact model, the physical model, the social model, and the evolutionary model. Each model has associated principles that have been distilled from the Distributed Cognition literature. These principles guide analysts to reflect on aspects of Distributed Cognition in data gathering and analysis. Questions that arise through reflection might include, for example, the following: Is there an "information buffer" that holds information for later use? What processes filter and change information? Is "situation awareness" good and why? How does the 'physical arrangement of equipment' impact information processing?

DiCoT-CL [9] was developed relatively recently by Furniss, Blandford, and others. It adds concentric layers to the original DiCoT framework, so that layers of sociotechnical system can be analyzed around a technology, e.g. a device and user at the center, then the device's use at the bedside, then its use at the ward level, then at the hospital level. Furniss performed an analysis of the design and use of a modern in-

patient blood glucose meter [19]. The focus on evaluating this medical device made it apparent that it was coupled to different layers of context. Conceptually this resonated with Grudin's [20] view of the computer reaching out. Furniss et al. apply this idea to conceptualize the medical device reaching out from interface issues between the device and the user, to issues at the bedside, to team issues at the ward level, to management issues at the hospital level [9]. DiCoT-CL adds concentric layers to DiCoT's five models. Figure 1 shows the different layers around the user-device interaction at the center, how each layer is divided into five segments, and where features of a sociotechnical system appear in the framework.

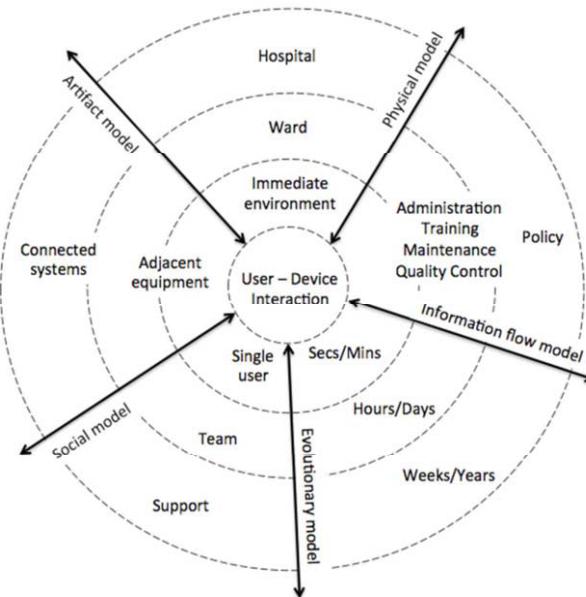


Figure 1. The DiCoT Concentric Layers (DiCoT-CL) Framework.

There are different centric layers around the user-device interaction, which is at its core. Each pie-shaped segment represents a different model. From the top, moving clockwise round, we have the physical model, the information flow model, the evolutionary model, the social model and the artefact model (reproduced from [9]).

DiCoT and DiCoT-CL have been built up through successive case studies. These case studies have mainly involved fieldwork, in which the design and use of technology has been evaluated in context, using observations and interview data. In each case study, the analyst who applies the framework often not only reflects on the results, but also on the applicability of the framework. Sometimes there is reason to add to the framework, e.g. an extra theme and more principles (e.g. [15]), and sometimes the emerging data and theory suggest new forms for the framework, such as the addition of concentric layers in DiCoT-CL.

2.3. Testing and Validation of the DiCoT-CL Framework

Development and testing of the framework has been iterative. Berndt et al. [21] report a case study that compares the learning and application of Contextual Design with DiCoT in the same setting, i.e. information flow in anaesthesia. Their results suggested that Contextual Design was easier to learn, but DiCoT encouraged deeper insights in this complex setting. Others have used DiCoT successfully in different contexts. For example, it has been used to analyze the work of agile software development teams [22], and the processing of patients within a hospital [23]. DiCoT-CL is a new development and has only been used in one case study [9, 19]. Further case studies will be developed to test the addition of the concentric layers to DiCoT. Future work will review original DiCoT principles to ensure they are well-structured and comprehensive, develop the social and evolutionary models, and provide training materials for the framework.

2.4. Example of How to Apply the DiCoT-CL Framework

When first engaging with complex sociotechnical systems, it is easy to get overwhelmed with information, particularly when new to the system. DiCoT helps to guide the analyst on where to focus data collection. The following example describes how to apply the framework in the field via an example evaluation of a modern inpatient blood glucose meter in an oncology ward [9]. This walkthrough is broken into three stages, and employing the framework will depend on familiarity with these stages.

In the first stage, the author [DF] shadowed a nurse to see what she did in relation to the blood glucose meter. She picked up the blood glucose meter reader, retrieved a case with its other paraphernalia inside, and started to do a quality-control check. The author noted down the detailed steps of this process and the equipment used as best he could in field notes, while asking questions at opportune times. This stage revealed preliminary task steps for an information flow model and notes on equipment use for the artefact model. The author then followed the nurse to do a blood glucose meter reading with a patient, and similarly, made careful observations and notes. Finally, over successive observations, the author observed, asked questions, and gathered more information and filled in these partial models and descriptions.

As a beginner one can work through the five DiCoT models to develop a description and schematic diagrams, e.g. an information flow diagram, a sketch of the device's interface, and the layout of equipment around a patient's bed, while reflecting on how this configuration of the system impacts its effectiveness and whether it could be improved. These models are developed iteratively. Through each iteration, describing the system via the models reveals gaps in understanding. Sketches will generate new questions. Further observations will reveal new issues, and the principles will encourage the analyst to think in different ways. DiCoT-CL will reveal areas where data is lacking. As the complexity of the picture builds up, intricate dependencies emerge between the models, which challenge the idea of a decomposition into separate models as in the first stage.

For example, an observation of a healthcare assistant lending a student nurse his or her personal barcode to use the blood glucose meter touches on the social, artefact, and information flow model. So, in which component model does this go? It doesn't really matter at first. What matters more is that this part of the process is noted and included

somewhere to start with; the models collectively build a picture rather than any one standing alone.

Once the analyst has more of a grasp of the framework in the context, she or he can refine and formalize the models. For example, DiCoT-CL revealed that the author had not applied the information flow model at the ward level, but what did this mean for the glucometer evaluation? Healthcare assistants had been observed writing bed numbers they had to attend to on tissue paper and cardboard trays as part of a blood glucose meter 'round' – and this seemed to belong within the information flow model. This highlighted that the device only supported single glucose readings; it did not support the user in doing multiple readings across the ward; such functionality could be a future design consideration. Here a reflective conversation between the data, the models, the principles, and within and between layers of DiCoT-CL can help drive new insights. Tensions between the data and the framework could also lead to its development, as noted in [15] above.

2.5. Limitations of the DiCoT-CL Framework

DiCoT-CL emphasizes the complex connections that a device or technology has with the context in which it is embedded. Therefore, there is some tension between emphasizing the context-dependent nature of a device within a specific context, and trying to evaluate its performance across different contexts. When evaluating technology across contexts the significant context-dependent features that impact the design and use of a device need to be recognized and managed.

3. Health Informatics Framework Example 2: Clinical Adoption Framework (CAF)

3.1. Purpose of the Clinical Adoption Framework (CAF)

The Clinical Adoption Framework (CAF) is a conceptual framework used to evaluate health IT adoption in healthcare organizations from a sociotechnical perspective [24].² The CAF represents health IT adoption as having three interrelated dimensions at the micro, meso, and macro levels. At each level, there is a feedback loop that can lead to further changes from the effects of the initial adoption. There is also a feedback loop across levels such that the adoption and effects at one level can influence the other levels. A basic premise of the CAF is that health IT adoption and its effects are not deterministic because they are dependent on the dynamic interplay of the factors within and across the three dimensions over time. Figure 2 shows CAF (source: <http://ehealth.uvic.ca/methodology/models/CAF.php>).

² See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

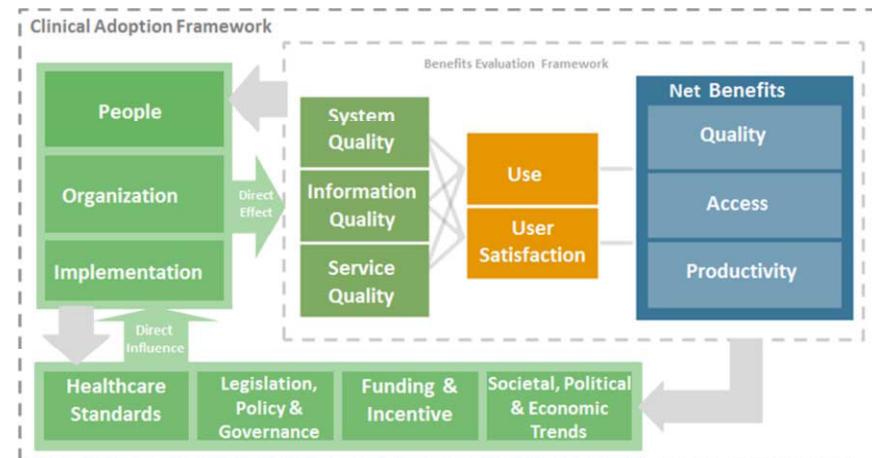


Figure 2. Clinical Adoption Framework.

3.2. Component Parts and Development of the Clinical Adoption Framework

The CAF is an extension of the Infoway Benefits Evaluation (BE) Framework that takes into account the contextual factors which influence health IT adoption. The BE Framework was created by Lau, Hagens & Muttitt (2007) to describe health IT adoption at the micro level by focusing on the health IT quality, its use and satisfaction, and net benefits [25], for Canada Health Infoway, a non-profit organization funded by the Canadian governments to accelerate the deployment of interoperable electronic health record systems (EHR) and ehealth solutions.

The BE Framework is an adaptation of the well-known Information Systems (IS) Success Model created by DeLone & McLean (2003) for business organizations [26]. One shortcoming of the IS Success Model is that it does not address the socio-organizational aspects. To account for these contextual factors, the CAF incorporated the meso and macro level dimensions with key measures from the Information Technology Interaction Model by Silver, Markus, & Beath (1995), the Unified Theory of Acceptance and Use of Technology Model by Venkatesh (2003), the Organizational Change Management Model by Kotter (1995) and the Health IT Risk Assessment Model by Pare, Sicotte, Jaana, & Giroard (2008) [27, 28, 29, 30]. The micro, meso, and macro dimensions of the CAF, the categories of measures in each of these dimensions, and an explanation of to what these measures refer are briefly described below in Table 2. Detailed explanation of the dimensions and measures are in Lau, Price, & Keshavjee (2009) [24].

Table 2. Dimension Levels and Measures of the Clinical Adoption Framework.

Dimension Level	Categories of Measures in the Dimension	Explanation of Measures
Micro	Quality of health IT	Accuracy, completeness and timeliness of the information, performance and security of the system, and responsiveness of the support services
	Use of health IT	Intended/actual health IT usage, user competency, and satisfaction in usefulness and ease of use
	Net benefits of health IT	Care quality in safety, appropriateness and effectiveness, access to care through provider/patient participation and service availability, and productivity in care coordination, efficiency and net cost
Meso	People	Individuals/groups, their characteristics and expectations, and roles and responsibilities related to health IT adoption
	Organization	The fit between health IT and the organization's strategy, culture, infrastructures, processes, and value.
	Implementation	Implementation refers to health IT adoption stages, project-management approaches, and extent of health IT-practice fit
Macro	Governance	Roles of governing bodies, legislations, and advocacy groups on health IT
	Funding	Remunerations, payments and incentives that influence health IT adoption
	Standards	health IT, organizational performance, and professional practice standards in place
	Trends	Public expectations, and socioeconomic and political influence on health IT

3.3. Testing and Validation of the Clinical Adoption Framework

The CAF underwent three testing/validation steps during its initial development. In the first of these, in 2009, Infoway held a *consultation session* with 23 health IT practitioners from across Canada to invite feedback on the CAF. The practitioners responded to whether the framework made sense, if concepts were missing or needed revisions, as well as their interest and effort needed to apply the framework in their organization. Based on the feedback, revisions were made to streamline the framework into its current form (Charlebois 2009) [31].

In the second, Oh (2009) [32] compared the CAF measures against 16 *published survey instruments*. They included 13 instruments from the Health IT Survey Compendium section of the Agency for Healthcare Research & Quality (AHRQ) Health IT website (AHRQ, 2010) [33] and three from Canada Health Infoway. Of the 16 instruments examined, only the Infoway System and Use Assessment Survey items mapped to all 20 micro-level measures. At the meso level the 16 instruments mapped between 0 and 11/12 of the measures. At the macro level they mapped poorly from 0 to 5/12 measures. No question items were found missing from the CAF which suggested it was sufficiently comprehensive for all aspects of HIT.

In the third, in a *meta-review* of 50 systematic reviews of health IT evaluation studies published in 1995-2008, Lau, Price, Kuziemsky & Gardner (2010) [34] mapped most of the evaluation measures from the published reviews to the micro-level of the CAF. They also identified measures that did not fit the micro level and created new categories which were patient/provider, implementation, incentive, policy/legislation, change improvement and interoperability. These factors mapped nicely under the meso and macro dimensions of the CAF.

The testing/validation results showed CAF has face validity as a multi-dimensional scheme. Therefore, CAF can be used to describe, understand and evaluate health IT adoption and its effects in healthcare organizations. Since its publication in 2009, the CAF has been applied, adapted, or mentioned in over 30 health IT related studies.

3.4. Example of How to Apply the Clinical Adoption Framework

The following example is to further demonstrate how the framework could be applied. The CAF was applied in a six-month post-implementation study of an electronic health record system (EHR) in two ambulatory clinics managed by a health region in a Canadian province [35]. The implementation of the EHR in these clinics represented the initial phase of a long-term plan by the health region to adopt EHRs in all of its ambulatory clinics throughout the region. The purpose of the study was to evaluate the impact of EHR adoption on the organization in order to guide subsequent implementation effort.

Four university researchers conducted the study over six weeks. They used a rapid evaluation method to examine selected micro and meso components of the CAF. The selected CAF components were deemed relevant and feasible by the researchers and clinic/IT executives given the stage of EHR adoption effort at the time. At the micro level, CAF components covered the EHR system, information and service quality, EHR usage and user satisfaction, and net benefit in terms of EHR-supported care coordination and efficiency were examined. In particular, system quality covered EHR functionality and usability; information quality covered EHR data accuracy, completeness and consistency; service quality covered EHR staff knowledge and responsiveness; and usage covered actual EHR use and its perceived usefulness. At the meso level, CAF components covered people, organization and implementation aspects for the clinics involved were examined. People covered clinic and EHR staff roles, expectations and experiences. Organization covered EHR-health IT infrastructure, strategy and process. Implementation covered EHR deployment process and EHR-practice fit.

The rapid evaluation method is a pragmatic field evaluation approach developed by the researchers as part of their eHealth evaluation research program. The method consisted of an EHR adoption survey, user assessment, usability/workflow analysis, document review, project risk assessment, data quality review, and group reflection [36]. Data collection took place over four weeks that included concurrent review of project documents and EHR data. EHR support staff organized interviews, assembled relevant documents, and extracted EHR data for the researchers. Notes taken during the interview, usability/workflow and focus group sessions were summarized and analyzed for common themes. The evaluation report was finalized in the last two weeks of the study.

Forty-three participants took part in the study that included clinicians and support staff from the two clinics, EHR support staff and health region executives involved with the project. Over four weeks the researchers completed 12 EHR adoption surveys, 14 usability/workflow sessions, 13 user assessment interviews, 11 project risk assessment interviews, 3 focus group sessions, and reviewed 65 project documents and 3 months of EHR data.

The study found that clinic staff perceived benefits in EHR-supported care coordination and efficiency, despite challenges stemming from early suboptimal deployment decisions surrounding EHR configurations, user training, clinic workflow,

data quality assurance, and data exchange with the regional EHR, which negatively impacted clinical work. For example, during the study, clinicians had to work with fragmented charts because some clinical documents were stored in the regional EHR, which required separate logins. The EHR had no mechanism to indicate whether a document was available or where it could be found. As a result, clinicians had to create workarounds that led to inconsistent EHR use. The researchers emphasized that the study represented only one point-in-time after the EHR was implemented in the clinics. Therefore, the attitude of the clinic staff toward the EHR could change over time if and when the identified issues were resolved. Overall, the CAF had proved useful in making sense of ways that EHR could add value to the organization.

3.5. Limitations of the Clinical Adoption Framework

The CAF is a complex scheme with multiple dimensions, categories, and measures that can be difficult to understand and apply in practice. More work is needed to explain and refine the respective components in ways that are relevant to practitioners involved with health IT adoption and evaluation. Second, there is little guidance available on how one should apply the CAF when studying health IT adoption. Having a how-to guide on the types of study methods and measures that can be used to examine health IT in a specific setting could facilitate its uptake in practice. Third, the CAF is new and has only been applied in a limited number of evaluation studies thus far. To be credible more studies are needed to demonstrate its validity and utility across different settings.

4. Conclusion

In the practice of evidence-based health informatics, the development of a framework, as well as its use in a live setting for real-world purposes must be conducted rigorously. A mix of expert consensus and some empirical observations rather than theory may be the basis for a new framework. Or vice versa. However, the important questions are what sort of expertise and how many experts were involved in its development? How many direct observations were made, in how many iterations, and in how many settings? Were validated theories or process models employed in the development of its components, as it was iteratively developed? How mature is the framework: How many times has it been put to the test in the field to guide the process that it purports to describe? Or, has it been used retrospectively to evaluate the completeness and/or success of that process? What were the outcomes of these efforts?

By their nature frameworks can be rigorously developed, yet how to employ them – where to start and what to do – is not always clear-cut without additional explanation or guidance materials. In addition to reading available literature and documentation, a suggestion is to contact the framework's developers. Request a discussion about the purpose of the framework and its parts to ensure that it is useful for the intended purpose, in the context in which it is to be applied, and how to use it effectively.

Recommended further readings

1. Ash JS, Stavri PZ, Kuperman GJ, A consensus statement on considerations for a successful CPOE implementation, *J Am Med Inform Assoc* **10**(3) (2003), 229-34. Note: A framework for computerized provider order entry implementation or retrospective evaluation. Could also be used as a framework to guide implementation of integrated EHRs, which include CPOE.
2. Holden RJ, Carayon P, Gurses AP, Hoonakker P, Hundt AS, Ozok AA, et al. SEIPS 2.0: A human factors framework for studying and improving the work of healthcare professionals and patients, *Ergonomics* **56**(11), (2013), 1669-86. Note: A framework that represents work system structure, process, and outcomes. Used to evaluate specific existing work systems, plan work system (re)design, and structure research data collection and analysis.
3. V.L. Patel, T.G. Kannampallil, R.R. Kaufman (Eds.) *Cognitive Informatics for Biomedicine: Human Computer Interaction in Healthcare*, Springer, Switzerland, 2015. Note: Addresses gaps on the applicability of theories, models, and evaluation frameworks of human computer interaction (HCI) and human factors for research in biomedical informatics.

Food for thought

1. Which classic theories are employed in development of the frameworks presented?
2. Could any of the frameworks presented here be considered Implementation Theories according to Nilsen's definition? Why?
3. Are you familiar with other health IT frameworks? For what purpose are they used, and how would you classify them according Nilsen's five categories? Why?

References

- [1] M. Reeves, K. Haanaes, J. Sinha, Navigating the dozens of different strategy options. *Harvard Business Review*, June 24, 2015. Online excerpt from *Your Strategy Needs a Strategy: How to Choose and Execute the Right Approach*. Harvard Business Review Press. 2015. Found at: <https://hbr.org/2015/06/navigating-the-dozens-of-different-strategy-options>, last access 11 February 2016..
- [2] I.J. Orlando, *The Dynamic Nurse-Patient Relationship: Function, Process and Principles*, G. P. Putman's Sons, New York, 1961.
- [3] H.C. Erickson, E. Tomlinson, M. Swain, *Modeling and Role Modeling: A Theory and Paradigm for Nursing*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
- [4] K.M. Swanson, Nursing as informed caring for the well-being of others, *Image: The Journal of Nursing Scholarship* **25**(4) (1993), 352-357.
- [5] U.S. Supreme Court Justice 378 U.S. 184 (1964).
- [6] P. Higgins, S. Moore, *Levels of theoretical thinking in nursing*, in Perspectives on Nursing Theory. 5th edition, in: P.G. Reed, B. Nelma, editors, Crawford Shearer, Wolters Kluwer Health, 2009, pp 49-56.
- [7] J.W. Kenny, *Relevance of theoretical approaches in nursing practice*, in: P.J. Christensen, J.W. Kenney, editors, *Nursing Process: Application of Conceptual Models*, 3rd edition, C.V. Mosby Company, St. Louis, 1990. pp 3-19.
- [8] P. Nilsen, Making sense of implementation theories, models and frameworks, *Implementation Science*, **10**(53) (2015), 1-14.
- [9] E. Berndt, D. Furniss, A. Blandford, Learning Contextual Inquiry and Distributed Cognition: a case study on technology use in anaesthesia, *Cognition, Technology & Work* **17**(3) (2014), 1-19.
- [10] H. Beyer, K. Holtzblatt, *Contextual Design*, Morgan Kaufmann, San Francisco, 1997.

- [11] A. Blandford, D. Furniss, *DiCoT: a methodology for applying Distributed Cognition to the design of team working systems*, in: S.W. Gilroy, M.D. Harrison (Eds.), *Interactive Systems, Design, Specification, and Verification*, 12th International Workshop, DSVIS 2005, Lecture Notes in Computer Science 3941, Springer, pp 26-38.
- [12] D. Furniss, A. Blandford, Understanding emergency medical dispatch in terms of Distributed Cognition: a case study, *Ergonomics* **49**(12-13) (2006), 1174-1203.
- [13] D. Furniss, A. Blandford, *DiCoT Modeling: From Analysis to Design*, in: Proceedings CHI 2010 Workshop, April 10-15, Atlanta, Georgia, Bridging the Gap: Moving from Contextual Analysis to Design, 2010.
- [14] D. Furniss, P. Masci, P. Curzon, A. Mayer, A. Blandford, Exploring medical device design and use through layers of distributed cognition: how a glucometer is coupled with its context, *Journal of Biomedical Informatics* **53** (2014), 330-41.
- [15] D. Furniss, P. Masci, P. Curzon, A. Mayer, A. Blandford, 7 Themes for guiding situated ergonomic assessments of medical devices: A case study of an inpatient glucometer, *Applied ergonomics* **45**(6) (2014), 1668-1677.
- [16] J. Grudin, *The computer reaches out: the historical continuity of interface design*, in: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, 1990. pp. 261-268.
- [17] E. Hutchins, *Cognition in the Wild*, MIT Press, 1995.
- [18] J. McKnight, G. Doherty, *Distributed cognition and mobile healthcare work*, in: Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction-Volume 2, British Computer Society, 2008. pp. 35-38.
- [19] A. Rajkomar, A. Blandford, Understanding infusion administration in the ICU through Distributed Cognition, *Journal of Biomedical Informatics* **45**(3) (2012), 580-590.
- [20] A. Rajkomar, A. Blandford, A. Mayer, Coping with complexity in home hemodialysis: a fresh perspective on time as a medium of Distributed Cognition, *Cognition, Technology & Work* **16**(3) (2014), 337-348.
- [21] A. Rajkomar, A. Mayer, A. Blandford, Understanding safety-critical interactions with a home medical device through Distributed Cognition, *Journal of Biomedical Informatics* **56** (2015), 179-94.
- [22] H. Sharp, H. Robinson, Collaboration and co-ordination in mature eXtreme programming teams, *International Journal of Human-Computer Studies* **66**(7) (2008), 506-518.
- [23] J. Werth, D. Furniss, *Medical equipment library design: Revealing issues and best practice using DiCoT*, in: Proc. International Health Informatics Symposium (IHI 2012), Miami, Florida, Jan 28-30, 2012.
- [24] F. Lau, M. Price, K. Keshavjee, From benefits evaluation to clinical adoption: making sense of health information system success in Canada, *Healthcare Quarterly* **14**(1) (2011), 39-45.
- [25] F. Lau, S. Hagens, S. Muttitt, A proposed benefits evaluation framework for health information systems in Canada, *Healthcare Quarterly* **10**(1) (2007), 112-228.
- [26] W. DeLone, E. McLean, The DeLone and McLean model of information system success. A ten year update, *Journal of Management Information systems* **10**(4) (2003), 9-30.
- [27] M.S. Silver, M.L. Markus, C.M. Beath, The information technology interaction model: A foundation for the MBA core course, *MIS Quarterly* **19**(3) (1995), 361-90.
- [28] V. Venkatesh, M.G. Morris, G.B. Davis, F.D. David, User acceptance of information technology. Toward a Unified view, *MIS Quarterly* **27**(3) (2003), 425-478.
- [29] J.P. Kotter, Leading change: why transformation efforts fail, *Harvard Business Review*, Jan 2007, 59-67.
- [30] G. Pare, C. Sicotte, M. Jaana, D. Girouard, *Prioritizing clinical information system project risk factors: a delphi study*, in: Proceedings of the 41st Hawaii International Conference on System Science. Waikoloa, HI. 2008.
- [31] M. Charlebois, Personal communication - *Validating the Clinical Adoption Framework responses*, Nov 10. 2009.
- [32] J. Oh, *Clinical Adoption Framework – Survey Mapping*, Dec 4, 2009 (Unpublished).
- [33] AHRQ, *Health IT Survey Compendium*, 2010, <http://healthit.ahrq.gov/health-it-tools-and-resources/health-it-survey-compendium>, last access 11 February 2016.
- [34] F. Lau, M. Price, C. Kuziemsky, J. Gardner, A review on reviews of health information system studies, *Journal of American Medical Informatics Association* **17** (2010), 637-645.
- [35] F. Lau, C. Partridge, G. Randhawa, M. Bowen, *Applying the clinical adoption framework to evaluate the impact of an ambulatory electronic medical record*, in: K.L. Courtney et al., *Enabling Health and Healthcare through ICT*, IOS Press, Amsterdam, 2013. pp. 15-20.
- [36] F. Lau, *EMR Evaluation Toolkit v3.0*, 2010, <http://ehealth.uvic.ca/resources/tools/tools.php>, last access 11 February 2016.

Ensuring the Quality of Evidence: Using the Best Design to Answer Health IT Questions

Charlene R. WEIR^{a,1}

^a*Department of Biomedical Informatics, University of Utah*

Abstract. The quality of logic in a research design determines the value of the results and our confidence regarding the validity of the findings. The purpose of this contribution is to review the principles of research design as they apply to research and evaluation in health IT. We review the architecture of research design, the definitions of cause, sources of bias and confounds, and the importance of measurement as related to the various types of health IT questions. The goal is to provide practitioners a roadmap for making decisions for their own specific study. The contribution is organized around the Threats to Validity taxonomy and explains how different design models address these threats through the use of blocking, factorial design, control groups and time series analysis. The contribution discusses randomized experiments, and includes regression discontinuity designs and various quasi-experimental designs with a special emphasis on how to improve pre/post designs. At the end, general recommendations are provided for improving weaker designs and general research procedures.

Keywords. Research design, reproducibility of results, program evaluation, social validity, research.

1. Introduction

The quality of logic in a research design is the unit of exchange in the world of science. The design determines the value of the results and our confidence in the findings for consumers of health information technology research. In other words, the design not only determines the nature of the evidence, it *is* the evidence.

Each field is unique and the types of design, the procedures used for research and the rules for accepting evidence vary. Those studying the role of drugs on disease have a different scope and level of control as compared to those who study the impact of reimbursement policies on patients seeking healthcare. Similarly, health IT research has a unique level of analysis and scope of concern requiring specialized approaches to research design.

For this contribution, the implicit assumption is that the manipulation of the information environment changes (*causes* a change) in providers' decisions, patients' outcomes or both. The purpose of this contribution is to review the principles of research design as they apply to research and evaluation in health IT. We will review

¹ Corresponding author: Prof. Dr. Charlene Weir, Department of Biomedical Informatics, University of Utah, School of Medicine, 421 Wakara Way, Salt Lake City, U.S. Charlene.Weir@utah.edu.

the architecture of design, the definitions of cause, sources of bias and confounds, and the importance of measurement as related to the various types of health IT questions. The goal is to provide practitioners with a roadmap for making decisions for their own specific study.

2. Types of Bias in Design

A well-accepted nomenclature for bias in design is the “Threats to Validity” approach described by Cook and Campbell [1] and later by Shadish, Cook and Campbell [2]. Threats to validity is a taxonomy of different sources of confounding that provide possible alternative explanations to the results.

A confound is a variable that is correlated with both the outcome and the intervention and provides an alternative explanation for the effect. For example, clinical knowledge impacts the speed at which providers can cognitively process an information display as well as the accuracy of their decisions. So, implementation of a dashboard on delirium can improve decision-making about acute mental status changes, but the effect may be confounded by the time spent educating the clinician about delirium during marketing and training and not the dashboard itself. In other words, time spent coaching the clinicians on the domain topic is an alternative explanation for improved decision making on the part of the group that received the intervention if time spent educating the physician or prior expertise on the topic is not controlled as a variable.

The “Threats to Validity” perspective includes internal, external and statistical validity, but the main focus on internal validity. Internal validity is the degree to which we could have confidence that the effect we see within a study (the difference between the control and the intervention) is due to the intervention. The most common confounds in health IT evaluation studies are computer skills, clinical expertise, variations in micro-culture within clinical settings, level of IT adoption of users, and the intensity of health IT implementation. In case-controlled studies, groups may differ on these variables if randomization is not part of the design.

The “Threats to Validity” approach is relatively comprehensive and is a very useful taxonomy to identify possible sources of bias when designing and evaluating all kinds of research. The basic list and a short description are presented in Table 1.

3. Types of Design

Designs vary naturally along a dimension of experimental control, ranging from low (e.g. a quality improvement study) to high levels of control (e.g. randomized control trial). Designs also vary in terms of the quality of evidence, and usually more control means a better design, but that is not always true. For example, there are designs where no random assignment is done, yet the quality of the evidence is very close to a study with a high degree of control, such as a randomized controlled trial.

For this contribution, designs will be presented in the order of the level of evidence for each. Relevant health IT research questions will be presented throughout the discussion. The randomized control trial or experiment will be presented first in significant detail, not only because it provides the highest level of evidence, but also because the design decisions involved are relevant in most other study situations.

Table 1. Threats to Validity [2].

Threat	Description
Ambiguous Temporal Precedence	Not clear whether the effect followed the cause in time. This problem occurs if variables are measured at the same time.
Group Differences or Selection Bias	The groups differ in a way that causes the effect; confounding exists and is unmeasured.
Regression to the Mean	Variables measured to be extreme the first time will, by chance alone, likely to be closer to the mean the second time they are measured.
History	Something happened externally at the same time as the intervention and caused the effect; secular trends exist that are unmeasured. For example, a minor flu sweeps the country at the same time as the implementation of an antibiotic prescribing decision support system.
Maturation / Learning.	Natural internal trends in the outcome variable can explain the effect, e.g. patients get better naturally, or IT users learn and become more skillful over time. Even IT systems “learn”, such as the improvement in computer support over time after implementation of a computerized physician order entry system (CPOE).
Attrition / Diffusion	The group composition changes over time as subjects quit (clinicians stop using the IT system) or move to another IT system.
Reactance or Hawthorne Effect	Subjects react to the experiment, either trying to please, are resistant or wish they were either in the control or experimental group.
Experimenter Bias or Lack of Blinding of Researchers	If the research team is not blind to the experimental condition, there is a high probability that there will be bias in data collection or analysis.
Instrumentation	The skill, reliability and content of how the outcome variables are measured differ between groups. For example, in a study examining the impact of CPOE on Adverse Drug Events (ADE), the ease by which electronic ADE notes are read versus handwritten ADE notes is a plausible reason for posited lack of a positive effect of the CPOE system.

3.1 Randomized Experiments

Random assignment to groups reduces the plausibility of an alternative explanation for research findings and more importantly it produces an unbiased average treatment effect. The treatment effect is the difference between what you found and what you *could have found* if you were able to go back in time and repeat your experiment with everything the same except for your intervention (contrapositive). It is important to note that it is a theoretical or probabilistic effect (it will be found on average to be true if exactly repeated many times). Sometimes randomization is assumed, such as when every other subject is assigned to receiving treatment. This research maneuver does not provide the best method to minimize group differences without confounding, is not truly random, and is not recommended.

Randomization ensures that group differences that may be the cause of the effect are usually rare, or, in other words, confounding variables are not generally or significantly correlated with the treatment condition. However, in any specific study, there may be group differences. Randomization simply ensures that, on average, those differences will be zero.

Experimental design also supports the logic of causal inference by controlling the sequence of events: randomization to groups occurs first, the intervention follows next and measurement of outcomes is last. It also supports causal inference by creating groups that differ only on the proposed causal variable, e.g. there is a control group that does not receive the intervention, but all other experimental procedures are the same. The difference between the two groups' results is the effect. It is important to note that in many non-randomized studies, the correlation between any unmeasured confound

and treatment assignment is highly unlikely to be zero. Statistical analyses will not be able to totally remove the effect of unmeasured, or even measured confounders.

Unit of Analysis. In many settings, including healthcare, the level of randomization is an important consideration. Randomization can occur at the level of the individual or the unit. In most clinical settings, providers are grouped into clinics or institutions, creating a natural correlation between those within a unit. Randomizing units to treatment conditions is a very appropriate strategy. For example, in a large healthcare system, 30 clinics could be organized by size into three groups of 10 each. Then, within each size group, the 10 clinics could be randomized to one of two conditions (five in each group). Size would then be a blocking variable in the analysis. Because implementation of Health IT always takes time, a very common and useful process is to group clinics into early, middle and late implementation. Randomization could then occur within those three groups. Because using aggregate units always decreases power, planning is very important. Power is defined as the statistical likelihood that you would detect a difference if one exists. Power is determined by the sample size (the larger the better), the strength of the intervention and the alpha level used to make decisions. Power is generally increased by adding units and not by increasing the number of individuals within units. Multilevel or hierarchical linear models are required for analysis. [3]

Choice of Control Groups. A key question to be considered in any design is the choice of a control group. Commonly, the control would be treatment as usual. This choice works if the usual treatment has a known effect and if the decision to be made is to use the new system or not. If health IT researchers are interested in the impact of a decision support system (DSS) on a provider's antibiotic prescribing, a possible control is those providers (or clinics) who do not receive the DSS with the outcome of interest being appropriate decision-making. However, often some form of the intervention will be implemented, regardless of the results of the study. In that case, a more refined process of choosing control groups is helpful. In the example above, a more interesting question might be a comparison between two forms of the DSS to see which would have a larger impact (especially if a DSS intervention is already going to happen). Or, perhaps, three levels of intensity of the DSS could be tested (e.g. alerts only, alerts plus education, and alerts plus education plus provider feedback). Providers (or clinics) could be randomly assigned to one of those three and the results would help IT specialists to know what components or intensity are necessary for the effect. A variation of this design is one that directly tests a "dose" effect. Perhaps the question is the degree of provider training necessary for adoption and the design could include groups that offer none versus one-on-one support.

Use of a Pre-Test. Pre-test measures are strongly encouraged by most experts [1, 2]. Pre-measures of the dependent variable(s) should always be done unless it would result in inappropriate sensitization (such as a knowledge or attitude test). Pre-tests (or baseline assessments) not only help in statistical analysis, but also are very useful in settings where attrition or diffusion (movement across the groups) might occur.

Factorial Design. A factorial design is one in which there are two or more independent variables. For example, we might be designing an alert about a patient's declining functional status and choose to represent it as narrative text versus reporting a functional status test crossed with the type of information in the alert (mental versus physical status). Figure 1 illustrates the basic design of this example which is a 2 by 2 factorial design. In factorial designs, participants are assigned to one (and only one) of the cells in the design, creating what is called a "between-subject design." They have

several advantages. First, they allow a direct test of the additive effects of different variables. For example, in the study comparing the form of an alert on functional status, a factorial design would provide a test of whether manipulating the display of information or the timing of the alerts is enough alone or whether both together create twice the effect. A crossed factorial design also supports a test for an interaction between variables, such as a question asking if an alert using narrative text is more effective when the provider logs onto the patient's chart, and whether the numerical-based presentation works better later when the provider is starting their note. So, in summary, a factorial design helps to answer many questions at the same time, making it an efficient design for studies asking a broad set of questions.

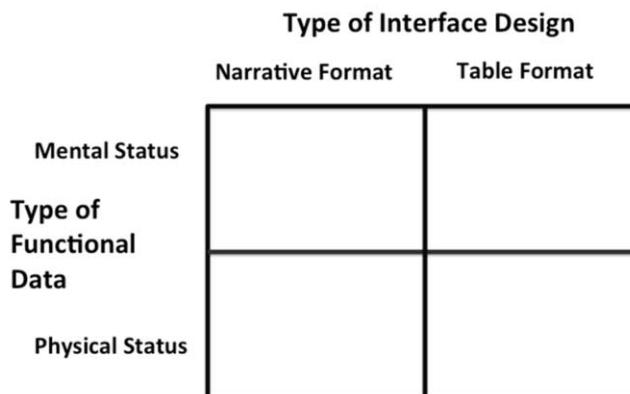


Figure 1. A pictorial representation a 2 x 2 factorial design with Factor 1 being the type of Interface with 2 conditions (narrative versus table format) and the second factor being the Type of Functional Data (mental status versus physical status).

Within-Subject and Crossover Designs. Because sample size is such an issue in experimental studies, researchers have derived methods that increase power, while preserving some of the strengths of a randomized trial. One common solution is to use the same subjects across all conditions. In other words, if you are testing the impact of a specific display, it is reasonable to have each subject respond to both types of displays if the order by which they are presented is randomized. When data is collected from the same subject in both experimental conditions, the design is referred to as a “within-subject” design. The main limitation of this approach is that participants are likely to “guess” your purpose and will change their responses on the second exposure. So, when using a within-subject design, careful attention must be given to minimizing contamination across the cells. Extending the time between measurements or changing to context can obscure the obvious links between conditions.

Randomization in the Field. Arguing for randomization may not be possible when the study is in a real clinical institution that is not your own or in situations where you are not in control; however, sometimes a researcher can take advantage of a naturally occurring situation, such as time-based interventions or wait lists. For example, in one study, the pharmacy department was trying to decrease the use of ineffective anti-dementia drugs and was systematically calling all patients and their families to discuss barriers to doing a discontinuance trial and to offer all of them a personalized patient portal first. The portal provided greater access to their clinical team. They were going

down the list alphabetically, but were easily convinced to select the patients in random order. Since the project was expected to take 6 months, it was easy to naturally compare patients who had access to the portal with those who didn't in terms of agreeing to doing a trial discontinuance.

In summary, experimental designs are considered the gold standard for providing evidence in research. Even when very strong observational designs are used, a randomized trial provides better evidence. However, there are significant limitations, including feasibility and ethical concerns. Often, it is simply not possible to randomize, such as when implementing provider order entry and the whole organization has to be involved. Or, it is not ethical to prevent part of the system from having access to an important intervention. And, just because randomization has occurred doesn't mean that threats to validity that occur *during* the course of the experiment won't bias the results, such as *attrition* (providers may quit using the program or decide to use the alternative) making the groups different. Or, providers in the intervention group may receive significantly more attention from the IT support staff than the control group resulting indirectly in increased skills. With some attention, health IT researchers can maximize the potential in the situation to conduct a randomized control trial, while still meeting the other operational goals of implementation.

3.2 Regression Discontinuity Design

Regression Discontinuity Design (RDD) is a research design that produces a quality of evidence nearly as high as a randomized control trial and is useful when randomization is not feasible [4]. It is particularly useful in health IT studies where there is a lot of data across a system that can be measured and specified before the intervention. It is used for those programs where only a subset of patients or providers is targeted for intervention (such as low performance).

In RDD design, groups are defined using a cutoff on a covariate (the assignment score) that is correlated with the outcome or is the outcome itself. The assumption is that the outcome variable is a continuous function of the assignment score. The variable that creates the best assignment value for a clear model is a pre-test of the outcome variable itself. The process is often used in education where the targets of a program are students who are performing at lower levels. The variable used for the cutoff is the pre-test on the performance variable. A clear cutoff of this variable is used to select a group (e.g. the bottom 50%). A local average treatment effect can then be estimated from observations on either side of the cutoff point and is a necessary condition for minimizing bias [5]. The treatment effects identified can be very similar to randomized studies [4]. The design is used in situations where it is considered ethical and/or desirable to offer treatment to those in need or who have low performance [6].

In healthcare, this design might translate to offering a decision support intervention to physicians who are at the lower end of a "compliance distribution" regarding a specific guideline, e.g. hypertension guidelines. Or, patients who are scored lower than the median cutoff score for understanding their hypertension could be provided access to an intensive online training. Then, a cutoff is chosen that determines a sharp dividing line between an intervention and control group. In the first example, that cutoff would be the providers who are 50% non-compliant with guidelines (across all of their patients). In the second case, it would those patients whose pre-test score of knowledge of hypertension were below the median score. When a regression line is computed between the assignment score (S) and the treatment outcome (average blood pressure of

a provider's panel of patients), the line will have a clear disjunction at the cutoff level if the program had the effect. Figure 2 illustrates the two stages of an RDD design. A provider's average across all of his patients is regressed on the pre blood pressures with group assignment as a variable. We can see a clear difference in the regression line for those in the treatment group as compared to those in the control group (dotted line) that disappears at the median dividing line between the treatment and control groups.

The assumptions for internal validity for RDD are: 1) The value of the pre-score is easily measurable and there is a clear cutoff that discriminates those in and those out of the different groups; 2) the relationship between the assignment variable and the outcome variable is a consistent regression line across all values of the identifying variable; and 3) the relationship between the two is correctly modeled. These assumptions are quite strong, however, and statisticians have developed some alternative analytic schemes [6].

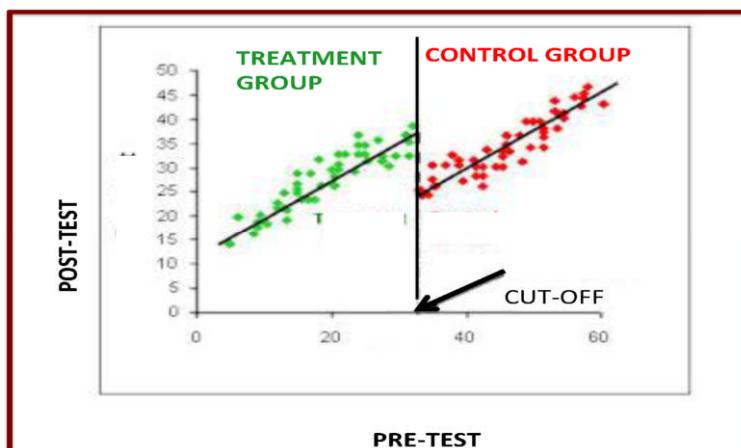


Figure 2. An example of a Regression Discontinuity Design (RDD) using a pre-test as the variable used for the cutoff at the 35% level. The line in the middle is the cutoff line. If an intervention had not occurred for those whose pre-test was below the 35% line, they would have been regressed along the dotted line.

Overall, power in a RDD design is lower and is estimated to be between 3 to 4 times smaller than a RCT. So, the sample size should be about 3 times larger than that estimated for a RCT.[7] Studies using electronic data will be at an advantage because there is often enough data to test for the clear time 1 to time 2 relationship underlying the basic model as well as to identify the group of patients or providers who meet the cutoff criteria [8]. The overall statistical analyses are beyond the scope of this contribution, but some useful references are provided at the end. An RDD design may be a very viable alternative in some health IT settings as identification of some covariates, the outcomes and the actual intervention may possibly all be automated. Dillender and Geneletti et al. are examples in the health and data field [9, 10].

3.3 Quasi-Experimental Designs

Quasi-experimental designs are those studies where subjects or units are not randomized. Table 2 describes the different types of quasi-experimental designs. Each design brings with it different types of threats to validity, some of which can be

mitigated by extending the design in targeted ways. Each of the descriptions in Table 2 comes with an image of the design with “O” referring to observation and “X” referring to intervention.

Table 2. Quasi-experimental designs.

Design	Description	Threats to Validity
Pre/Post	O X O	History, Maturation, Instrumentation, Regression, Attrition
Interrupted Time Series	O ₁ O ₂ O ₃ ... O _n X O ₁ O ₂ O ₃ ... O _n	Attrition, Instrumentation
Multiple Baseline	O _{a1} O _{a2} O _{a3} ... O _{an} X O _{b1} O _{b2} O _{b3} ... O _{bn}	Instrumentation, Attrition,
Control Groups: non-randomized	O X O O X O	Selection bias, interaction effects with Instrumentation, Reactance, Attrition, Diffusion

Pre/Post Design. The O X O design would be a traditional pre/post design where the outcome is measured prior to the intervention then measured post implementation. In this contribution, we will review some of the basic forms. Harris et al. provide a thorough description of quasi-experimental designs in informatics [11-13]. The pre-post designs are the most common in informatics and the most fraught with threats to validity. The basic framework for a pre/post design is that the researcher measures the outcome at least once before the intervention and then measures it again at least once post intervention. These designs are very common in health IT evaluations, but have some significant limitations. One of the major issues is that informatics interventions do not occur in a single discrete unit of time. They are often continuously updated as part of the implementation process and as part of user feedback and limitation of resources. Deciding what time period would constitute the “pre” measurement might not be too difficult, but determining the “post” measurement could introduce bias. In addition, the exact nature of the “independent variable” may not be well defined as it evolves over time.

There are some simple ways to improve the pre-post design, but the most basic and effective is to do multiple pre-measures of the outcome variables. This process will allow for understanding of secular trends, thereby revealing regression to the mean tendencies, maturation (getting sicker or more well anyway) and/or learning trends. This design will also provide a strong argument against history (something happened at the same time as the intervention that could have caused the effect). In addition, two to three pre measures will allow a researcher to see a more finely nuanced picture of the effects of implementation strategies and user feedback in some cases. However, it is vulnerable to instrumentation effects as research processes change over time and measurement reliability increases (thereby decreasing the variation around each measurement).

Interrupted Time Series. The best and most valid design in this group is the Interrupted Time Series. In this design, many measures (as many as 30 or more pre/post) can be done to provide strong baseline trend data for valid statistical analysis. The tendency is to aggregate a variable of a year’s timeline to provide a “pre” score. However, even taking 12 monthly scores would improve the reliability of the measures and interpretability of the results. These results look very similar to what quality improvement specialists call “control charts.”

Multiple Baseline. Most researchers in applied fields have the tendency to measure only those variables that they expect to change and no others. Many evaluation experts advise to measure many dependent variables in order to detect both positive and

negative effects. In addition to adding those variables to a design, it is also very useful to add outcome measures that are not expected to be influenced by the intervention. For example, if we were implementing a patient portal, we might measure patient satisfaction, patient knowledge and patient compliance with medications – all of which the intervention would be expected to impact. In contrast, it would also be useful to measure patient's knowledge about a disease that they don't have, or satisfaction with their insurance company's coverage in order to show the effect has more precision. Similarly, if providers are targeted in a decision support system (DSS) intervention, we would measure provider's knowledge and compliance with DSS recommendations, but also their compliance with guideline recommendations not targeted in the specific DSS. In all cases, the best approach is to take multiple measures before and after the intervention. Figure 3 illustrates the results from a sample multiple baseline study.

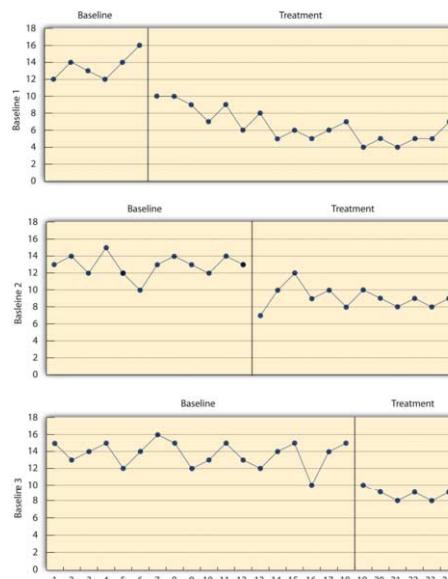


Figure 3. Example of a multiple baseline with 3 in-patient units during a CPOE intervention. The x-axis is in months and the y-axis reflects days of stay. The vertical line is the month in which the intervention began.

Control Groups, Not Randomized. The addition of a comparison group helps a great deal in increasing the validity of the study design. The more similar the groups are, the stronger an inference can be made that group differences do not provide an alternative explanation to the basic findings of the study. Similarly, assessing the groups for possible confounding characteristics pre/post also helps improve the logic of the design. Important group differences include age, gender, skills, experience, and attitudes as well as clinic-level characteristics such as staffing, patient types, location and size. These variables can be used as covariates in the analysis of outcomes, thereby providing some statistical control. However, it is important to remember that the lack of randomization leaves the study vulnerable to unmeasured confounders.

There are several tactics that help mitigate the threat of unmeasured confounding across multiple groups. First, the inclusion of multiple control or comparison groups helps inform the interpretation of the effect. Even better is the inclusion of 4 - 5 units where implementation is staggered, but the effect is measured from the beginning of

the first unit implementation. This design manipulation makes a strong statement regarding confounders. For example, if a large healthcare institution plans to implement Computerized Provider Order Entry (CPOE) across 4 hospitals and 8 clinics, the researcher could take advantage of the staggered times for implementation which naturally occur by simply measuring the outcome variables (e.g. adverse events, length of stay, or costs) at all settings from the beginning and at the same time. If the measured outcomes change in the same staggered pattern as the implementation (regardless of the baseline) then the inference that CPOE caused the change is supported.

Another common variation to a non-random comparison group design that is very useful in applied settings is the wait list. Often, clinical interventions are staggered across time in order to match the resources available, thereby creating an implicit wait list. Since the providers or patients on the wait list are assumed to be similar to the patients who first get the treatment or intervention, then outcomes can be compared between the two groups after the first group receives the intervention and before the second one begins.

4. General Recommendations for Conducting Experimental Research

In general, three factors are essential to effective research, including the choice of dependent variables, piloting and the intervention itself. The first is a careful selection of outcome variables, using theoretically-based measures with measured reliability and validity. As noted above, evaluation studies need to include a wide consideration of measured outcomes, stakeholders and contexts.

Second, taking the time to do extensive piloting is essential. Maintaining the integrity and fidelity of research procedures is as important to the validity of the results as is the design itself. Fine-tuning the intervention is important, including conducting usability testing across different roles and settings, ensuring that instructions are understandable, that the research staff is trained and reliable and that the measurement forms have been tested for ease of use are all components of good piloting. Changing the way that data is collected during a research project is a major threat to validity and very common in informatics research. Data collected electronically is not really the same as data collected from handwritten logs. Research assistants learn and become more accurate over time and machines wear out over time. Observers and coders become fatigued and their data can suffer in a cyclic pattern. All of these factors contribute to time-varying confounds and require careful scrutiny and frequent inter-rater reliability checks during the course of the study pilot work can take as long as the study itself, but will greatly improve the quality of the findings.

Third, the intervention itself needs be analyzed in terms of intended and unintended mechanisms of action. In the field of informatics, researchers often fail to specify the components of an intervention, the importance of the implementation strategies as part of the intervention and the continuous feedback cycle of change.

5. Conclusions

Research design is often viewed as confusing, mysterious and inaccessible. Hopefully, the threats to validity approach described here will provide informatics researchers with

tools to analyze the quality of any proposed design and to guide their work. The design itself cannot take the place of a sound theoretically approach, of understanding prior literature, of sufficient piloting, and, of course, developing a powerful and effective intervention. In the end, it is the creativity of the research team that makes good science.

Recommended further readings

1. E. Layman, V. Watzlaf, *Health Informatics Research Methods: Principles and Practice*, AHIMA, Chicago, 2009.
2. C. Friedman, J.C. Wyatt, *Evaluation Methods in Medical Informatics* (2nd Edition), Springer Science Business Media, New York, 2010.
3. J. Pearl, *An Introduction to Causal Inference*, CreateSpace Independent Publishing Platform, 2015.
4. W. Shadish, T. Cook, D. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Houghton Mifflin Company, Boston, MA, 2002.

Food for thought

1. Why is it important to publish even non-significant results?
2. Why is it not appropriate to dismiss a study with a small sample size even though the results are significant?
3. Think of some historical events that could confuse our interpretation of a pre-post intervention assessment of a decision-support intervention!
4. Discuss how the constantly changing aspects of a health IT intervention may impact of conclusions of a study examining the results of such an intervention!

References

- [1] T. Cook, D. Campbell, *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Rand McNally, Chicago, 1979.
- [2] W. Shadish, T. Cook, D. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Houghton Mifflin Company, Boston, MA, 2002.
- [3] S. Raudenbush, A. Bryk, *Hierarchical Linear Models (Second Edition)*, Sage Publications, Thousand Oaks, 2002.
- [4] W. Shadish, M. Clark, P. Steinera, Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random and Nonrandom Assignments, *J of the Am Stat Assoc* **103**(484) (2008), 1334-44.
- [5] J. Rubin, Assignment to Treatment on the Basis of a Covariate, *Journal of Educational and Behavioural Statistics* **2**(1) (1977), 1-26.
- [6] G. Imbens, T. Lemieux, Regression Discontinuity Designs: A Guide to Practice, *Journal of Econometrics* **142**(2) (2008), 615-35.
- [7] H. Lee, T. Munk, *Using regression discontinuity design for program evaluation*, Section on Survey Research Methods, JSM, 2008, pp. 1678-82. <https://www.amstat.org/sections/srms/Proceedings/y2008/Files/301149.pdf>, last access 11 February 2016.
- [8] P. Schochet, *Technical Methods Report: Statistical Power for Regression Discontinuity Designs in Education Evaluations in (NCEE 2008-4026)*, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education Washington DC, 2008.
- [9] M. Dillender, The effect of health insurance on workers' compensation filing: Evidence from the affordable care act's age-based threshold for dependent coverage, *J Health Econ* **43** (2015), 204-28.

- [10] S. Geneletti, A.G. O'Keeffe, L.D. Sharples, Bayesian regression discontinuity designs: incorporating clinical knowledge in the causal analysis of primary care data, *Stat Med* **34**(15) (2015), 2334-52.
- [11] A.D. Harris, J.C. McGregor, E.N. Perencevich, J.P. Furuno, J. Zhu, D.E. Peterson, J. Finkelstein. The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics, *J Am Med Inform Assoc* **13** (2006), 16-23.
- [12] E. Layman, V. Watzlaf, *Health Informatics Research Methods: Principles and Practice*, AHIMA Press, Chicago, 2009.
- [13] C. Friedman, J.C. Wyatt, *Evaluation Methods in Medical Informatics* (2nd Edition), Springer Science Business Media, New York, 2010.

Mixed Methods: A Paradigm for Holistic Evaluation of Health IT

Philip J. SCOTT^{a,1}

^aCentre for Healthcare Modelling and Informatics, University of Portsmouth, United Kingdom

Abstract. This contribution offers an overview of the ‘third research paradigm’, its historical roots and its relevance for health informatics. Using illustrative studies, we explore the concepts of triangulation and integration of quantitative and qualitative data and refute common philosophical objections to mixing different types of knowledge. We consider how the mixed method paradigm relates to two programme design and evaluation frameworks that are important for health informatics: realist evaluation and Theory of Change. We discuss how to manage practical challenges to this approach and explain how mixed method studies support an evidence-based approach to real world policy, planning and investment decisions.

Keywords. Evidence-based practice, epistemology, informatics, outcome and process assessment, program evaluation, qualitative evaluation, quantitative evaluation, randomized controlled clinical trials, research design, theoretical effectiveness.

1. Introduction

Given the socio-technical nature of information systems, mixed method designs are widely regarded as essential for their holistic investigation [1]. The research locus for health information systems is a field of complex interaction, opaque mechanisms of action, contested definitions of success, context sensitivity and unpredictable unintended consequences. Hence, evaluating the ‘whole picture’ in health informatics surely needs a multi-dimensional synthesis [2].

What is the mixed methods paradigm? The *Journal of Mixed Method Research* defines its topic as “research in which the investigator collects and analyzes data, integrates the findings, and draws inferences using both qualitative and quantitative approaches or methods in a single study or program of inquiry” [3]. It also has been characterized as “the use of whatever methodological tools are required to answer the research questions under study” [4], thus differentiating it from solely qualitative or quantitative approaches.

¹ Corresponding author: Dr. Philip Scott, Centre for Healthcare Modelling and Informatics, University of Portsmouth, Buckingham Building, Lion Terrace, Portsmouth, PO1 3HE, United Kingdom, Philip.scott@port.ac.uk.

The objectives of this contribution are to:

- Give a brief overview of the mixed method paradigm, signposting the seminal literature.
- Analyze two example health informatics studies that employed mixed methods.
- Refute three philosophical objections to mixed methods: the ‘incompatibility thesis’, the ‘disguised positivism critique’ and the ‘holism misapprehension’.
- Explore how the mixed method paradigm relates to two theory-driven programme design and evaluation frameworks: realist evaluation and Theory of Change.
- Consider some practical difficulties in resourcing and executing a mixed method project and reflect upon the impact of mixed method studies on real world policy and planning.

2. Mixed methods: the ‘third paradigm’?

2.1 History

Although social researchers have been combining quantitative and qualitative methods in various ways since at least the 19th century [5], the formal concept of a distinct mixed methods research paradigm is contemporary. Mixed methods historians often cite the 1959 proposal for “multiple operationalism” (a validation process correlating independent quantitative variables and methods) [6] as a prototypical development.

The geodetic term ‘triangulation’ was reportedly first adopted in this sense in 1966, to mean the use of multiple quantitative measurement processes as a means to provide “the most persuasive evidence” about a theoretical proposition [7]. The idea of research triangulation was progressively extended, and by the end of the 1970s had gained a broader usage that included combining *qualitative and quantitative* methods and data [8]. Despite the historical record of methodological eclecticism in actual research practice and the emergence of concepts like triangulation to justify their co-existence, academic debates in the 1970s-80s were characterized as ‘paradigm wars’ between supposedly irreconcilable philosophies and practices of qualitative and quantitative research [9]. One response to these disputes was the explicit formulation of mixed methods as a ‘third paradigm’, which emerged from the 1990s onwards as a new methodological movement [4].

The ‘third paradigm’ quickly bloomed, gaining its own dedicated journal in 2007 [3, 10] and forming an international research association [11]. It has generated a substantial body of literature, with several handbooks and digests of key sources [4, 12, 13]. The “current orthodoxy” now recommends a multi-method approach [14]. However, as we discuss further below, recurrent objections persist to fundamental aspects of this ‘movement’.

2.2 Defining characteristics

What differentiates the mixed methods research paradigm? As noted above, an essential mixed methods argument is that the research questions dictate the appropriate techniques: neither interpretivist/constructivist/qualitative nor post-positivist/quantitative methods are privileged. Mixed methods studies are “interested in both

narrative and numeric data” [4]. The basic rationale for the mixed methods approach is holism [5, 15].

By definition, the aim of holistic research is to get ‘the whole picture’. Neither an insightful thick description nor a reductive controlled trial with bounded statistical inferences is, on its own, a complete account. Countable things should be counted with reliability and validity. Unquantifiable things should be studied with authenticity and fidelity. Neither component is inherently superior to the other, though the objectives of a particular study will determine the relative contribution of each. The greater value offered by holism is that ‘the whole is greater than the sum of its parts’: a synergy between interpretivist/constructivist and post-positivist worldviews and between qualitative and quantitative methods and data.

Despite their association with “vile positivism” [14], randomized controlled trials (RCTs) remain necessary to answer many research questions [16] (particularly about *efficacy* or *effectiveness* – does intervention/system X actually result in outcome Y changing significantly)² and to provide data that can be subject to quantitative meta-analysis. Cluster RCTs can be used to evaluate complex interventions [17], where the unit of intervention is ward, practice or some other healthcare organizational entity – typically the unit of deployment for an information system.

Of course, RCTs do not guarantee ‘purity of knowledge’. Industry funding and publication bias are known to contribute to conclusions that significantly favour experimental interventions both in primary research and meta-analysis [18]³ and apparently influence published opinion about whether to use systematic reviews in policy-making [19]. These significant criticisms do not invalidate the RCT as a study design, but they do demonstrate the need to incorporate qualitative findings to inform design and interpretation. In practice, many studies aim for this in the ‘discussion’ section of their published findings. What the mixed methods paradigm proposes is to make this explicit from the conception of the study and through each stage of its execution.

On the other hand, some research questions simply cannot be answered by a solely RCT design – when the objective is *explanation* of real-world phenomena. For example, if the purpose is to determine how aspects of organizational context affect adoption of a system and realization of planned benefits, or why error rates change seasonally in certain care services, it is clear that qualitative methods must be the principal techniques. Nonetheless, the full picture may be enhanced with quantitative data: What are the adoption rates in comparative services? What are the seasonal trends for medication errors? A mixed methods approach defends qualitative evaluation against the criticism of being mere reportage.

How can qualitative and quantitative methods and data be meaningfully combined? Common types of combination are triangulation, integration and dialectic. Mixed methods textbooks illustrate in detail how various kinds of method and data combination have been designed and executed (for instance, chapters 7 and 11 in [4]). These configurations typically use Morse’s notation to indicate methodological dominance (capital letters, “QUAN” or “QUAL”) and sequence (“→”) or concurrence (“+”). For example, “QUAN → Qual” would denote a quantitatively driven project

² See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

³ See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: *ibid.*

followed by a qualitative study, whereas “QUAL + Quan” would mean a qualitatively driven study that has concurrent qualitative and quantitative components.

Triangulation can be as simple as comparing and contrasting qualitative and quantitative responses in a questionnaire, or between a quantitative RCT and in-depth qualitative interviews or focus groups [15, 20]. *Integration* aims to go beyond ‘compare and contrast’ to achieve a richer synthesis where qualitative and quantitative methods and data are interdependent and “mutually illuminating” [21].

The *dialectic* stance explicitly recognizes the different ‘voices’ and worldviews present in the situation under study and welcomes ‘divergence and dissonance’ so that analytical dialogue can generate new perspectives and insights [22]. Whichever approach is used, the combined analysis can be presented in a narrative, tabular or graphical format [20].

3. Example mixed methods studies in health informatics.

Mixed methods have been used in various health informatics evaluations, addressing topics such as the unintended consequences of computerized physician order-entry (CPOE) [23, 24], use of smartphones for clinical communication [25] and clinician adoption of summary care records [26] and point-of-care systems [27]. In this section we discuss two examples, to illustrate how mixed methods can work in practice.

Wu and colleagues [25] report a mixed method evaluation of using smartphones for team communication in hospital general medicine units. They gathered quantitative data about frequency of smartphone usage (calls and emails) and qualitative data from semi-structured interviews and ethnographic observations. Their analysis presents thematic analysis of the qualitative data and descriptive statistics about the volume and frequency of smartphone usage. The study noted improvements in team efficiency but negative experience in the volume of interruptions and some aspects of professional behaviour (such as inappropriate smartphone usage whilst dealing with a patient). Divergent views were noted about which communicated incidents were genuinely urgent and about inter-professional relationships (with nurses having negative perceptions and doctors reporting positive views). The study integrates some of the statistics with the qualitative analysis; for example, connecting the observations and interview comments about the perceived level of interruptions with the actual usage data. The study draws primarily qualitative conclusions, so can be characterized as a “QUAL + Quan” design.

In [23], Ash and colleagues (“POET” – the Physician Order Entry Team) summarize their four year programme of investigation into the unintended consequences of CPOE. This is a strong example of a flexible, iterative study design that used mixed methods to seek a holistic understanding. Following their earlier work on CPOE success factors, the group developed and piloted qualitative semi-structured interview and observational techniques to explore CPOE unintended consequences at one institution. Then they organized a conference of invited experts to gather additional data and elicit stakeholder advice on the next stage of field work. After the subsequent main phase of qualitative observations and interviews at five sites, POET created a framework of nine types of CPOE unintended consequences. This framework was then used in comparison with results from a short survey instrument to gauge clinician expectations about forthcoming CPOE implementations at three further sites, showing that neither ‘average’ clinicians nor leaders had a true picture of the likely impact.

POET next developed a telephone survey instrument with both qualitative and quantitative (nominal and Likert scale) content, aimed at all US hospitals using CPOE, followed by a second conference of invited experts to validate the results and plan dissemination of the learning.

The POET research programme was predominantly qualitative, with quantitative data limited to descriptive statistics and correlation analysis in the telephone survey study. The various reports offer a narrative integration of the qualitative and quantitative findings. The overall programme could be described as “QUAL → Quan → Qual”. POET explicitly drew upon a mixed methods rapid assessment approach used for public health interventions, and have proposed how this can be used in other clinical information system evaluations [28].

4. Mixed-up thinking? The ‘incompatibility thesis’ and other stories.

4.1 *Epistemological incoherence?*

A common argument used against the mixed method paradigm, still raised [5, 29] (and implied in [30, 31]), is that qualitative and quantitative methods and data derive from contrasting philosophical worldviews (a set of beliefs and values giving rise to a particular outlook upon life and reality, typically either interpretivist/constructivist or post-positivist in this context) which are incompatible (or ‘incommensurable’, as Kuhn characterized successive scientific paradigms [32]). This ‘incompatibility thesis’ therefore asserts that qualitative and quantitative methods and data cannot meaningfully be combined and to do so would be ‘epistemologically incoherent’ [33]. This argument is sometimes limited to specific quantitative methods such as RCTs or experimental designs in general. Various descriptors of the allegedly incommensurable ontologies and epistemologies are used. Additional dimensions of incompatibility are sometimes added, such as ‘axiology’ – beliefs or theories about values.

Numerous authors have discredited this argument [2, 4, 33-35]. The three main grounds of refutation we address are: logical fallacy, pragmatist philosophy and research praxis.

The logical fallacy is exposed by decomposing the incompatibility argument into four steps:

1. Realist and relativist ontologies (and objectivist and constructivist epistemologies, respectively) are fundamentally irreconcilable belief systems.
2. Therefore, knowledge from one belief system cannot meaningfully be integrated with knowledge from the other belief system.
3. Quantitative knowledge is intrinsically bound to a realist ontology and objectivist epistemology. Qualitative knowledge is intrinsically bound to a relativist ontology and constructivist epistemology.
4. Therefore quantitative and qualitative knowledge cannot meaningfully be integrated.

The first two points of the argument are addressed below. The simplest refutation of the fallacy is that the third step of the argument is palpably false. Worldviews are held by people, not by methods and data [36]. The fact that a quantitative (or experimental) method produces a set of numbers does not prescribe how ‘real’ the referents of those data points are, or what kind of knowledge is claimed by reporting

those numbers. That is down to the interpretation of the results. Equally, the narrative themes of a qualitative study may or may not be ‘real’, and qualitative knowledge may or may not be reliable and transferable. The debate over ‘realist RCTs’ (see section 5.1, below) illustrates this point.

Furthermore, pragmatist philosophy suggests that the first two steps of the incompatibility thesis are also false [2]. Pragmatism does not make absolute knowledge claims, but presents only ‘warranted assertions’; the values and confidence levels of the ‘warrant’ are contingent. Pragmatists contend that human knowledge does not have *a priori* foundations: ontology and epistemology are developed empirically and have no privilege over any other data or argument.

Finally, actual research practice demonstrates that qualitative and quantitative methods and data have in fact been successfully integrated in many ways for many years. Clinical practice is a prime example of everyday combination of multiple worldviews (for example, patient, carer, doctor, nurse, medical device or algorithm) and qualitative and quantitative methods and data. To borrow a ‘vile positivist’ concept, the incompatibility thesis has been empirically falsified.

4.2 Disguised positivism?

It has been charged that the mixed method paradigm is “positivism in drag” [9]. An early form of this criticism was that “Mixed-method designs are direct descendants of classical experimentalism. They presume a methodological hierarchy in which quantitative methods are at the top and qualitative methods are relegated to ‘a largely auxiliary role...’ ...it excludes stakeholders from dialogue and active participation in the research process” [37].

In context, such critique seems not so much about the general mixed method paradigm but more about specific incarnations of mixed method study as defined by certain US research funding agencies. The idea that using mixed methods means excluding stakeholders from dialogue and participation is simply wrong. Nor does mixed methods thinking put experimentalism on a pedestal. However, this observation raises a legitimate question. Is the orthodox ‘evidence hierarchy’ universally applicable outside the field of medicine? We raise this question again in section 5.

4.3 Spurious holism?

A final philosophical objection is the ‘holism misapprehension’. This term is introduced to counterpoint the proposition that “Holism, in any form of inquiry, is neither obtainable nor desirable” [34]. The authors of this quotation were arguing against what we might call ‘absolute holism’: the unachievable ambition to assess every single aspect of a phenomenon.

Their argument is clearly right in the sense they intended: a map or model is by design only a partial representation and can only be such. However, it is worth reflecting on the more limited (and original) sense of holism that we adduce to justify the mixed methods paradigm. According to the Oxford English Dictionary, the word ‘holism’ was coined in 1926 as a biological term for ‘the tendency in nature to produce wholes from the ordered grouping of unit structures’ [38]. This ‘wholeness’ is what mixed methods evaluation is seeking: a rounded, multi-perspective understanding, not some unattainable definitive completeness. Holism, then, is inherently an idea about organic systems – a direct conceptual resonance with information systems evaluation.

5. Relation of the mixed methods paradigm to theory-driven programme evaluation

The term “programme” is used here to describe a portfolio of interventions that seeks to change individual or group behaviour, or organizational structure and performance, to meet some political, social, commercial or philanthropic aim. Various frameworks have been used to inform the design, development and evaluation of such change programmes.

What programme frameworks have in common is the recognition that changing human and institutional behaviour is highly complex, with multiple interacting contextual factors, so pathways and mechanisms of change need to be unpacked in some detail to understand what is really happening or may happen. Although different terminology is used in each framework, each is trying in some way to produce a ‘programme theory’: with *these* assumptions, how is it that the desired goals can be achieved in *this* context for *this* programme. This is clearly a relevant approach for health informatics, given that its fundamental *raison d'être* is to change healthcare for the better through improvements in the use of information. In this section, we consider two widely used frameworks: realist evaluation and Theory of Change.

5.1 Realist evaluation

Realist evaluation aims to answer the question ‘what works for whom, in what circumstances and in what respects, and how?’ [14, 39, 40].⁴ Its name derives from its original philosophical roots in Bhaskar’s critical realism, a central concept of which is that the objects of investigation are *mechanisms* that can be activated to produce particular *outcome patterns* in certain *contexts*. A mechanism may exist but not be activated, be activated but not observed, or be activated but affected unpredictably by other mechanisms or by the context. Realist evaluation has three principal characteristics: an emphasis on theory and explanation, a multi-method approach and a focus on the context-mechanism-outcome (C-M-O) configuration [14]. The realist hypothesis (of what is happening in the C-M-O pattern) is also called the ‘programme theory’ (an expression which has other meanings in other frameworks).

The mixed method paradigm and realist evaluation are, broadly speaking, quite consistent. However, there remain echoes of the incompatibility thesis that currently inhibit realist evaluation. Although realist evaluation claims to be ‘method neutral’, there is disagreement among its advocates whether ‘realist RCT’ is a meaningful concept [31, 36, 41]. The originators of realist evaluation were highly skeptical of experimental designs for programme evaluation, and contemporary followers remain only ‘cautiously supportive of quasi-experimental designs’ [31]. They reject the ‘realist RCT’ and seek to limit application of the term ‘realist’ to what they endorse as such. The mixed method paradigm suggests a more open choice of methods and does not exclude study designs on philosophical grounds. It is irrelevant that RCTs historically derive from a positivist background with a successionist model of causality. What the RCT results mean is down to interpretation. RCTs control for context, but how they do so is a matter of detailed design; factorial studies or mediation analysis can be used to

⁴ See also: T. Otte Trojel et al., Going beyond systematic reviews: Realist and meta-narratives reviews, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

explore contextual factors rather than ‘design them out’. RCT methods may legitimately be used by researchers with positivist, pragmatist or realist worldviews. This more open minded approach has been adopted in the UK Medical Research Council guidance on process evaluation of complex interventions [42].

5.2 Theory of Change

Theory of Change has its roots in international development projects [43]. It is not in fact “a” theory, but an approach or a way of thinking to unpack and articulate “the” Theory of Change for a particular programme. It includes analysis of context (political, organizational, social and environmental), actors (implementers and ‘subjects’ of the change), assumptions (about the participants or the mechanism or effectiveness of the proposed interventions and indicators) and rationale (evidence or hypothesis that the interventions will work as anticipated). It provides a way of prospectively articulating programme theory in a graphical model with explanatory narrative, based on extensive discussion with stakeholders.

Theory of Change starts from the long-term goals and maps backwards to necessary pre-conditions, causal outcome pathways, interventions, assumptions, rationale and measurable indicators. It is both a process and a product; the product is “a working model against which to test hypotheses and assumptions about what actions will best bring about the intended outcomes” [44]. As such, it offers a framework both for programme design and for evaluating complex interventions [45]. Theory of Change is methodologically neutral and has accommodated both qualitative techniques and studies using randomized controls and other experimental designs [46].

In its present form, Theory of Change has a more prescriptive process than realist evaluation, (though in practice taking various forms and being flexibly applied [43]), but less epistemological baggage. Reportedly, Theory of Change studies in practice often address the ‘implementation theory’ (how the intervention should be operationalized to meet its objectives) rather than the ‘programme theory’ (the theoretical causal relationship between the mechanisms of an intervention and the desired behavioral outcomes) [46]. The mixed method paradigm would seem to work quite naturally with the Theory of Change.

6. Practical challenges for mixed method studies

There are several factors that pose practical difficulties in resourcing and executing a mixed method project and may limit the impact of mixed method studies on real world policy and planning. We comment briefly on issues related to funding and resources, the research team, and the relative position of mixed methods studies in systematic reviews.

Perhaps the most obvious constraints are the usual ones – money, people and time. By definition, mixed method studies are trying to do more than mono-method studies and are therefore likely to need more people (with a mix of skills), be more complicated and take longer than a solely qualitative or quantitative evaluation. This raises the issues of affordability, project management, timescale and flexibility of study design. Funders may take some convincing that the benefits of a longer, more expensive and more convoluted project (with some parts likely to be iterative and not fully definable at the proposal stage) are truly justified. This problem might be

addressed by breaking the evaluation into stages and seeking funding for each phase in turn, or finding a scheme to finance preliminary work that may lead to a larger programme (e.g. [47]).

Once a mixed methods project is operational, there may be team working issues [48]. Even if there is no formal incoherence at the methodological level, the ghost of rival epistemologies may haunt interpersonal relationships between specialists from qualitative or quantitative traditions. The tacit knowledge, assumptions and discourse that each individual brings to the team may trigger a culture clash. There are likely to be disagreements not just about the practical logistics of the study but its framing, purpose and priorities. There may, for example, be tension between the 'hard' thinking style of health economists and the 'soft' thinking style of sociologists – or vice versa, or either group may deliberately resist categorization and 'act up' by adopting what they see as the worldview of their 'opponents'. The only solution to this is careful team recruitment that is sensitive to personality types and a management style that cultivates an open minded team dynamic that embraces diversity and consciously tolerates ambiguity.

The consideration of the mixed methods approach as a means to seek holistic knowledge naturally leads to reflection upon the translation of such multi-faceted knowledge into policy formation and programme planning. That in turn suggests reflection on systematic review methods and the hierarchy of evidence. 'Standard' approaches (such as the Cochrane and Campbell Collaborations⁵) have tended to be associated primarily with quantitative methods, with systematic reviews of RCTs at the top of the pyramid (e.g. [49]). Alternative attempts to blend different kinds of evidence have included realist review [39] and meta-narrative synthesis [50]⁶, with the latter having some echoes of the dialectic approach mentioned in section 2.2 (see also the suggested further reading). Where should mixed method studies fit in the evidence hierarchy? Is not an evaluation that integrates an RCT *and* qualitative data stronger evidence than an RCT alone? Does the evidence pyramid need new layers to grade systematic reviews of mixed method studies? Perhaps health informatics needs its own evidence hierarchy, learning from other disciplines? These are questions to ponder.

7. Conclusions

While the 'paradigm wars' are notionally in the past, there remain some tensions between the instincts and preferences of qualitative and quantitative researchers and the institutional contexts within which they operate. Although the barriers are perhaps less insurmountable than a few decades ago, traditional academic job titles and career paths tend to sustain this contrived methodological and philosophical divide. For example, the labels 'statistician' and 'qualitative researcher' can tend to imply particular backgrounds and contrasting cultural orientations.

The increasing awareness of 'programme theory' in various frameworks is a positive development that supports adoption of the mixed methods paradigm: the

⁵ See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁶ See also: T. Otte Trojel et al., Going beyond systematic reviews: Realist and meta-narratives reviews, in: *ibid.*

mechanisms of change need both to be explained and measured in various ways. Mono-method approaches do not offer the depth and richness that a holistic mixed methods design and evaluation can bring. Government funding agencies, such as the UK Department for International Development, set a powerful example in the expectation that programme applications have a clear Theory of Change (for instance, [51]). As this pattern spreads to other research funders, the case for mixed methods will become correspondingly stronger.

The mixed methods research paradigm is well established and widely adopted, with health informatics perhaps ahead of the general information systems discipline in its use of this paradigm. Practical research questions, not abstract philosophy, must surely take precedence in the selection of methods. Mixed method evaluations are essential for evidence-based health informatics.

Recommended further reading

1. The Campbell Collaboration, *Background*, <http://www.campbellcollaboration.org/background/index.php>, last access 11 February 2016.
2. Center for Theory of Change, *Publications*, <http://www.theoryofchange.org/library/publications>, last access 11 February 2016.
3. M. Eccles, J. Grimshaw, M. Campbell, C. Ramsay, Research designs for studies evaluating the effectiveness of change and improvement strategies, *Qual Saf Health Care* **12**(1) (2003), 47-52.
4. S. Shepperd, S. Lewin, S. Straus, M. Clarke, M.P. Eccles, R. Fitzpatrick, G. Wong, A. Sheikh, Can we systematically review studies that evaluate complex interventions?, *PLoS Med* **6**(8) (2009), e1000086.
5. G. Wong, T. Greenhalgh, G. Westhorp, J. Buckingham, R. Pawson, RAMESES publication standards: realist syntheses, *BMC Medicine* **11**(1) (2013), 21.
6. M. Zachariadis, S. Scott, M. Barrett, Methodological implications of critical realism for mixed-methods research, *MIS Quarterly* **37**(3) (2013), 855-879.

Food for thought

1. How could the mixed method paradigm guide evaluation design in a scenario where there are contradictory stakeholder expectations about the purpose and scope of the study?
2. What types of research questions might suggest a dialectic rather than a synthetic integration of mixed method findings?
3. How might a revised hierarchy of evidence for health informatics be constructed?

References

- [1] V. Venkatesh, S.A. Brown, H. Bala, Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems, *MIS quarterly* **37**(1) (2013), 21-54.
- [2] P.J. Scott, J.S. Briggs, A pragmatist argument for mixed methodology in medical informatics, *J Mix Methods Res* **3**(3) (2009), 223-241.
- [3] Sage Publications, *Journal of Mixed Methods Research*, <http://us.sagepub.com/en-us/nam/journal-of-mixed-methods-research/journal201775>, last access 11 February 2016.

- [4] C. Teddlie, A. Tashakkori, *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*, Sage Publications, 2009.
- [5] S. Hesse-Biber, Mixed Methods Research: The “Thing-ness” Problem, *Qual Health Res* **25**(6) (2015), 775-788.
- [6] D.T. Campbell, D.W. Fiske, Convergent and discriminant validation by the multitrait-multimethod matrix, *Psychol Bull* **56**(2) (1959), 81.
- [7] R.B. Johnson, A.J. Onwuegbuzie, L.A. Turner, Toward a definition of mixed methods research, *J Mix Methods Res* **1**(2) (2007), 112-133.
- [8] N.K. Denzin, *The research act: a theoretical introduction to sociological methods*, McGraw Hill, New York, 1978.
- [9] L.S. Giddings, Mixed-methods research: Positivism dressed in drag? *J Res Nurs* **11**(3) (2006), 195-203.
- [10] A. Tashakkori, J.W. Creswell, Editorial: The New Era of Mixed Methods, *J Mix Methods Res* **1**(1) (2007), 3-7.
- [11] MMIRA, *Mixed Methods International Research Association*, <http://mmira.wildapricot.org>, last access 11 February 2016.
- [12] S. Ozawa, K. Pongpirul, 10 best resources on... mixed methods research in health systems, *Health policy and planning* **29**(3) (2014), 323-327.
- [13] S.N. Hesse-Biber, R.B. Johnson, *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*, Oxford University Press, 2015.
- [14] R. Pawson, A. Manzano-Santaella, A realist diagnostic workshop, *Evaluation* **18**(2) (2012), 176-191.
- [15] T.D. Jick, Mixing qualitative and quantitative methods: Triangulation in action, *Adm Sci Q* **24**(4) (1979), 602-611.
- [16] J.L. Liu, J.C. Wyatt, The case for randomized controlled trials to assess the impact of clinical information systems, *J Am Med Inform Assoc* **18**(2) (2011), 173-80.
- [17] P. Craig, P. Dieppe, S. Macintyre, S. Michie, I. Nazareth, M. Petticrew, *Developing and evaluating complex interventions: new guidance*, Medical Research Council, London, 2008.
- [18] V. Yank, D. Rennie, L.A. Bero, Financial ties and concordance between results and conclusions in meta-analyses: retrospective cohort study, *BMJ* **335**(7631) (2007), 1202-5.
- [19] S.R. Forsyth, D.H. Odierna, D. Krauth, L.A. Bero, Conflicts of interest and critiques of the use of systematic reviews in policymaking: an analysis of opinion articles, *Syst Rev* **3** (2014), 122.
- [20] A. O'Cathain, E. Murphy, J. Nicholl, Three techniques for integrating data in mixed methods studies, *BMJ* **341** (2010), c4587.
- [21] M.D. Fetter, L.A. Curry, J.W. Creswell, Achieving integration in mixed methods designs—principles and practices, *Health Services Research* **48**(6pt2) (2013), 2134-2156.
- [22] J. Greene, J. Hall, *Dialectics and pragmatism: Being of consequence*, in: A. Tashakkori, C. Teddlie, editors, *SAGE Handbook of Mixed Methods in Social and Behavioral Research*. 2nd ed, SAGE Publications, London, 2010.
- [23] J.S. Ash, D.F. Sittig, R. Dykstra, E. Campbell, K. Guappone, The unintended consequences of computerized provider order entry: Findings from a mixed methods exploration, *Int J Med Inform* **78** (2009), S69-S76.
- [24] J.I. Westbrook, J. Braithwaite, A. Georgiou, A. Ampt, N. Creswick, E. Coiera, R. Iedema, Multimethod evaluation of information and communication technologies in health in the context of wicked problems and sociotechnical theory, *J Am Med Inform Assoc* **14**(6) (2007), 746-55.
- [25] R. Wu, P. Rossos, S. Quan, S. Reeves, V. Lo, B. Wong, M. Cheung, D. Morra, An evaluation of the use of smartphones to communicate between clinicians: a mixed-methods study, *J Med Internet Res* **13**(3) (2011), e59.
- [26] T. Greenhalgh, K. Stramer, T. Bratan, E. Byrne, J. Russell, H.W. Potts, Adoption and non-adoption of a shared electronic summary record in England: a mixed-method case study, *BMJ* **340** (2010), c3111.
- [27] K. Zheng, R. Padman, M.P. Johnson, H.S. Diamond, Understanding technology adoption in clinical care: Clinician adoption behavior of a point-of-care reminder system, *Int J Med Inform* **74**(7-8) (2005), 535-543.
- [28] C.K. McMullen, J.S. Ash, D.F. Sittig, A. Bunce, K. Guappone, R. Dykstra, J. Carpenter, J. Richardson, A. Wright, Rapid assessment of clinical information systems in the healthcare setting: an efficient method for time-pressed evaluation, *Methods Inf Med* **50**(4) (2011), 299-307.
- [29] B. Blackwood, P. O'Halloran, S. Porter, On the problems of mixing RCTs with qualitative research: the case of the MRC framework for the evaluation of complex healthcare interventions, *J Res Nurs* **15**(6) (2010), 511-521.
- [30] T. Greenhalgh, J. Russell, Why do evaluations of eHealth programs fail? An alternative set of guiding principles, *PLoS Med* **7**(11) (2010), e1000360.
- [31] B. Marchal, G. Westthorp, G. Wong, S. Van Belle, T. Greenhalgh, G. Kegels, R. Pawson, Realist RCTs of complex interventions - an oxymoron, *Soc Sci Med* **94** (2013), 124-8.

- [32] T. Kuhn, *The structure of scientific revolutions*, University of Chicago Press, Chicago, 1962.
- [33] K.R. Howe, Against the quantitative-qualitative incompatibility thesis, *Educ Res* **17**(8) (1988), 10-16.
- [34] J. Paley, R. Lilford, Qualitative methods: an alternative view, *BMJ* **342** (2011), d424.
- [35] G. Biesta, *Pragmatism and the philosophical foundations of mixed methods research*, in: A. Tashakkori, C. Teddlie, editors, SAGE Handbook of Mixed Methods in Social and Behavioral Research. 2nd ed, SAGE Publications, London, 2010.
- [36] C. Bonell, A. Fletcher, M. Morton, T. Lorenc, L. Moore, Methods don't make assumptions, researchers do: a response to Marchal et al., *Soc Sci Med* **94** (2013), 81-2.
- [37] N.K. Denzin, Y.S. Lincoln, *Introduction: The discipline and practice of qualitative research*, in: N.K. Denzin, Y.S. Lincoln, editors, *The Sage Handbook of Qualitative Research*, Sage, London, 2005.
- [38] Oxford English Dictionary, "holism, n.", Oxford University Press, 2015.
- [39] R. Pawson, *Evidence-Based Policy: A Realist Perspective*, SAGE Publications, London, 2006.
- [40] K.L. Salter, A. Kothari, Using realist evaluation to open the black box of knowledge translation: a state-of-the-art review, *Implement Sci* **9** (2014), 115.
- [41] C. Bonell, A. Fletcher, M. Morton, T. Lorenc, L. Moore, Realist randomised controlled trials: a new approach to evaluating complex public health interventions, *Soc Sci Med* **75**(12) (2012), 2299-306.
- [42] G.F. Moore, S. Audrey, M. Barker, L. Bond, C. Bonell, W. Hardeman, L. Moore, A. O'Cathain, T. Tinati, D. Wight, J. Baird, Process evaluation of complex interventions: Medical Research Council guidance, *BMJ* **350** (2015), h1258.
- [43] I. Vogel. *Review of the use of 'Theory of Change' in international development*, DFID, London, 2012.
- [44] D.H. Taplin, H. Clark, E. Collins, D.C. Colby. *Theory of Change - Technical papers*, ActKnowledge, New York, 2013.
- [45] M.J. De Silva, E. Breuer, L. Lee, L. Asher, N. Chowdhary, C. Lund, V. Patel, Theory of Change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions, *Trials* **15** (2014), 267.
- [46] A. Blamey, M. Mackenzie, Theories of Change and Realistic Evaluation: Peas in a Pod or Apples and Oranges? *Evaluation* **13**(4) (2007), 439-455.
- [47] National Institute for Health Research, *Programme Development Grants*, <http://www.nihr.ac.uk/funding/programme-development-grants.htm>, last access 11 February 2016.
- [48] A. O'Cathain, E. Murphy, J. Nicholl, Multidisciplinary, interdisciplinary, or dysfunctional? Team working in mixed-methods research, *Qual Health Res* **18**(11) (2008), 1574-85.
- [49] Oxford Centre for Evidence Based Medicine, *Levels of Evidence*, <http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009>, last access 11 February 2016.
- [50] T. Greenhalgh, H.W. Potts, G. Wong, P. Bark, D. Swinglehurst, Tensions and paradoxes in electronic patient record research: a systematic literature review using the meta-narrative method, *Milbank Q* **87**(4) (2009), 729-88.
- [51] UK Department for International Development, *Building Resilience and Adaptation to Climate Extremes and Disasters (BRACED) Programme: Full proposal application form*, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/286953/full-proposal-app-form.docx, last access 11 February 2016.

Evaluation of People, Social, and Organizational Issues – Sociotechnical Ethnographic Evaluation

Bonnie KAPLAN^{a,1}

^a Yale Center for Medical Informatics, Yale University, New Haven, CT, USA

Abstract. Sociotechnical approaches are grounded in theory and evidence-based. They are useful for evaluations involving health information technologies. This contribution begins with an overview of sociotechnical theory and ethnography. These theories concern interactions between technology, its use, people who use or are affected by it, and their organizational and societal situations. Then the contribution discusses planning and designing evaluations, including frameworks and models to focus an evaluation, and methodological considerations for conducting it. Next, ethical issues and further challenges and opportunities are taken up. Concluding case examples, referenced throughout, illustrate how good evaluations provide useful results to help design, implement, and use health information technologies effectively.

Keywords. Evaluation studies, organizational studies, ethnography, medical informatics, qualitative research, qualitative evaluation, organizational culture, organizational case studies.

1 Introduction

Successful implementation involves interactions and mutual adjustments among an information technology application and the organization, people, and practices where it is used. Sociotechnical evaluation analyzes this interplay between technologies and social and technical systems. It emphasizes how people, organizations, professions, culture, work practice, ethical issues, social and political environment, and the like, all interact and change each other over time. Sociotechnical analyses assess how information technology and workflow influence each other; how clinical and patient roles relate to technological use; how useful and usable health information technologies are; and what consequences, patient safety issues, or user responses might occur. They involve considering these interdependent elements as a holistic dynamic network rather than as fixed pre-defined separate domains [1,2,3,4].

For example, Example 1 indicates that using images, and incorporating clinical images into on-line electronic patient records, depends not only on the computer system, but also on interwoven issues of expertise, trust and relationships among colleagues, clinical knowledge of individual patients, institutional priorities, how

¹ Corresponding author: Prof. Dr. Bonnie Kaplan, Yale Interdisciplinary Center for Bioethics, Box 208293, 238 Prospect Street, New Haven, CT 06511, United States, bonnie.kaplan@yale.edu.

conveniently system access fits into a busy and frequently interrupted day full of *ad hoc* conversations, ways images are interpreted and their clinical meanings negotiated, and other *socio* parts of sociotechnic. That is how sociotechnical systems work, and how sociotechnical analyses can be helpful.

2 Sociotechnical Theory

Sociotechnical approaches incorporate theories and evidence from multiple disciplines. Key theoretical features include examining technologies as they actually are used in natural settings to investigate how technical and physical work settings affect their use; how users negotiate, re-negotiate, interpret, and re-interpret features of the technology; and relationships among the social and technical components of these emergent processes as they unfold over time. The approaches are based on an understanding that a new information technology and the social system where it is introduced change each other as different parties pursue different goals [4,5]. These approaches are not deterministic, nor do they understand technological development in terms of a rational, linear sequence. Instead, they emphasize evolving processes and interactions so that no factor acts in isolation from others, or has a uni-directional impact. They see processes and causes interacting in multiple causal directions and relationships.

Sociotechnical principles developed as part of the Tavistock Institute's post-WW II analysis of British industries. They emphasized designing work for workers' interests and quality of working life [4]. By the 1990s, sociotechnical ideas had been introduced into health informatics, as were social interactionist approaches – approaches that consider relationships between system, individual, and organizational characteristics and effects among them – which now would be labeled "sociotechnical" [1,2,3,6,7,8,9]. Sociotechnical theory in health informatics, then, has roots in traditional sociotechnical research, ergonomics, social construction of technology, technology-in-practice, and social informatics [5]. To these antecedents, I would add theories of change.

Informatics systems introduce change which may be welcome, or disruptive, to the individual and the organization. Sociotechnical theory conceptualizes organizational change as interacting components – for example, Leavitt's well-known diamond model of people, task, technology, and structure [10] – each responding to a +change in any other so as to maintain organizational homeostasis, with the interactions themselves being most important. Other theories of change based on the foundational work of Rogers [11] and Lewin [12] characterize it as a dynamic process that proceeds through stages involving multiple actors with different concerns and perceptions of benefit. These actors include experts, sponsors, and people adopting (or not adopting) the change. These actors are connected and communicate through various social, organizational, social, and cultural channels. The change occurs, then, at individual, group, organizational, and cultural levels. Any of the stages, actors, system components, and units of analysis could be the focus of evaluation.

3 Ethnography

Ethnographic approaches explore how users experience health information technology and why they interact with it as they do. They involve getting to know and documenting the people and culture by spending time and participating in the setting

under study [13]. Ethnography expresses findings in terms meaningful to the people involved. This enables people to recognize themselves and thereby makes those findings more convincing and relevant. Ethnographic sociotechnical evaluation can help prevent difficulties through better needs analysis, system design and implementation practices, understanding what people do when working with the technologies, and identifying why they view and use the technologies in those ways.

Ethnography involves starting with a sense of what to investigate and progressively sharpening the investigation as more is learned. This is different from beginning with immutable testable hypotheses, *a priori* research questions or categories, and a pre-set research design. Instead, the study evolves and changes according to what is learned as it proceeds. Because sociotechnical systems are dynamic, freezing a research design before beginning may turn out to poorly match the situation at hand as it develops. Ethnographic methods are particularly valuable in natural, uncontrolled settings. They allow for adjusting a study in a fluid environment where unanticipated findings emerge and situations change.

Methodologically, Examples 1 and 2 are ethnographic. Ethnographies tend to emphasize the people involved and explore their situations. The main general investigative questions are:

- (1) What is happening here?
- (2) Why is it happening?
- (3) How has it come to happen in this way?
- (4) What do the people involved think is happening?
- (5) How are they responding to what is happening?
- (6) Why are they responding that way? [14]

The key question is “Why?”: Why are the people who are involved actually involved; why do they think and react as they do; why do they use the technology as they do; why are they interacting as they are; what meanings do they attribute to the technology, health and disease, their roles, and what they do; and why those meanings?

To answer, ethnographic work uses open-ended evaluation questions, qualitative data collection and analysis, interpretive and multi-level data analysis, a focus on the lived experience and its meaning to those involved, emergent findings, and making tacit knowledge and practice manifest. Because it enables a deep understanding of what is going on, wiser decisions and actions may be based on those findings, and theoretical insights may be developed [14].

In Example 2, sociotechnical approaches revealed emergent, unexpected findings involving more general interrelationships between work and technology use. The analysis reinforces the sociotechnical stance that the technology does not stand alone, the social system (in this case, laboratory management, laboratory work, and hiring practices) does not stand alone, but the two mutually affect and change each other. The ethnographic approach enabled better understanding of how laboratory work was understood. This resulted from resolving seemingly divergent findings from multiple sources of data through an interpretation that accounted for all data, in this case, the job-orientation model that relates how people think of their job to how they think about computer systems introduced into their work. This rich result contributed to theory.

4 Theory Development

Example 2 also exemplifies other theoretical points. The evaluation contributed to the idea that “the same” system is not the same for all concerned, which also was found in an evaluation of an automated telephone counseling system [15]. Similarly, “success” may be defined and experienced differently among different groups and individuals at different times [16]. Further, as also evident in studies that contributed to the idea of the importance of fit between a technology and an organization, “fit” has to be produced actively and changes over time [3,17].

The two examples contributed to another theoretical insight as well. The findings inspired a framework helpful in future studies: the 4Cs of communication, care (or, if outside of clinical institutions, whatever else the mission of the organization is), control, and context [2,8,18]. In Example 1, on-line images improved communication and care, raised control issues, and occurred in the different contexts of a government and academic medical center. The laboratory information system in Example 2 also improved communication and, therefore, care; highlighted control issues; and took account of the context of different laboratories and technologists in the job-orientation model. Frameworks like 4Cs can be useful for evaluation planning and design.

5 Planning and Designing Ethnographic Sociotechnical Evaluation

The multiplicity of interacting systems and sub-systems presents a wide range of choice for how to design an evaluation. To choose among the possibilities, decisions are needed concerning how to focus an evaluation, when to evaluate, and how to evaluate.

5.1 What to Evaluate

To answer the key evaluation question of what is happening and why, it is hard to know at the outset what of all the activity and who of all those involved will be important. Theories, models, and frameworks can help to target what is most relevant for the situation at hand. They provide a lens through which situations can be analyzed and understood; highlight what is important; explain how various factors, influences, and considerations interrelate; help organize and explain findings; and lead to predictions for further investigation and planning. Their power comes from emphasizing only some aspects of the area under study. Because each necessarily leaves out aspects that may turn out to be important, it can be helpful to use more than one theory, model, or framework. Sociotechnical evaluation lends itself to just that. 4Cs, discussed in Section 4, brings attention to issues of communication, care, control, and context. Siting and Singh’s model focuses on hardware and software; clinical content; human-computer interface; people; workflow and communication; organizational policies, procedures, and culture; external rules, regulations, and pressures; and system measurement and monitoring [19]. An additional set of evaluation questions, based on those of Anderson and Aydin [7], could be:

- (1) Does the system work as designed?
- (2) Is it used as anticipated?
- (3) Does it produce the desired results?

- (4) Does it work better than what it replaced?
- (5) Is it cost-effective?
- (6) How well have individuals been trained to use it?
- (7) What are changes in departmental interaction, delivery of care, patient safety, control and power in the organization, or the healthcare system at large?
- (8) How do the system and these changes relate to the practice setting?

Combining theories, models, or frameworks can help an evaluator choose potential evaluation questions. What purpose the evaluation serves also is important when choosing a focus. Table 1 gives some examples.

Table 1. How evaluation purpose can affect evaluation focus.

If the purpose of the evaluation is	The evaluation could focus on
• Technical	• System requirements
• Economic	• Cost/benefit
• Clinical	• Patient Care
• Education	• Students' grades, learning outcomes
• Research	• Access to literature, data
• Policy	• Cost, utilization
• Usefulness	• User satisfaction, degree of use

Just as the system, the users, and the context interact and shape each other, the evaluation context and environment affect how the study is conducted over time. These include:

- (1) purpose of the system, which may be for research and development, a demonstration project, or a commercial product;
- (2) organizational commitment, which might be to continue, maintain, or quash the system, or to evaluate it;
- (3) who the client is;
- (4) how evaluation results will be used;
- (5) budget and time frame;
- (6) evaluator skills and expertise;
- (7) who the research subjects are; and
- (8) the people who are involved.

Considerations about these people include:

- (7) how the need for the system and for the evaluation was determined, and by whom;
- (8) what needs the system and the evaluation meet, and whose needs they are;
- (9) who will be using the system, doing data entry, or receiving outputs;
- (10) what users' attitudes towards the system, and towards the evaluation, are;
- (11) who was involved in needs assessment, design, and testing, and why those where the people involved;
- (12) whether potential users perceive a need for the system;
- (13) whose interests the system or the evaluation serves, or appears to others to serve; and
- (14) what different parties want to know.

Knowing the environment and people involved can alert the evaluator to considerations that should be examined further.

5.2 When to Evaluate

Sociotechnical ethnographic evaluations can be done at any stage, or multiple stages, of system development or implementation. When to evaluate depends on the purpose of the evaluation, as in the two examples. There is no need for concern that study results or even conducting the study will affect the object of study. It will. A moving target is assumed. Evaluation, then, can be used to influence needs assessment, analysis, design, implementation, and how a system is used without “tainting” either the process or the rigor of the study. In fact, it is wise to feed what is learned back into the process so that it proceeds more smoothly.

5.3 How to Evaluate

Choosing methods depends on evaluation questions, evaluator skills and expertise, and budget and time table. The theoretical underpinnings of sociotechnical approaches suggest methods and research designs that are flexible and encourage emergent, unexpected findings. Rather than the usual impact studies that characterize much medical research – randomized controlled trials and experimental designs to test hypotheses – interactionist (i.e. where subsystems and system components interact over time) sociotechnical study designs are preferred. Table 2 indicates some ways impact and interactionist studies differ.

Table 2. Differences between impact studies and interactionist studies.

	<i>Impact</i>	<i>Interactionist</i>
<i>Epistemology</i>	Objectivist	Objectivist or Subjectivist
<i>Purpose</i>	Factors	Process
	Variance	
<i>Methods</i>	Quantitative	Qualitative
<i>Causality</i>	Uni-directional	Multi-directional
<i>Question</i>	What	Why

Sociotechnic approaches examine how peoples’ practices are situated in their environments and how the actors and technological change interact. These studies are best done *in situ* using methods appropriate to naturalistic settings and changing circumstances. Ethnographic sociotechnical evaluation is interactive not only in examining interactions among the social and technical components of the system under study, but also among components of research design. What should be studied and what the research questions are depends on the purposes, methods, conceptual concepts, and validity issues involved, and each of these shapes the others [20]. Study design, then, should be longitudinal, modifiable, and flexible over time. Because evaluation can help direct a project, it can be both formative and summative, and should focus on a variety of concerns reflecting the various actors involved. Employing multiple methods is beneficial because different data sources provide different data [18]. Different informants may have different focuses, report processes that are different from what the evaluator observes, and behave differently from the way they indicate on surveys or in laboratory settings. The challenge is to make sense of these

differences. If data do not converge, a richer understanding develops through accounting for apparent contradictions, as in the laboratory information system study (Example 2). Multiple methods and data sources lead to robust results.

Qualitative methods were used in the two examples. They are especially valuable for sociotechnical ethnographic evaluation. Data collection methods include participant observation; observation; unstructured or semi-structured interviews; focus groups; surveys with open-ended questions; analysis of artifacts like documents, images, texts, or drawings; and the researcher's own impressions and reactions. Analysis methods include coding, contextual or narrative analysis, analytic memos, and displays. Data analysis involves constantly integrating and analyzing voluminous, mostly textual, data from multiple sources. Interpretations and hypotheses continually are formulated, tested, and verified or discarded through a process of on-going data analysis and writing that assesses whether they make sense in light of existing and future data. What seems most interesting, relevant, or important progressively becomes clearer [14].

Qualitative data analysis software is a boon to managing and analyzing the volumes of data an evaluation study produces, but it does not do the analysis *per se*. The evaluator still needs to figure out how to interpret data. It helps in this process to focus on:

- (1) how people use words and what they mean by them—what is meant by “work” in the laboratory (Example 2) or “see[ing] what’s really going on” in an image (Example 1);
- (2) what people say and do, and under which circumstances they say and do it—how the clinicians in the second imaging study (Example 2) negotiated what images meant;
- (3) how people justify or give reasons for what they say, do, believe, etc.—comments laboratory technologists wrote about why the new system was a “hassle” or improved reporting (Example 2);
- (4) what does not seem to make sense (the puzzles)—how a laboratory technologist’s job does not change when the technologists’ tasks change (Example 2); and
- (5) how to make sense of *all* the data.

Focusing this way helps produce evaluations that get at what it means to the people involved to use health information technologies. Paying close attention to who the people are, what they think, what they do in real-life settings, and how they differ, helps explain how all that interacts with health information technology development, implementation, adoption, and use – in other words, how the social and technical subsystems interact. The end result, then, goes well beyond simply reporting data. It requires solving puzzles by accounting for all data in a way that focuses on what the technology means to the participants, why it means that, and what the implications are. Explaining the data in this way helps make tacit knowledge, assumptions, meanings, and values explicit, so they can be taken into account. It tells a coherent, compelling story that is useful, and makes theoretical contributions by both drawing on theory to produce an interpretation and also, as in the examples, possibly develop new theory.

5.4 How to Validate Evaluation Results

Qualitative researchers collect rich data and produce interpretations that account for it all through a process known as triangulation. Particular attempts are made to collect

data that may contradict the developing interpretation. Data is continually collected until no new information seems to be possible, which is known as reaching saturation. The people involved in the study are asked for feedback and for their responses to the developing interpretation in a process known as member-checking, and what they say becomes new data [14]. A neutral partner can review data and how it is interpreted. Similarly, research team members can test each other's ideas, methods, and interpretations. Eventually, reviewers and other researchers judge the work, just as in any other form of research. Reproducability is impossible; every situation, evaluator, and study is different. The goal is transferability, so that significant insights can be developed, theoretical contributions can be made, and the knowledge gained can be applied elsewhere.

6 Sociotechnical Ethnographic Evaluation Research Ethics

Evaluators face ethical decisions even before beginning an evaluation and thereafter. In addition to usual research ethics issues, additional concerns arise in sociotechnical ethnographic evaluation. A few of them are mentioned here. Special considerations involve informed consent, privacy and confidentiality, social justice, practitioner research, power, reciprocity, relevance, and how the research is used [21].

As in other fieldwork, interpersonal relationships develop between evaluators and participants, raising questions of just what those relationships should be. The evaluator may be privy to material that those involved did not give consent for or see people who were not asked for permission. When a new technology is introduced it is hard to anticipate how people will react, making consent even more problematic [22]. The evaluator may observe what could be unethical behavior, or be asked to engage in behavior that some may consider unethical. A sociotechnical viewpoint involves sensitivity to ethical questions like who defines, and should define, the evaluation questions, interpretation, and use of results, and whose interests are served by the evaluation. The evaluation, too, likely will involve the goals, values, and assumptions incorporated into the technology, how it is implemented, how people are expected to use it, and effects expected from it, also raising ethical concerns.

7 Future Challenges and Opportunities

To date, evaluation mostly concerns visible, tangible health information technologies in physical settings. Newer developments—virtual health care delivery, distributed integrated health care organizations, virtual workers, fluid organizational boundaries, social networks, telehealth and mobile health applications, avatars and artificial intelligence—make *in situ* studies more difficult, especially if health care delivery crosses jurisdictional boundaries. Adding to the complexity is the need for multi-site studies that include community, home, or other non-academic locations with geographic or national variation. There is room for sociotechnical evaluation study designs and methods that address these challenges while also contributing to much-needed methods to assess patient outcomes better [3].

8 Conclusion

Sociotechnical ethnographic evaluation focuses primarily on the people in addition to the technology. Contributors to system “success” *are* sociotechnical. By focusing on technologies as they actually are used, in the settings in which they are used, and seeing how people negotiate and reinterpret the technologies as the social and technical systems interact with each other, sociotechnical ethnographic evaluation can contribute to theory and practice while improving health information technologies and patient care.

Example 1 – Clinical Imaging Systems

Administrators and clinicians differed about the value of a new system that integrated patient record textual, numeric, and image data [23]. This raised an administrative control issue concerning decisions about continued development. Also, previously the department where an image was produced kept the image, but now images were available to all, which potentially created another control issue.

In a week of interviews and observations, we investigated what clinicians thought about the benefits of the system. Clinicians told us that having the images available as part of the on-line patient record improved communication and consultation, so improved clinical decisions, and hence, patient care. Because “a written report won’t convey everything,” and “you don’t know [if the report] is an accurate description,” now, clinicians said, they “know what’s there,” they could “look through a patient and see,” “see what’s really going on.” That way, they did not need to repeat procedures. They could plan treatment better and give students “real” experience through these images.

Elsewhere, I spent a week shadowing a physician as he performed his daily activities. The purpose was to identify how clinical images are used in an academic medical center planning to develop a stand-alone imaging system [24]. The physician was interrupted constantly, frequently telephoned for information, talked about patients with other doctors he met fortuitously on the stairway, and consulted with Pathology and Cytology after receiving reports that slides were “not diagnostic,” or “inadequate for evaluation.” The person reading the slides told the physician that he had a “gut feeling” that the cells indicated cancer, though “quantitatively it was a little short” and showed him why. At his weekly radiology conference, they discussed each patient’s images, asking each other about the patient and what they saw, or thought they saw, on the image. For the physician I shadowed, mutually viewing images was improving communication and clinical decision making, and seeing the images was better than reading a report. However, reading an image was not a matter of “see [ing]what’s really going on,” but of interpreting the image in light of expertise and experience, clinical knowledge of that particular patient, and discussing all that.

In these studies, clinicians thought of the benefits of viewing images as a whole, not as a separate part of patient care. They thought having those images improved care and decision making. They considered the images objective, talking of them as showing, all by themselves, what was “really going on.” Yet, the studies indicate that what an image means and what clinical decisions should be based on it depends on far more than simply having the image. In these evaluations, the meanings of those images were being negotiated through collegial interchanges, though neither clinicians nor

system developers acknowledged it. Even though the same could be said about paper and film-based images, health information technologies often are premised on a belief that providing information alone is enough because it speaks for itself. This belief affects design, implementation, and use.

Example 2 – Clinical Laboratory System

We investigated the impact of a new system on laboratory work in a longitudinal study ranging from pre- to post-implementation. More in line with a sociotechnical ethnographic approach, we also sought to identify what happens when an academic medical center converts from a manual to automated system for ordering clinical laboratory tests and reporting test results [25]. The study included interviews, observations, participant observation, and surveys.

Technologists' responses to scaled-response survey questions indicated no change in laboratory work. Nevertheless, it was clear from their comments in open-ended questions that work was changing. Some technologists reported being "happy" because of fewer abusive telephone calls. They also liked the more legible, timely, and complete laboratory reports. Others, instead, reported on the "hassle" of having to interrupt their work to enter test results into the computer. We realized that the first group of technologists thought of their job as providing laboratory test results, an outcome- or product-oriented view of laboratory work. The other group of technologists thought of their work as doing laboratory tests, a more process-oriented view in which they saw the new computer system as a "hassle" that took them away from the laboratory bench. This job-orientation model applied not only to individual technologists, but also to the fit between system and different laboratories. The same laboratory information system used in all the laboratories was not "the same" for everyone, nor even every laboratory. Instead, it was viewed differently in ways that related to job orientation. Moreover, it was apparent that being able to work with the computer system was a new criterion for being a laboratory technologist.

The findings can be reported in terms of improving communication between the laboratories and the clinicians by producing better and more timely laboratory reports, thereby improving care. Laboratory technologists fielded fewer telephone calls asking for laboratory results. Control issues arose over laboratory work, and the different context of each laboratory was related to how technologists viewed the new system. In particular, how the laboratories, as well as individual technologists and laboratory directors, saw the nature of laboratory work was key to understanding their reactions. In interviews, directors had told us that the new system would not change technologists' jobs. If they had realized that there were different views of laboratory work, that laboratory work was now different, and that these differences would matter in how technologists and laboratories related to the new system, they could have prepared staff better.

Recommended further readings

1. M. Berg, Patient care information systems and health care work: a sociotechnical approach, *International Journal of Medical Informatics* **55** (1999), 87-101.
2. B. Kaplan, N. T. Shaw, Future directions in evaluation research: people, organizational, and social issues, *Methods of Information in Medicine* **43** (2004), 215-231.
3. B. Kaplan, J.A. Maxwell, *Qualitative research methods for evaluating computer information systems*, in: Evaluating the Organizational Impact of Healthcare Information Systems, 2nd ed., J.G. Anderson, C.E. Aydin, eds., Springer, New York, 2005. pp. 30-55.
4. M. H. Harrison, R. Koppel, S. Bar-Lev, Unintended consequences of information technologies in health care - an interactive sociotechnical analysis, *Journal of the American Medical Informatics Association* **14** (2007), 542-549.
5. S. Sawyer, M. Jarrahi, *Sociotechnical approaches to the study of information systems*, in: Computing Handbook: Information Systems and Information Technology, 3rd ed., v. 2, H. Topi, A. Tucker, eds., Chapman and Hall/CRC, Boca Raton, FL, 2014. pp. 5-1 – 5-27.

Food for thought

1. What are the distinguishing features of sociotechnical theory? What advantages and disadvantages would each feature bring to an evaluation?
2. How might ethnography influence evaluation? What are the pros and cons?
3. What are the benefits and pitfalls of using models, theories, or frameworks to focus an evaluation?
4. How would you address the challenges you would expect to face in qualitative data collection and analysis?
5. How would you design a sociotechnical ethnographic evaluation outside an institutional setting, for example, of a smartphone application for managing a diabetic teenager's diet or an elderly person's depression? What evaluation questions would you investigate? How would you go about investigating them? What ethical challenges might arise?

References

- [1] M. Berg, Patient care information systems and health care work: a sociotechnical approach, *International Journal of Medical Informatics* **55** (1999), 87-101.
- [2] B. Kaplan, Evaluating informatics applications - some alternative approaches: theory, social interactionism, and call for methodological pluralism, *International Journal of Medical Informatics* **64** (2001), 39-56.
- [3] B. Kaplan, N. T. Shaw, Future directions in evaluation research: people, organizational, and social issues, *Methods of Information in Medicine* **43** (2004), 215-231.
- [4] S. Sawyer, M. Jarrahi, *Sociotechnical approaches to the study of information systems*, in: Computing Handbook: Information Systems and Information Technology, 3rd ed., v. 2, H. Topi, A. Tucker, eds., Chapman and Hall/CRC, Boca Raton, FL, 2014. pp. 5-1 – 5-27.

- [5] M. H. Harrison, R. Koppel, S. Bar-Lev, Unintended consequences of information technologies in health care - an interactive sociotechnical analysis, *Journal of the American Medical Informatics Association* **14** (2007), 542-549.
- [6] M. Berg, C. Langenberg, I.V.D. Berg, J. Kwakkernaat, Considerations for sociotechnical design: experiences with an electronic patient record in a clinical context, *International Journal of Medical Informatics* **52** (1998), 243-251.
- [7] J.G. Anderson, C.E. Aydin, *Overview. Theoretical perspectives and methodologies for the evaluation of health care information systems*, in: Evaluating the Organizational Impact of Healthcare Information Systems: 2nd ed., J.G. Anderson, C.E. Aydin, eds., Springer, New York, 2005. pp. 5-29.
- [8] B. Kaplan, *Social interactionist framework for information systems studies: the 4Cs*, in: IFIP WG8.2 and WG 8.6 Joint Working Conference on Information Systems: Current Issues and Future Changes, T.J. Larson, L. Levine, J.I. DeGross, eds., International Federation for Information Processing, Helsinki, 1998. pp. 327-339, <http://ifipwg82.org/node/26>, last access 11 February 2016.
- [9] J.G. Anderson, C.E. Aydin, B. Kaplan, *An analytical framework for measuring the effectiveness/impacts of computer-based patient record systems*, in: Proceedings: Twenty-Eighth Hawaii International Conference on Systems Science, J.F. Nunamaker, R.H. Sprague, eds., IEEE Computer Society Press, Los Alamitos, Cal, 1995. pp. 767-776.
- [10] H. J. Leavitt, *Applying organizational change in industry: structural, technological and humanistic approaches*, in: Handbook of Organizations, J. G. March, ed., Rand McNally, Chicago, 1965. pp. 1144-1170.
- [11] E.M. Rogers, *Diffusion of Innovations*, 5th ed., Free Press, New York, 2003.
- [12] K. Lewin, *Group decision and social change*, in: Readings in Social Psychology, 3rd ed., E. Maccoby, T. M. Newcomb, E. L. Hartley, eds., Henry Holt, New York, 1958. pp. 197-211.
- [13] S. Reeves, A. Kuper, B.D. Hodges, Qualitative research methodologies: ethnography, *BMJ* **337** (2008), a1020, 512-514.
- [14] B. Kaplan, J.A. Maxwell, *Qualitative research methods for evaluating computer information systems*, in: Evaluating the Organizational Impact of Healthcare Information Systems, 2nd ed., J.G. Anderson, C.E. Aydin, eds., Springer, New York, 2005. pp. 30-55.
- [15] B. Kaplan, R. Farzanfar, R.H. Friedman, Personal relationships with an intelligent interactive telephone health behavior advisor system: a multimethod study using surveys and ethnographic interviews, *International Journal of Medical Informatics* **71**(1) (2003), 33-41.
- [16] M. Berg, Implementing information systems in health care organizations: myths and challenges, *International Journal of Medical Informatics* **64** (2001), 143-156.
- [17] J. Aarts, H. Doorewaard, M. Berg, Understanding implementation: the case of a computerized physician order entry system in a large Dutch university medical center, *Journal of the American Medical Informatics Association* **11**(3) (2004), 207-216.
- [18] B. Kaplan, Addressing organizational issues into the evaluation of medical systems, *Journal of the American Medical Informatics Association* **4** (1997), 94-101.
- [19] D.F. Sittig, H. Singh, A new sociotechnical model for studying health information technology in complex adaptive healthcare systems, *Quality and Safety in Health Care* **19** (2010), i68-i74.
- [20] J.A. Maxwell, *Qualitative Research Design: An Interactive Approach*, Sage, Thousand Oaks, CA, 1996.
- [21] I.F. Shaw, Ethics in qualitative research and evaluation, *Journal of Social Work* **3** (2003), 9-29.
- [22] B. Kaplan, S. Litewka, Ethical challenges of telemedicine and telehealth, *Cambridge Quarterly of Healthcare Ethics* **17** (2008), 401-416.
- [23] B. Kaplan, H.P. Lundsgaarde, Toward an evaluation of a clinical imaging system: identifying benefits, *Methods of Information in Medicine* **35** (1996), 221-229.
- [24] B. Kaplan, Objectification and negotiation in interpreting clinical images: implications for computer-based patient records, *Artificial Intelligence in Medicine* **280** (1995), 439-454.
- [25] B. Kaplan, D. Duchon, Combining qualitative and quantitative approaches in information systems research: a case study, *Management Information Systems Quarterly* **12** (1988), 571-586.

From Usability Engineering to Evidence-based Usability in Health IT

Romaric MARCILLY^{a,1}, Linda PEUTE^b and Marie-Catherine BEUSCART-ZEPHIR^a

^a*INSERM CIC-IT, Lille; Univ Lille Nord de France; CHU Lille; Lille, France*

^b*Academic Medical Center, University of Amsterdam, Center for Human Factors Engineering of Health Information Technology (HIT-Lab), Amsterdam, The Netherlands*

Abstract. Usability is a critical factor in the acceptance, safe use, and success of health IT. The User-Centred Design process is widely promoted to improve usability. However, this traditional case by case approach that is rooted in the sound understanding of users' needs is not sufficient to improve technologies' usability and prevent usability-induced use-errors that may harm patients. It should be enriched with empirical evidence. This evidence is on design elements (what are the most valuable design principles, and the worst usability mistakes), and on the usability evaluation methods (which combination of methods is most suitable in which context). To achieve this evidence, several steps must be fulfilled and challenges must be overcome. Some attempts to search evidence for designing elements of health IT and for usability evaluation methods exist and are summarized. A concrete instance of evidence-based usability design principles for medication-related alerting systems is briefly described.

Keywords. Usability, human engineering, medical informatics, health informatics, evaluation studies as topic, evidence.

1. Introduction

Studies on Human Factors and usability of Health Information Technology (health IT) are increasingly demonstrating their importance to health IT design, development and implementation [1]. Even if Human Factors and usability are often closely associated, they however do not refer exactly to the same discipline.

According to the International Ergonomics Association, "*Human Factors (or ergonomics) is the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance.*" [2]. Human Factors has a holistic view of the work system. This work system is "*comprised of five elements: the person performing different tasks with various tools and technologies in a physical environment under certain organizational conditions*" [3]. The "tool" (or product or technology) as a topic of research can be described by several characteristics amongst which is usability.

¹ Corresponding author: Dr. Romaric Marcilly, CIC 1403, Maison Régionale de la Recherche Clinique, 6 rue du professeur Laguesse, 59037 Lille Cedex, France, romaric.marcilly@univ-lille2.fr or romaric.marcilly@gmail.com.

Usability is then looked upon as “*the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specific context of use*” [4]. Usability thereby concerns the elements of the graphical user interface, their arrangement, navigational structures, the behaviour of the system in response to users' actions along with the completeness of functions and the work model implemented in the system [5]. Gradually, usability has become a research field in its own right but with the same theoretical, methodological, and empirical roots as Human Factors.

This contribution focuses on how usability research may lead to evidence-based usability practice in the field of health IT.

2. Why is it necessary to consider usability in health IT evaluation?

There are three main categories of reasons accounting for the growing importance of considering usability in the design and implementation of health IT.

2.1. Usage and safety of use reason

Usability is an intrinsic characteristic of a technology that impacts end-users' interaction with the technology; it leads to higher work efficiency in case of good usability, but in case of poor usability it may also slow down user performance, decrease users' satisfaction, and expose users to use errors [6;7]. Then, through its influence on the user, the usability of a technology will indirectly impact the other components of the work system in which this technology is implemented (incl. ensuring patient safety) and the whole work system performance [6-8]. Ultimately, usability flaws in a technology may (i) lead users to reject the technology and / or (ii) even cause harm to patients.² Case studies have identified usability flaws that have had consequences on the quality of the usage of the technology, and subsequently on the outcome of the usage.

For instance, a drop-down menu in a Computerized Physician Order Entry (CPOE) proposing 225 options for medical dosing frequency compels a physician to scroll through the whole list of options. This promoted errors especially for uncommon drug programs. Confused by apparently similar labels, the physician selected the wrong dosing frequency options. As a consequence, a patient received four times excess of Digoxin inducing ventricular fibrillation. Several studies showed that usability-induced use-errors led to patient harm or death: radiation over-dosage errors during radiotherapy [9], dispensing errors with pen injectors [10], or falsely implanted total knee arthroplasties [11]. These insights have led to a growing interest in the effect of the usability of a technology on the system use outcome.

2.2. Regulatory reason

The safety concern led the European Commission to reinforce the "ergonomics" essential requirement for CE marking: the EU revised Medical Device Directive

² See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

(MDD) [12] explicitly requires a safety-oriented usability engineering process to be integrated in the design and development lifecycle of medical devices. In order to adhere to this directive, international standards suggest to implement the User Centred Design process (UCD) during the technology design and development lifecycle (e.g. [13]). Those regulations first dealt with medical devices, and then have been progressively extended to specific types of medical software (e.g. Clinical Decision Support System (CDSS)) [12]. Now, international committees recommend applying UCD to all types of medical systems (including software) (e.g. [1;14]).

2.3. Impact evaluation reason

Over the last decade, requirements over Health Technology Assessment including cost-benefit and medico-economic analyses have been increasing. As a consequence, more and more technologies are expected to undergo some sort of clinical investigation demonstrating their safety and positive clinical impact. However there is one major difference between clinical trials of drugs and clinical investigation studies of health IT and medical devices: the latter are user-dependent. Their efficacy and efficiency depends on their proper use by the end-users (clinicians or patients). When important usability flaws plague the human-machine interface of a product, besides potential erroneous use, users may adopt workaround behaviours to adapt to the poorly usable tool (e.g. [7]). Many of those behaviours are quite personalised and variable. This introduces major potential biases in clinical studies of health IT, as erroneous use, workarounds, and other adapting behaviours inevitably modify the technology efficacy and efficiency. Careful consideration of usability before and during clinical investigation of health technology may help uncover those hidden or intermediary variables and explain puzzling contradictory results [15].

3. Usability Engineering: the User-Centred Design Process

Health technologies should be designed following a safety-oriented UCD process [12;16] in order to ensure that the resulting product is (i) safe to use, (ii) compliant with regulations, and (iii) usable enough to be properly used by end-users, which is a major condition for the technology to achieve its intended (clinical and organisational) impact.

The UCD process is an iterative design and evaluation strategy that considers end-users (i.e. clinicians or patients) by taking into account their needs and by involving them in design and evaluation activities [4]. As described in Figure 1, this process includes four main iterative tasks that may be categorized into specification and evaluation activities.

3.1. Specification activities

First, a sound and precise analysis of the work system in which the technology is to be implemented has to be carried out, including the analysis of the cognitive tasks performed by the end-users [17]. Results depict the whole work system including work partners and the collective and collaborative aspects: needs of the end-users are deduced and potential room for improvement for the current work system is identified. The analysis also allows foreseeing how the technology will support the tasks to be

performed, fulfil users' needs and ultimately improve the work system's efficacy and efficiency. On this basis, specifications for the technology under design are formulated.

Once the context of use has been analyzed, a supplementary source of information, i.e. existing usability design principles, can be used to refine the specifications. Those principles gather knowledge on human capabilities and limitations in a given context. They are more or less generic/specific, some being applicable to any kind of technology and context of use (e.g. [18] for interactive systems), others to a unique type of technology (e.g. [19] for medical alerting systems). Those principles are no substitute for the work system analysis; they provide designers with complementary Human Factors information relevant for the technology under design. Recent studies have shown that applying usability design principles reduces user workload, improves the efficiency of technology, and increases user satisfaction [20].

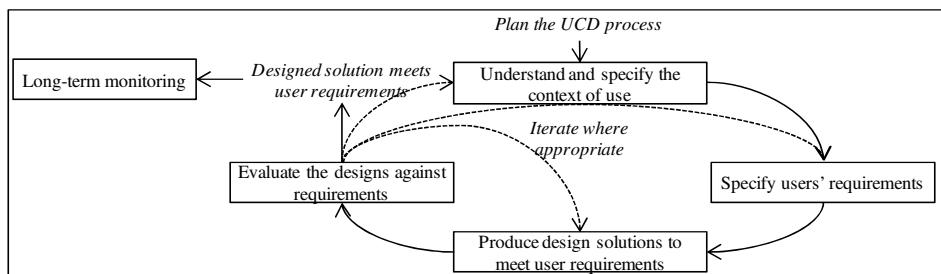


Figure 1. The User-Centred Design process adapted from the ISO 9241-210 [4].

3.2. Evaluation activities

Usability evaluation pursues two main purposes depending on the stage of the system development lifecycle they take place in [13]:

- **Formative usability evaluation** (or "usability verification") consists of iterative and fast evaluation rounds aiming at identifying and fixing usability flaws of the successive versions of the product under development. It applies to early mock-ups and prototypes up to the pre-final version of the product.
- **Summative usability evaluation** (or "usability validation") aims at validating the usability of the final version of the product before its release for clinical use.

Three types of usability evaluation methods are recommended by standards [4]:

- **Expert evaluations** are in-lab methods performed by usability experts without involvement of any end-users. Those methods include heuristic evaluation, where usability experts analyze a user interface by comparing it against a set of usability principles (e.g. [18]), and cognitive walkthrough, where experts step through a user interface for a task, note goals, actions, system responses and potential problems [21]. Those methods require three to five Human Factors experts working in parallel, and enable uncovering of a large number of flaws in a small amount of time. Those methods are part of a prospective approach of the usability: the evaluators' expertise offers insight on what usability problems users might face, in order to fix these problems before the technology is actually used. Experts must own a sound expertise in usability and also in the clinical activity supported by the technology under evaluation. To cope with the problem of clinical expertise of the

evaluators, usability experts sometimes perform the usability evaluation with a clinical expert.

- **User testing and simulation** methods involve observing representative end-users interacting with the technology while carrying out representative tasks. These methods are often carried out as controlled observations in which the behaviours and interactions of the users with the technology are recorded for detailed analysis [22]. They are also often associated with the “think-aloud” method that is considered as the gold standard providing the best insight on user's interaction [23], or with the eye-tracking method [24]. The main difference between user testing and simulation rests on the ecological validity of the evaluation environment: user testing takes place in an office or in a usability lab while simulation requires locating the study in real or realistic settings.³ Both methods can be applied as soon as an interactive mock-up is available; however, due to the costs inherent to the simulation, it is better to perform simulation with (close to) a final product. In terms of results, those methods enable observation of users facing usability flaws and how those flaws impact the usage (including use errors) and the work system (including safety issues).
- **Post market surveillance** is the method with the higher possible fidelity. It enables gathering usability feedbacks once the technology is implemented and used. Data can be collected through direct observation, users' questionnaire or interview, or review of incidents reports [25]. Data collected provide information on usage problems and negative outcomes likely induced by usability flaws in the system. Unintended usage of the technology and workaround behaviours can also be observed. However, the complexity of the work system in which the system is implemented can make it difficult to determine how the usability of the system impacts users and clinical outcomes and which usability issues are root causes.

Those methods have their own specificities and are not equivalent in terms of detection power of usability issues and in terms of types of issues detected [26]. They are often combined together or with other methods (e.g. log analysis, focus groups) and their results are triangulated in order to get a more complete representation of the quality of the technology in terms of usability [27]. Although insights from pre implementation usability evaluations inform redesign of the system, post implementation study is then a necessary step in order to get information on the effectiveness of the pre implementation usability evaluations.

4. Grounding User Centred Design (UCD) in evidence

For several decades UCD has been promoted by reference books, scientific publications, standards and is now imposed by European Union regulation for medical devices and some types of health IT. There is no more need to advocate that carefully taking into account usability during the design process can be beneficial to the design of Health Technologies: it facilitates usage and contributes to fulfilling the medical intention while preventing use-errors leading to patient harm.

³ See also: S. Jensen, Clinical simulation as an evaluation method in health informatics, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

But recommending or imposing usability engineering (UCD) does not mean that it is actually applied to all medical devices and all health IT. Indeed, several recent publications report negative outcomes and patient harm due to usability issues in various types of health technology [9-11;28]. This shows that manufacturers do not (properly) apply UCD so as to decrease the risk of usability-induced use errors. One cause is that manufacturers do not understand how to apply properly UCD. In order to convince all stakeholders, it is necessary to go from an "artisanal" (on a case to case basis) approach towards a UCD grounded in empirical evidence. The evidence will allow drawing upon guidelines for applying the UCD efficiently for each type of health IT and context of use and at each step of the design process. The following sections describe how evidence-based usability knowledge can be produced along with a concrete instance of this knowledge.

4.1. Definition of evidence-based usability

By analogy to evidence-based medicine, evidence-based usability is defined as "*the conscientious, explicit and judicious use of current best evidence in making decisions in design of interactive systems in health care by applying usability engineering and usability design principles that have proven their value in practice*" [29]. This evidence deals with two main topics:

- The design elements of the technology: what are the usability design principles for a given type of technology for which positive value has been demonstrated in practice? What are the instances of usability flaws (violations of those principles) known for this technology (usability mistakes not to make) and what are their consequences on the user and the work system?
- The usability evaluation methods: which method(s) is (are) most suitable at each step of the design process and each type of technology? In which conditions of application are those methods the most efficient? Which combinations of methods have proven their value in practice?

Even if the awareness of designers and researchers in health IT on the need for evidence is increasing, evidence-based usability is still at its infancy. Several steps must be completed and challenges must be overcome to achieve this evidence.

4.2. Steps to get evidence-based usability

The steps to get evidence-based usability are not fundamentally different from those in Health Informatics[30] but some specificities must be pointed out:

- **Perform high quality evaluations.** The main stimulus for evidence is the result of usability and socio-technical evaluations of health IT: descriptions of usability flaws and of their consequences. To ensure the validity of those results, it is necessary to apply properly the right study design⁴ and evaluation method taking precautions against potential biases⁵.

⁴ See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁵ See also: J. Brender, Theoretical basis of health IT evaluation, in: *ibid.*

- **Report evaluations precisely and completely.** The descriptions of the technology evaluated, of the context of evaluation, and of the evaluation method must be reported exhaustively along with the whole set of usability results to allow later meta-analyses.⁶
- **Identify and gather relevant high quality studies.** Scientific publications must be considered. However, not all usability (and socio-technical) evaluations of health IT are published due to non-disclosure agreement and publication reporting biases. To improve the coverage of existing data, grey literature, users' feedbacks to manufacturers, and incidents reports databases (e.g. MAUDE [31]) should be examined too.⁷ Descriptions of incidents may provide information on the consequences of usability flaws on the user and in terms of patient safety.
- **Extract relevant data.** Data must be extracted in a standardized way. Data may be quantitative (e.g. number of errors) but, most of the time, they are qualitative (e.g. description of the usability flaws). For qualitative data, it is necessary to pay the greatest attention to the extraction process to ensure reproducibility.
- **Compare and synthesize publications' findings.** Meta-analyses can be performed (e.g. by comparing the severity of usability flaws in different tools). To go a step further with qualitative data, qualitative comparison analyses should be used to identify the causal contribution of various conditions to an outcome of interest [32]: it allows establishing cause-consequences links between usability flaws, usage problems, and negative outcomes.
- **Learn lessons in terms of usability evidence.** The evidence must present the value of usability methods and design principles, and the conditions of validity of the results. Since one learns better from one's mistake, the evidence of the negative impact of violating usability design principles (i.e. usability flaws) or not applying properly usability evaluation methods must be presented too.
- **Disseminate evidence-based usability knowledge.** The evidence should be disseminated during the Health Informatics curriculum or through training of designers. Moreover, a database should be developed that would contain the formulation of the evidence along with the data supporting and contradicting it.⁸

4.3. Challenges to overcome

The road towards evidence-based usability is paved with challenges to be faced:

- **Uneven quality of evaluation.** Despite good practices guidelines [33]⁹, manufacturers favour quick and dirty methods (e.g. questionnaire targeting perceived usability) over validated methods (e.g. usability test). Providing evidence on the value of validated usability methods will promote their use.
- **Poor reporting quality.** Overall, usability studies on health IT are poorly reported [34] (e.g. not all usability issues are reported). Existing reporting guidelines [35]¹⁰ do not completely fit the specificities of usability evaluations (e.g. no mention of

⁶ See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁷ See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: *ibid*

⁸ See also: A. Georgiou, Finding, appraising and interpreting the evidence, in: *ibid*.

⁹ See also: P. Nykänen et al., Quality of health IT evaluations, in: *ibid*.

¹⁰ See also: E. Ammenwerth et al., Publishing health IT evaluation studies, *ibid*.

the iterative process). Guidelines dedicated to Human Factors and usability [36] should be used. Similarly, incident reports lack details and are therefore difficult to interpret. Incident reporting forms should be structured so that usability characteristics of the technology incriminated are described precisely.

- **Lack of taxonomy.** Health IT lacks a recognized taxonomy. Consequently, labels of the technology evaluated may be subject to discussion (e.g. what does "medication-related CDSS" refer to: an alerting system, order sets, clinical reminders?). Therefore the scope of the evidence related to that technology may be confused.
- **Difficulties to identify usability studies.** "Usability" and "Human Factors" are not MeSH terms. This issue may bias the identification of usability studies. Moreover, usability evaluations are often part of larger studies that mention seldom "Usability" in the title, the keywords, or the abstract. Authors should be encouraged to explicitly identify usability activities in their paper.
- **Distinguish the origin of usability issues.** Usability issues may originate in features of the technology but also in the local setting of this technology. Telling this difference may be a difficult but is a crucial task in order not to attribute a usability issue to a feature of a product when it comes from its parameterization. Therefore, reports should highlight as far as possible the origin of the usability issues.
- **Difficulties to access manufacturers' databases.** Manufacturers do not share users' feedbacks and results of homemade usability evaluations with Human Factors researchers. This policy prevents researchers from accessing and analyzing large valuable repositories. A win-win cooperation mode should be defined to encourage manufacturers to share those data with the Human Factors community.

4.4. Examples for available evidence

This section describes the few available examples for evidence both on design elements and on the usability evaluation methods.

4.4.1. Evidence on design elements

Several reviews aimed at identifying the positive and negative usability characteristics of a given health IT. Those reviews focused on CPOE [37], Electronic Medical Records [38], medication-related alerting systems [5] and M-health applications [39]. Those reviews are not equally useful. Only the first three ones matched the usability flaws they identified with usability design principles. The review on M-health applications defined a list of usability characteristics generic to mobile applications, not specific to a type of application; moreover, its results mixed usability flaws, usage problems and design principles. Therefore, it is not possible to build directly evidence on the design elements for a specific type of mobile applications.

One example of a more structured review is [5] that identified the usability flaws of medication-related alerting systems and then complemented them (i) by an analysis of the consequences for the user and for the work system of those flaws [40] and (ii) by a matching with existing usability design principles [41]. Table 1 and Figure 2 present excerpts of the results from this review. Based on those results, a database could be provided to designers to make them aware (i) of the known usability mistakes and their

consequences to be prevented when designing a medication-related alerting system and (ii) of the existing usability design principles useful to prevent those mistakes.

Table 1. Excerpts from the database of usability issues related to medication alerting system (details in [41]).

Usability flaws	Usage problems	Negative outcomes	Related Usability Design Principles
#1 Compatibility, alert presentation issue: "CPOE provides feedback on drug allergies, but only after medications are ordered." [7]	Behavioural issue: "Some house staff ignored allergy notices (...) and, most important, post hoc timing of allergy information." [7]	Workflow issue: "House staff claimed post hoc alerts unintentionally encourage house staff to rely on pharmacists for drug allergy checks, implicitly shifting responsibility to pharmacists." [7]	Fit the clinicians' workflow.: Alert must be displayed at the appropriate time during the decision making. [42]
#2 Insufficient guidance: "Physician (MD) orders [VA] aspirin - 162 mg. An order check [alert] appears. Says duplicate drug order. Non-VA ASPIRIN. [Alert] mentions 325mg. MD is looking at it also and [appears] confused" [43]	Behavioural issue: "MD clicks through [the alert] [accepts order]" [accepts without understanding the alert] [43]	Patient safety issue: "MD goes back to the medication list. Aspirin is now listed both under VA list and non-VA medication list" [double order of aspirin] [43]	Provide "means to advise, orient, inform, instruct, and guide the users throughout their interactions with a computer, including from a lexical point of view." [18]

In summary, existing reviews may provide the basis for evidence for design elements of health IT but the work is still to be up-dated and completed. As for other potential sources of evidence, there is still no in-depth analysis of incident reports that identify the positive and negative usability characteristics of health IT.

4.4.2. Evidence on usability evaluation methods

Some publications systematically analyzed the usability methods used for health IT. Most of them draw a picture of the type of usability methods used to develop and evaluate health IT [34], according to the stage of the System Development Life Cycle and the type of technology evaluated [44], or for a specific type of technology ("technology-based diabetes intervention platform" [45]). One specific study showed interest in the advantages and problems of usability evaluation methods applied to health collaborative systems [46]. Finally, the impact of usability evaluation and subsequent redesign on the task-completion time has also been evaluated [47]: the results of this review pointed towards a trend in improved task efficiency after modifications based on the results from usability evaluation.

In summary, the evidence published on usability evaluation methods is still weak (mostly lists of methods used) and a long road still needs to be travelled to be able to know (i) amongst all the existing ways to instantiate a specific method, which one is the most efficient for a given technology at each specific step of the design process, and (ii) whether some methods (and combinations of methods) are best suitable for a given technology and for specific parts of the UCD than others.

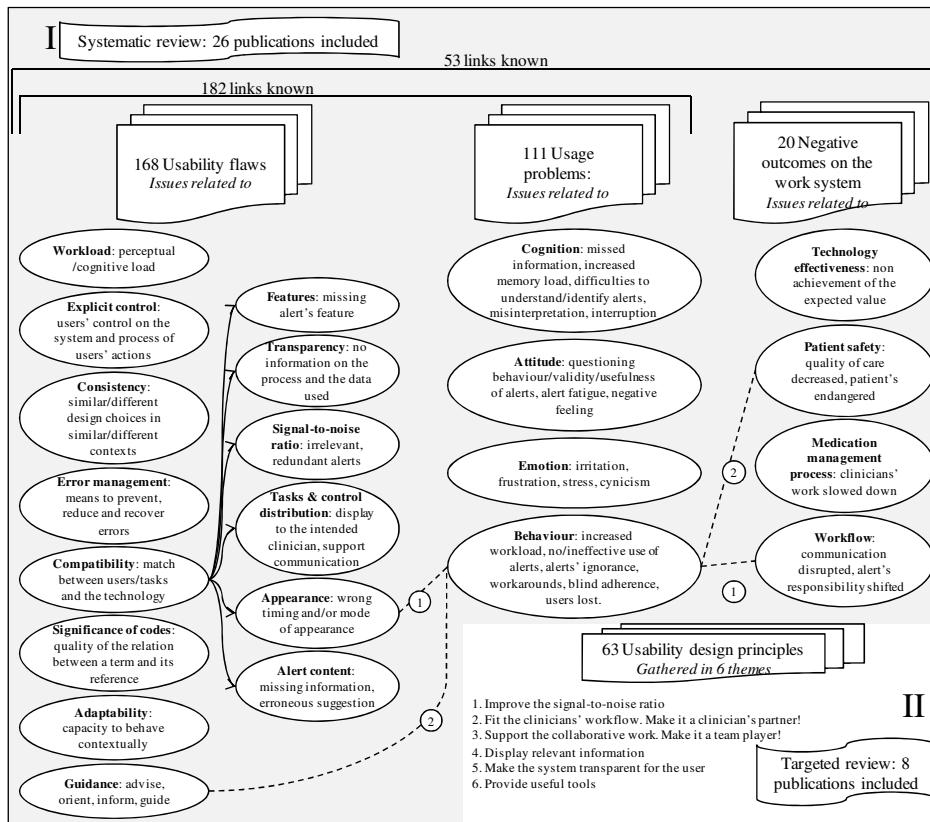


Figure 2. Representation (I) of the types of usability flaws reported for medication alerting systems and of their consequences for the user (usage problems) and the work system (negative outcomes); and (II) of the main themes of usability design principles known for that technology (based on [41]). Dotted lines synthesize two instances of cause-consequences chains reported in the literature between usability flaws in the alerting system, usage problems experienced by the users and negative outcomes in the work system (cf. Table 1).

5. Discussion

In this contribution we elucidate on usability as a critical factor of success and safe use of health IT. The UCD approach should be applied to ensure easy-to-use, efficient, satisfying, and non error-prone technology. Currently, stakeholders in the application of UCD do not apply UCD for it to be efficient for each type of technology and at each step of the design process. Therefore it is still possible to apply UCD erroneously and design technologies that can induce use-errors due to low usability. Guidelines based on empirical evidence are thus needed to help designers or evaluators avoid design flaws by choosing appropriate usability design principles and (combinations of) usability evaluation methods which usefulness and efficiency have been proven empirically.

Some attempts to get evidence-based usability knowledge exist. They proceed through systematic searching, critical appraisal and synthesis of the usability literature. Even if those attempts are limited, they are nonetheless valuable and provide the first

steps towards evidence-based usability practice. However, the road towards evidence-based usability is full of pitfalls. Measures must be adopted to help search for evidence.

Developing evidence-based usability knowledge is not an end in itself. It is necessary to make it available to designers and evaluators to ultimately improve health IT usability and to avoid usability-induced use errors, and thus to protect patients, users, and organisations. Thereupon, several questions must still be discussed: for instance, to whom precisely must the evidence be provided? Under which format? When in the project time-line? How generic or technology- or context-specific should the evidence be? The challenges to get evidence and the questions to discuss cannot be overcome and answered by individuals. Achieving and spreading evidence require the active involvement of the whole Human Factors and usability community in Health Informatics along with the support of manufacturers.

Recommended further readings

1. B. Séroussi, M.-C. Jaulent, C.U. Lehmann, editors, *Yearbook of medical informatics 2013: Evidence-based health informatics*, Stuttgart, Schattauer, 2013.
2. G. Salvendy, editor, *Handbook of Human Factors and Ergonomics*, 4th ed., John Wiley and Sons, Hoboken, New Jersey, 2012.
3. P. Carayon, editor, *Handbook of Human Factors and Ergonomics in Health Care and Patient Safety*, 2d ed., CRC Press, Boca Raton, Florida, 2012.
4. M. B. Weinger, M. E. Wiklund, D. J. Gardner-Bonneau, editors, *Handbook of Human Factors in Medical Device Design*, CRC Press, Boca Raton, Florida, 2010.

Food for thought

1. Should formative and summative evaluations results be considered equally when searching for evidence on Human Factors and usability?
2. How do in-lab and field studies differ in providing sight on usability knowledge?
3. What metrics would ensure that usability actually improves or reduces the safety and the beneficial effect of a health technology?
4. What policy and institutional processes should become normative requirements to ensure that systems are developed in user-friendly formats?

References

- [1] British Standards Institution (BSI), *The growing role of human factors and usability engineering for medical devices*, white paper, <http://www.bsigroup.com/en-GB/medical-devices/resources/whitepapers>, last access 11 February 2016.
- [2] Human Factors and Ergonomics Society, *Definitions of Human Factors and Ergonomics*, <http://www.hfes.org/Web/EducationalResources/HFEdefinitionsmain.html>, last access 11 February 2016.
- [3] P. Carayon, H.A. Schoofs, B.T. Karsh, A.P. Gurses, C.J. Alvarado, M. Smith, et al., Work system design for patient safety: the SEIPS model, *Qual Saf Health Care* **15** Suppl 1 (2006), i50-i58.
- [4] International Standardization Organization, *Ergonomics of human system interaction - Part 210: Human centered design for interactive systems* (Rep N°9241-210), Geneva, International Standardization Organization, 2010.

- [5] R. Marcilly, E. Ammenwerth, F. Vasseur, E. Roehrer, M.C. Beuscart-Zephir, Usability flaws of medication-related alerting functions: A systematic qualitative review, *J Biomed Inform* **55** (2015), 260-71.
- [6] Institute of Medicine (IOM), *Health IT and patient safety: building safer systems for better care*, The National Academies Press, Washington DC, 2012.
- [7] R. Koppel, J.P. Metlay, A. Cohen, B. Abaluck, A.R. Localio, S.E. Kimmel, et al., Role of computerized physician order entry systems in facilitating medication errors, *JAMA* **293**(10) (2005), 1197-203.
- [8] D.F. Sittig, A socio-technical model of health information technology-related e-iatrogenesis, *AMIA Annu Symp Proc* 2008, pp. 1209-10.
- [9] D. Ash, Lessons from Epinal, *Clin Oncol (R Coll Radiol)* **19**(8) (2007), 614-5.
- [10] M. Grissinger, Pen Injector Technology Is Not without 'Impending' Risks, *Pharmacy and Therapeutics* **35** (2010), 245-66.
- [11] J.K. Fakler, Y. Robinson, C.E. Heyde, T. John, Errors in handling and manufacturing of orthopaedic implants: the tip of the iceberg of an unrecognized system problem? *Patient Saf Surg* **1**(1) (2007), 5.
- [12] European Parliament Council, *Council Directive 2007/47/EC*, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:247:0021:0055:en:PDF>, last access 11 February 2016.
- [13] International Electrotechnical Commission, *Medical devices - Application of usability engineering to medical devices (Rep N° 62366)*, Geneva, International Electrotechnical Commision, 2007.
- [14] HIMSS Usability Task Force, *Promoting Usability in Health Organizations: Initial Steps and Progress Toward a Healthcare Usability Maturity Model*, Healthcare Information and Management Systems Society, 2011.
- [15] L. Watbled, M.C. Beuscart-Zephir, S. Guerlinger, L. Douze, E. Lepage, S.J. Darmoni, et al., Work System Characteristics Impacting the Performance and Quality of the Discharge Letter Process, *Stud Health Technol Inform* **218**(2015), 145-50.
- [16] J. van der Peijl, J. Klein, C. Grass, A. Freudenthal, Design for risk control: the role of usability engineering in the management of use-related risks, *J Biomed Inform* **45**(4) (2012), 795-812.
- [17] K.J. Vicente, *Cognitive work analysis*, Lawrence Erlbaum Associates, Mahwah, NJ, 1999.
- [18] D.L. Scapin, J.M.C. Bastien, Ergonomic criteria for evaluating the ergonomic quality of interactive systems, *Behaviour and Information Technology* **6**(4-5) (1997), 220-31.
- [19] M. Zachariah, S. Phansalkar, H.M. Seidling, P.M. Neri, K.M. Cresswell, J. Duke, et al., Development and preliminary evidence for the validity of an instrument assessing implementation of human-factors principles in medication-related decision-support systems – I-MeDeSA, *J Am Med Inform Assoc* **18** Suppl 1 (2011), i62-i72.
- [20] A.L. Russ, A.J. Zillich, B.L. Melton, S.A. Russell, S. Chen, J.R. Spina, et al., Applying human factors principles to alert design increases efficiency and reduces prescribing errors in a scenario-based simulation, *J Am Med Inform Assoc* **21**(e2) (2014), e287-e296.
- [21] J. Horsky, D.R. Kaufman, V.L. Patel, Computer-based drug ordering: evaluation of interaction with a decision-support system, *Stud Health Technol Inform* **107**(Pt 2) (2004), 1063-7.
- [22] J.M. Bastien, Usability testing: a review of some methodological and technical aspects of the method, *Int J Med Inform* **79**(4) (2010), e18-e23.
- [23] L.W. Peute, N.F. de Keizer, M.W. Jaspers, The value of Retrospective and Concurrent Think Aloud in formative usability testing of a physician data query tool, *J Biomed Inform* **55** (2015), 1-10.
- [24] A. Eghdam, J. Forsman, M. Falkenhav, M. Lind, S. Koch, Combining usability testing with eye-tracking technology: evaluation of a visualization support for antibiotic use in intensive care, *Stud Health Technol Inform* **169** (2011), 945-9.
- [25] M. Wiklund, J. Kendler, A.Y. Strohlic, *Usability testing of medical devices*, CRC press, 2010.
- [26] M.W. Jaspers, A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence, *Int J Med Inform* **78**(5) (2009), 340-53.
- [27] M.F. Walji, E. Kalenderian, M. Piotrowski, D. Tran, K.K. Kookal, O. Tokede, et al., Are three methods better than one? A comparative assessment of usability evaluation methods in an HER, *Int J Med Inform* **83**(5) (2014), 361-7.
- [28] Y.Y. Han, J.A. Carcillo, S.T. Venkataraman, R.S. Clark, R.S. Watson, T.C. Nguyen, et al., Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system, *Pediatrics* **116**(6) (2005), 1506-12.
- [29] R. Marcilly, L.W. Peute, M.C. Beuscart-Zephir, M.W. Jaspers, Towards Evidence Based Usability in Health Informatics? *Stud Health Technol Inform* **218** (2015), 55-60.
- [30] E. Ammenwerth, Evidence-based Health Informatics: How Do We Know What We Know? *Methods Inf Med* **54**(4) (2015), 298-307.

- [31] U.S.Food and Drug Administation (FDA), *MAUDE - Manufacturer and User Facility Device*, <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>, last access 11 February 2016.
- [32] B. Rihoux, Qualitative Comparative Analysis (QCA) and Related Systematic Comparative Methods: Recent Advances and Remaining Challenges for Social Science Research, *International Sociology* **21**(2) (2006), 679-706.
- [33] P. Nykänen, J. Brender, J. Talmon, N. de Keizer, M. Rigby, M.C. Beuscart-Zephir, et al., Guideline for good evaluation practice in health informatics (GEP-HI), *Int J Med Inform* **80**(12) (2011), 815-27.
- [34] L.W. Peute, R. Spithoven, P.J. Bakker, M.W. Jaspers, Usability studies on interactive health information systems; where do we stand? *Stud Health Technol Inform* **136** (2008), 327-32.
- [35] J. Talmon, E. Ammenwerth, J. Brender, N. de Keizer, P. Nykanen, M. Rigby, STARE-HI -statement on reporting of evaluation studies in health informatics, *Yearb Med Inform* (2009), 23-31.
- [36] L.W. Peute, K.F. Driest, R. Marcilly, S. Bras Da Costa, M.C. Beuscart-Zephir, M.W. Jaspers, A Framework for reporting on Human Factor/Usability studies of Health Information Technologies, *Stud Health Technol Inform* **194** (2013), 54-60.
- [37] R. Khajouei, M.W. Jaspers, The impact of CPOE medication systems' design aspects on usability, workflow and medication orders: a systematic review, *Methods Inf Med* **49**(1) (2010), 3-19.
- [38] M. Zahaby, D.B. Kaber, M. Swangnetr, Usability and Safety in Electronic Medical Records Interface Design. A Review of Recent Literature and Guideline Formulation, *Human Factors* **57**(5) (2015), 805-834.
- [39] B.C. Zapata, J.L. Fernandez-Aleman, A. Idri, A. Toval, Empirical studies on usability of mHealth apps: a systematic literature review, *J Med Syst* **39**(2) (2015), 1.
- [40] R. Marcilly, E. Ammenwerth, E. Roehrer, S. Pelayo, F. Vasseur, M.C. Beuscart-Zephir, Usability Flaws in Medication Alerting Systems: Impact on Usage and Work System, *Yearb Med Inform* **10**(1) (2015), 55-67.
- [41] R. Marcilly, *Towards a usability knowledge base to support health information technology design and evaluation: application to medication alerting systems*, PhD thesis, Lille 2 University, 2014.
- [42] J. Horsky, G.D. Schiff, D. Johnston, L. Mercincavage, D. Bell, B. Middleton, Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions, *J Biomed Inform* **45**(6) (2012), 1202-16.
- [43] A.L. Russ, A.J. Zillich, M.S. McManus, B.N. Doebbeling, J.J. Saleem, Prescribers' interactions with medication alerts at the point of prescribing: A multi-method, in situ investigation of the human-computer interaction, *Int J Med Inform* **81**(4) (2012), 232-43.
- [44] P.Y. Yen, S. Bakken, Review of health information technology usability study methodologies, *J Am Med Inform Assoc* **19**(3) (2012), 413-22.
- [45] C.R. Lyles, U. Sarkar, C.Y. Osborn, Getting a technology-based diabetes intervention ready for prime time: a review of usability testing studies, *Curr Diab Rep* **14**(10) (2014), 534.
- [46] L. Pereira de Araujo, M.M. Mattos, *A systematic literature review of evaluation methods for health collaborative systems*, in: Proceedings of the IEEE 18th International Conference on Computer Supported Cooperative Work in Design, 2014. pp. 366-9.
- [47] J. Wills, K. Hurley, *Testing usability and measuring task-completion time in hospital-based health information systems: a systematic review*, in: Proceedings of the 28th IEEE Canadian conference on electrical and computer engineering (CCECE), 2012. pp. 1 - 4.

Participatory Design, User Involvement and Health IT Evaluation

Andre KUSHNIRUK^{a,b,1}, Christian NØHR^b

^a*School of Health Information Science, University of Victoria, Victoria, Canada*

^b*Department of Development and Planning, Aalborg University, Aalborg, Denmark*

Abstract. End user involvement and input into the design and evaluation of information systems has been recognized as being a critical success factor in the adoption of information systems. Nowhere is this need more critical than in the design of health information systems. Consistent with evidence from the general software engineering literature, the degree of user input into design of complex systems has been identified as one of the most important factors in the success or failure of complex information systems. The participatory approach goes beyond user-centered design and co-operative design approaches to include end users as more active participants in design ideas and decision making. Proponents of participatory approaches argue for greater end user participation in both design and evaluative processes. Evidence regarding the effectiveness of increased user involvement in design is explored in this contribution in the context of health IT. The contribution will discuss several approaches to including users in design and evaluation. Challenges in IT evaluation during participatory design will be described and explored along with several case studies.

Keywords. User-computer interface, software design, electronic health records, evidence-based practice.

1. Introduction

Work in health care has always been closely dependent on advanced levels of knowledge, and the way in which professionals work is not always apparent. Work may be interpreted differently and work descriptions do not reveal all aspects of work practices [1]. The late Professor Branko Cesnik of Monash University often used the slide in Figure 1 to express that the knowledge applied in health care activities arises from interaction rather than evidence. Instead of performing a literature search in e.g. Medline prior to making a decision to act, it is more common to discuss the issue with a co-worker, or ask a more senior colleague.

When designing or evaluating health IT systems it is essential to explicate the underlying knowledge that is determining the health care professionals' decisions to act. For this reason it is important to give the end users a prominent position particularly in design projects. However there are a multitude of methods to involve end users in design processes. In Figure 2 three of the dominant schools for involving users in IT sys-

¹ Corresponding author: Dr. Andre Kushniruk, School of Health Information Science, University of Victoria, Human and Social Development Building, 3800 Finnerty Road, Victoria, British Columbia, Canada, V8P 5C2, andrek@uvic.ca.

tem design are depicted. The three schools vary in the extent to which the user is involved in decision making about design.

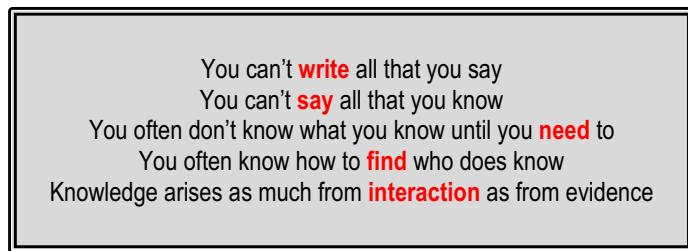


Figure 1. In performing health care work knowledge arises from interaction
 (from Professor Branko Cesnik, Monash University).

The *user-centered design* approach became widely used after Donald Norman and Stephen Draper in 1986 published their book: “User-Centred System Design: New Perspectives on Human-Computer Interaction” [2]. Two years later Norman published his seminal book “The Psychology Of Everyday Things (POET)” [3], which later was revised to “The Design of Everyday Things” [4], where he urged designers to study people, to take their needs and interests into account. The user-centered approach is also inherent in traditional usability testing and evaluation.² The methodological challenges for the user-centered design process are how to understand users’ need and design for these needs. The user-centered approach acknowledges the importance of user input into design and the characteristics of user-centered design include: (a) an early focus on observing and understanding users and tasks in design, (b) empirical evaluation and measurement of user interactions, and (c) iterative design processes (involving cycles of design, evaluation and re-design) [5]. This may involve the designer/developer observing and making notes about user (e.g. health professional or eHealth consumer) preferences, interactions and needs (as depicted in Figure 2a) using a variety of methods ranging from usability testing to observational methods such as time-motion studies [6].

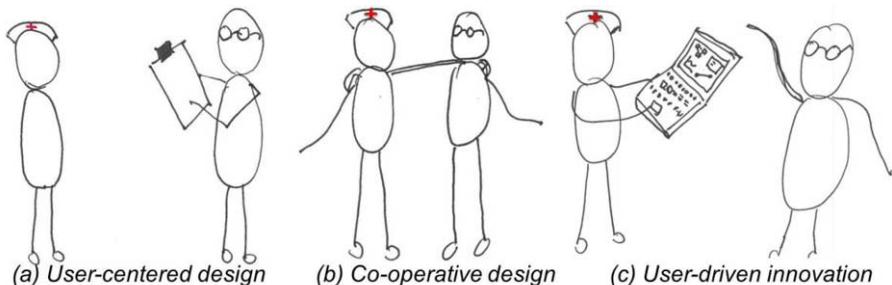


Figure 2. Three different schools of user involvement in IT design.

² See also: R. Marcilly et al., From usability engineering to evidence-based usability in health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics 222, Stud Health Technol Inform, IOS Press, Amsterdam, 2016.

The *co-operative design approach* emerged as recognition that as work activities become complex, human activities involve coordination and co-operation among many individuals with different areas of expertise. When the number of people involved in a work process exceeds a few, the complexity of coordinating increases several times. CSCW (Computer Supported Cooperative Work) is a central research field to address how collaborative activities and their coordination can be supported by means of computer systems [7]. The main challenge for the co-operative design approach is how to co-operate with users in the design process. This is depicted in Figure 2b, where user and designer/developer work together to come up with designs and modifications to design. This may involve creating a “design process where both users and designers are participating actively and creatively, drawing on their different qualifications” [8]. Other aspects of this approach include creation of prototypes that can be shown to users and used to simulate future work situations or studied in real settings (i.e. “*in-situ*” simulations) or in real life. Use of prototyping and simulations³ allow the users to explore and experience future work situations involving technology. This in turn supports discussion among the users and the designer/developers through a co-operative process. As an example, work by Jensen and colleagues, in development of a laboratory where clinical simulations are conducted with end users, has allowed for design and redesign of a wide range of clinical information systems [9].

User driven innovation is an example of participatory design where the basic idea is to engage the users to innovate and develop products themselves. Here the user becomes the central player on the design team and as shown in Figure 2c where key aspects of design decision making emanate from the user(s) themselves and the role of the developer becomes that of supporting and facilitating this creative user process. Participatory design involves direct involvement of users in the design of technologies [10]. Thus there is an emphasis on direct input of users in the design process and users are actively involved in decision making about design. It is not uncommon that users are the real source of innovations in an array of areas. Von Hippel studied this systematically and recommended that user driven innovation projects focus on “lead users” as the primary source of innovation [11]. A specific method to work systematically with the innovation process is described by Kanstrup and Bertelsen in their handbook on user innovation management [12]. This method involves application of a set of user innovation management (UIM) techniques to facilitate user innovation including step-wise approaches to understanding users and their contexts to generate design concepts from. As will be discussed, a major challenge for the user driven innovation approach is how to create space for user innovation, collect and sort out user-innovations, and transform these into new products.

2. Rationale for User Participation

There are two main rationales for participation of end users in design and evaluation of health IT. A pragmatic rationale is to increase IT system functionalities and service quality. This rationale stresses the need for users and developers to learn together through continuous mutual learning processes. The designers are responsible for point-

³ See also: S. Jensen, Clinical simulation as an evaluation method in health informatics, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

ing out technological options, and the users are the source of knowledge about their practices and the use situation [10].

A second rationale is political and reflects a commitment to give voice to those whose future we are to design. The basic premise is to *empower*, *emancipate* and *enhance* the health professional, the patient and/or the citizens in relation to application of health IT systems [13]. Participation needs to happen, because those who are to be affected by the changes resulting from designing and implementing health IT systems should, as a basic human right, have the opportunity to influence the design and implementation processes [10].

The two rationales behind participation in design or evaluation of health IT cannot be parameterized to obtain evidence of their efficiency or effectiveness – from this perspective it is basically a matter of commitment and taking a stand. Healthcare has often been slow to empower users as equals and this has been a strong rationale for increased user participation in design of new healthcare systems and applications.

As described above, the degree of end user participation in design can be seen as being on a continuum from considering the user as a “subject of study” in user-centered design, to users playing a more collaborative role in co-operative design, to the move to users driving the design process itself in true participatory design. In the following section we will provide examples of design projects within health informatics that vary along this continuum of user involvement in design. This will be followed by a discussion of the evidence in the literature about the benefits of user involvement and also the challenges and issues that arise as the degree of user input and involvement in design is increased.

3. An Example of a User-centered Design and Evaluation Project

A key component of user-centered design is continual and iterative input from end users through the evaluation of user interactions with developing prototypes and system designs. Early work in this area in health informatics came from Kushniruk and colleagues who applied and extended usability engineering methods to the design and refinement of healthcare information systems such as electronic health records, decision support systems and patient clinical information systems [14].

The first work in this area involved classic usability testing methods whereby representative end users (e.g. physicians and nurses) were observed as they were asked to carry out representative tasks (e.g. entry and retrieval of information about medications) using early system designs and prototype information systems. This work involved video recording users as they interacted with the systems under study while verbalizing their thoughts (i.e. “thinking aloud”). Thus the approach involved the designer/developers observing end users, noting their problems and issues through analysis of their observations, and refining system designs based on their analyses of end user interactions.

In a series of studies examining design of an electronic health record system (EHR) for use in clinical contexts, 16 physicians were asked to interact with a prototype version of the system and to think aloud while using it to carry out representative tasks (e.g. entering and retrieving patient data) using the system [14]. The screens were recorded as digital videos (using freely available screen recording software) and audio recordings of their thinking aloud were fully transcribed. In addition, physical actions can be recorded using an external camera (see Figure 3 for an example showing a

health professional being recorded as she works with a computer system). Using a video coding scheme (described in [15]) the user interactions were analyzed at a fine-grained level to identify usability problems and potential inefficiencies and flaws with the design of the system. This resulted in identification of a range of specific usability issues, including user navigation problems, difficulty qualifying medical findings in the system, and difficulty in representing temporal sequences. The results were summarized and presented to the design team, resulting in a modified user interface, which was in turn tested again to ensure that the issues identified were resolved.

With this user-centered approach to design, users were involved in the process early on and their interactions with evolving prototypes and early system designs were recorded and analyzed. However, their direct input into design decision making was limited, with some direct user suggestions being incorporated into redesign, but the majority of “fixes” coming from results of empirical analysis of user interactions by the design and evaluation team. The approach was shown to be effective, and many subsequent usability studies following this iterative user-centered approach have shown substantial reduction of user problems from one iteration to the next in the design and implementation of systems such as EHRs in a range of clinical settings, with one evaluation project showing a ten-fold decrease in coded usability problems during one iteration [15]. The user-centered approach has also been effectively applied to the analysis of systems designed for use by patients and lay people [31]. In addition, the approaches to conducting such user-centered evaluation to feedback input into iterative cycles of redesign have been modified and packaged to become low-cost and rapid in their application [16], which is leading to increased dissemination of user-centered design methods in healthcare (see Figure 3).

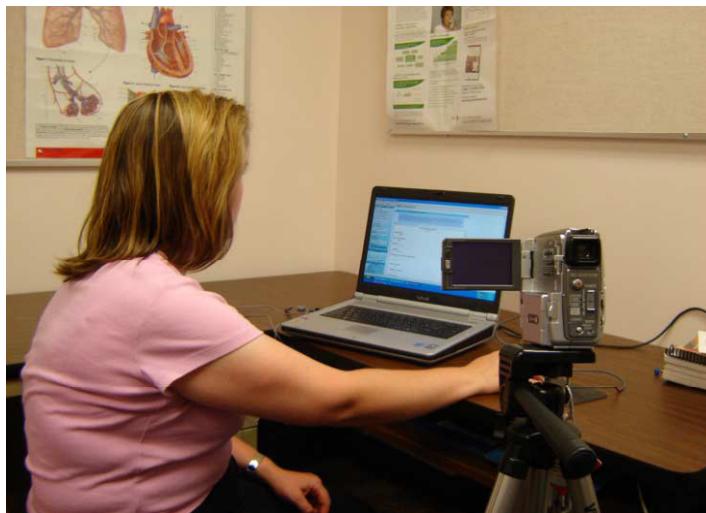


Figure 3. Example of low-cost rapid usability engineering set up for video recording health professionals as they work.

4. An Example of a Co-operative Design and Evaluation Project

As noted above, healthcare IT projects are recognized for being complex, typically involving multiple users and highly variable contexts of use. To address these issues we need to ensure systems are not only free from usability problems but that they serve to support and enhance complex healthcare workflow and practices [15]. To address these issues evaluative projects have emerged that may include multiple levels of analysis to consider not only surface level usability problems, but also the impact of systems and technology on workflow, inter-professional collaboration, healthcare outcomes and patient safety. This has necessitated the application and development of new approaches in healthcare IT for supporting co-operative design. Along these lines low-cost methods for conducting evaluations involving multiple users in simulated as well as real life healthcare settings and contexts (i.e. “in-situ” methods) have emerged [16][18]. These methods extend the usability testing methods employed in user-centered design to include recording of users in real work settings and collaborative environments. Such work has been aimed at better understanding the complex interaction and interplay among multiple users (e.g. physicians, nurses, patients) in multiple contexts (e.g. hospital care, home care) of use. To carry out this type of design one approach has been the development of simulation laboratories [9], while other researchers have moved the study to the actual location(s) where the technology will be used (e.g. in a particular hospital or home setting).

In an extension of the work described above for user-centered design of an EHR, it was discovered that increased and new types of user input would be needed to determine how to effectively modify and extend the design of the EHR for use in real clinical practice (i.e. during use with patients present in the room during clinical consultations). Along these lines, the application of the “simulated patient” approach (used in medical education to assess resident-patient interactions) was extended to be including in clinical simulations that involved physicians interacting with prototype EHRs while interacting with actors playing the role of patients. This involved video recording not only the computer screens but also the full doctor-patient interaction (e.g. dialogue between the doctor and the patient). The earliest work along these lines in healthcare IT was able to detect how an EHR system affected doctor-patient interaction and clinician decision making through video analysis of the interaction during several clinical scenarios [17]. The results were used to modify the user interface of the EHR to include features that users desired (such as easier navigation through the system using a navigation map feature).

It should be noted that during design sessions, the users who had interacted with the EHR system during the simulations were also included in design discussions to provide direct input and feedback into modifications of the system. Thus the roles of users in the project included interacting with the prototype and system being developed, as well as directly interacting with the design team during debriefs and design meetings to provide their continual input in a co-operative design process. This hybrid approach to user involvement (i.e. involving both observation of users, and also their direct input and feedback into design decision making) has since proven useful in a wide range of projects [18], including in the design and evaluation of a medication information system, and in the study of personal health records where users worked closely with designers in verbalizing and documenting their information needs during post-task interviews and cued-recall sessions after interacting with a health information system [19].

5. An Example of User Driven Innovation

User driven innovation is an excellent example of participatory design. In user driven innovation, the key issue is to create the space for *users* to be able to innovate and to transform these innovation ideas into new products or usable systems. The innovations should be grounded in user's needs, values, and knowledge.

Kanstrup and Bertelsen have outlined three central themes for organizing and conducting user-driven innovation and presented a set of techniques to support user-driven innovation processes [12]. First a co-operation between the users and the designers must be established – participating users must be carefully selected and a plan made for the innovation process. Second the context of the innovation has to be explored by gaining insight into current problems and needs, and also generating visions for future solutions. Thirdly ideas for the possible futures should be sketched and presented to decision makers. The second phase can be particularly challenging in the health care domain as health care institutions and facilities constantly are short of resources and taking clinicians away for design activities always means taking their time away from patient care. However, design games can be an activity that makes participating in innovation projects achievable as it has a high output using little clinician time for participating.

In the early 1990's design games were introduced to provoke development of a shared understanding among users and designers [20] and to form dialogue that supported mutual learning of the current practice and generate new design ideas [21]. Kanstrup and Christensen point out that the opening of the mind that gaming generates can be explained by Bateson's reflection on fun and seriousness, fantasy and games [22-23]. When playing games we are moving down unknown paths and thereby discover new aspects and generate new ideas. In games you can challenge the rules of current work practice by adding randomness to achieve a certain degree of "muddle". In Bateson's metalogues⁴ he makes the point: "If we didn't get into muddles, our talks would be like playing rummy without first shuffling the cards" [23, p. 26]

In the European project PSIP (Patient Safety through Intelligent Procedures in medication), which aimed to identify and prevent adverse drug events (ADE), a participatory design approach using design games was employed. The main objective of the project was to develop innovative knowledge based on data mining results and to deliver to professionals and patients a contextualized knowledge fitting the local risk parameters in the form of alerts and decision support functions. The design of these decision support functions was targeted by a design game approach. A PSIP design game was constructed to create space for clinician users to participate with their design ideas for clinical decision support functions [24]. The game was played by two teams (green team and blue team) each consisting of two nurses and two physicians (see Figure 4). They had the following items available:

- A box called "the PSIP machine" was made as a physical artifact that they could point to, hold, discuss the functionality of – a machine they could attribute any ability they wanted.
- A set of laminated scenarios for situations with medical errors for focusing and situating discussions and designs against medical errors.

⁴ A metalogue is a conversation about some problematic subject.

- Printed cards describing pre-made functionalities for clinical decision support plus blank cards for the group to describe newly invented functionalities.

The task for the competing teams was to build machines that could help to prevent the errors in the scenarios. There were no limits to the ability of the machines, they can do whatever the participants can imagine. The rules of the game were:

- Participants have two minutes to read the scenario.
- The team has 10 minutes to discuss the scenario and design a “PSIP machine” using the functionality cards or the blank cards.
- The teams have two minutes to present their machine to each other.
- After one hour the designed machines were evaluated and scored by the competing group.

The PSIP design game included three principles: Focus, produce, and prioritize. The scenarios presenting the problems to solve were derived from the database of reported ADEs, they made the participants focus on a very specific task. The competitive elements and the time limits stimulated the creativeness to produce results, and the evaluation and scoring prompted a prioritization among the solutions.

The game process, the evaluation and scoring were documented with video photos and notes by the facilitators. The teams designed eight different machines for clinical decision support preventing medical errors. The machines presented what the teams found most important in order to prevent medical errors. In the succeeding debriefing the design principles were derived from the central themes – see table 1.

This design game provided central knowledge and ideas for future design of clinical decision support systems based on negotiations among expert clinical users about complex practices. The game resulted in eight different machines derived through discussions and priorities for further design that the players pointed out. The two-hour workshop occupied only very little clinician time but produced a very important outcome for the further input into the design process.



Figure 4. The PSIP design game.

Table 1. Central themes and design principles from the played game [24].

Theme	Design Principle
Risks related to lack of integration of information	Design for integrated information: If physicians are to gather information from several different information systems, the risk that some information will be missing is high. Hand written information is dangerous and must be avoided.
	Integrated information must be used intelligently for clinical alerts, i.e. in cases of interaction.
Risks related to misreading and analysis of measurements and test results	Design graphical diagrams for visualizing measurements and lab results: Diagrams, e.g. a curve, will at a glance reveal if a measurement is out of normal range.
Risks related to rigid information system	Design for optimizing prescription: Information systems experienced as rigid and a disturbance (vs. a support) of the clinical work tend to lead to bad data discipline and workarounds.
Risks related to interruptions	Design for calm working environments when prescribing dispensing and administering medicine.
Risk related to misreading of medicine	Design for barcode readings or other types of scanning for verification.

6. The Benefits of User Participation

In the general IT literature lack of user input during design has been identified as being the single biggest contributing factor in the failure of complex IT systems to be adopted by users [25]. This finding has been found to hold in a number of different domains and is nowhere more salient than when considering complex healthcare IT projects, which have been associated with a high failure rate internationally [15]. Thus there is clear evidence that lack of user input is detrimental to the likelihood of system success and end user adoption.

Regarding the impact of varying degrees of user involvement in design, Kujala [6] as well as Damodaran [26] have collated results of studies from a number of different areas. Reported benefits of increased user involvement during design (particularly from participatory design) have included: (a) improved system quality as a result of better and more accurate user requirements gathering, (b) greater likelihood of inclusion of features users actually want, while avoiding addition during design of costly features users did not want, (c) higher levels of user acceptance of the resultant system developed with greater user input, (d) improved understanding of the resultant system by end users leading to lessened training needs and fewer usage issues, and (e) a higher level of participation in decision making by users in the organization to which they belong.

7. Selecting Tools and Techniques for User Participation

Regarding the issue of selection of tools and techniques for supporting greater user involvement in design of healthcare IT, a growing body of literature has documented an array of design and evaluation methods that can be employed [10]. Muller and colleagues [30] have characterized participatory design practices along several dimensions in creating a taxonomy of methods that can be used for selecting an appropriate ap-

proach for a particular design project. Along one dimension there is a range between designers participating in the user's world, which includes ethnographic observation⁵ and contextual inquiry and sessions envisioning future solutions (e.g. which could be held for example in an eHealth consumer's home or a health professionals' clinical environment). Other types of studies may be conducted in simulated environments (i.e. usability and simulation laboratories) using mock-ups, prototypes and theatres for design, where users may directly participate in design activities using computer supported tools, prototypes and methods.⁶

Also, the point in time along the system development life cycle where the product being developed is located is another important dimension, with methods such as use of design games, envisioning exercises, ethnographic methods and contextual inquiry being potentially very useful during the early phases of design. Methods particularly applicable later in the design process include co-operative prototyping, co-operative evaluation, participatory analysis of usability data and participatory customization of healthcare IT. Issues that cut across all phases of the life cycle include decisions regarding location of design and evaluation activities, the selection of user participants and the assessment of the appropriate user participant group size [10].

Evaluation methods used include many used in traditional system development. However, they differ fundamentally in the extent to which the boundary between designer and user is crossed and the degree, extent and nature of the user input into design. Therefore, evaluative methods such as focus groups, interviews, observation and methods adapted from usability engineering such as clinical simulations are applicable for evaluation during participatory design [10]. In addition, participatory and interpretive evaluative methods may also be applied, borrowing from advances in areas such as contextual inquiry, participative ethnographic methods, and video ethnography.

8. Issues and Challenges

Despite the reported benefits of greater user involvement in the design process, the issues of *how* (i.e. what methods to use), *when* (i.e. when during the system development life cycle) and *where* (e.g. in laboratory settings, real-world settings etc.) to engage users in design have remained active research questions, with varying evidence about the optimal approaches to bringing users into the design process in healthcare IT development [27]. In particular, the issue of assessing how representative the users (selected or volunteered to be involved in design process) are of the projected user population becomes a complex question when designing large scale systems (i.e. systems such as public health systems, which may have hundreds of thousands of users and a great many categories of different user types).

This has led Pilemalam and Timpka to discuss a need for a third generation of participatory design in healthcare (with the first generation being focused on the ideology of collective system design, and the second generation of participatory design shifting

⁵ See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁶ See also: R. Marcilly et al., From usability engineering to evidence-based usability in health IT, in: ibid.

towards commercial and IT applications [28]). According to Pilemalam and Timpka, participatory design “has traditionally presumed a certain degree of homogeneity as regards the information system target group”. To address this issue they propose a hybrid approach for large scale health projects which may involve elements of both user-centered design (e.g. usability testing studies with a larger number of users) alongside traditional participatory design processes with smaller groups of users. In the literature other issues have appeared including the following: (a) the perception by some developers that there may not be enough time in the system development life cycle to involve users as extensively as they would like, (b) obtaining access to representative users (particularly in healthcare) may be difficult, (c) there is the potential to have too many “user voices” leading to difficulty in obtaining consensus, and (d) users may need to be educated about aspects of design in order to work more collaboratively with the design team [6].

The issue of how and when to consider differing “voices” within design and evaluation needs to be considered in the context of the systems development life cycle of health information systems. There are currently a number of open challenges and issues regarding increased user input into design processes. These include the following: (a) how to define and recruit users for participatory design and related approaches; (b) when and how users can most effectively be brought into design and evaluation processes throughout the system development life cycle; (c) the representativeness of the user and user groups involved, (d) what evaluative methods can best be applied when designing and developing systems using the participatory approach; and (e) how to translate user input into improved systems. Kensing and Blomberg also echo some of these issues and concerns and have identified three main areas of challenges: (a) the politics of design in terms of the degree of ability of users to influence and shape the design of systems they will end up using, (b) the nature of user participation, and (c) the selection of the right tools and techniques for effective user participation [29].

9. Discussion and Conclusion

Lack of user satisfaction with healthcare information systems has been a serious issue in the area of health informatics. Indeed, consistent with the general literature on system design and adoption, lack of user input into design has been shown to be one of the most significant factors associated with failure of systems to be adopted by end users [25]. Evidence relevant to designing more effective systems involving greater user input has shown that approaches such as user-centered and participatory design can improve the effectiveness and adoption of a wide range of information systems [6].

In this contribution we have explored a number of different approaches to increasing user input into system design for the improved design and evaluation of healthcare information systems. Although the approaches vary, the common and clear thread is the need for increased user involvement in design. A number of challenges and issues exist for effectively increasing user input into design. In particular further work is needed to determine what aspects of design are best enhanced through increased user involvement as well as what methods are most appropriate for facilitating increased user involvement through the different phases of the system development life cycle.

The challenges of incorporating effective evaluation into participatory design are varied, including issues of how to incorporate user needs, and how to incorporate evaluation of designs using both low and high fidelity prototypes in mobile, home or natu-

realistic environments. Innovative approaches such as use of design games and other methods involving collection of user artifacts, photographs and video clips have also been employed in a number of projects in healthcare in Denmark, Canada and internationally and are promising. The projects described in this contribution are examples of projects where end user input was significant for obtaining a successful outcome. Future work should include an emphasis on understanding both the benefits and challenges of increasing user involvement so that users can be most effectively incorporated in the design of healthcare software.

Recommended further readings

1. J. Simonsen, T. Robertson, *Routledge International Handbook of Participatory Design*, Routledge, New York, 2013.
2. L. Botin, P. Bertelsen, C. Nøhr (Eds.), *Techno-Anthropology in Health Informatics. Methodologies for Improving Human-Technology Relations*, Studies in Health Technology and Informatics, Vol. 215, IOS Press, Amsterdam, The Netherlands, 2015.
3. A.M. Kanstrup, P. Bertelsen, *User Innovation Management, a handbook*, Aalborg University Press, Aalborg, Denmark, 2011.

Food for thought

1. What are the key points in a system's conception and completion that user participation is important?
2. What methods do you think are most effective for involving users in design?
3. How can knowledge about the importance of increased user input be translated into practices in the healthcare software industry?
4. What group(s) of strategically important users are likely to be overlooked in involving users in design participation?

References

- [1] L. Suchman, Representations of work: Making work visible, *Communications of the ACM* **38** (1995), 56.
- [2] D.A. Norman, S.W. Draper, *User Centered System Design; New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum and Associates, New York, 1986.
- [3] D.A. Norman, *The Psychology of Everyday Things*, Basic Books, 1988.
- [4] D.A. Norman, *The Design of Everyday Things*, MIT Press, Cambridge Mass, 2002.
- [5] J.D. Gould, C. Lewis, Designing for usability: Key principles and what designers think, *Communications of the ACM* **28**(3) (1985), 300-311.
- [6] S. Kujala, User involvement: A review of the benefits and challenges, *Behaviour & Information Technology* **22**(1) (2003), 1-16.
- [7] P.H. Carstensen, K. Schmidt, Computer Supported Cooperative Work: New Challenges to Systems Design, in Kenji I, editor, *Handbook of Hum. Factors/Ergonomics*, Asakura Publishing, Tokyo, 2003.

- [8] J. Greenbaum, M. Kyng, *Design by doing*, in: Design at Work: Cooperative Design of Computer Systems, J. Greenbaum, M. Kyng, eds., Lawrence Erlbaum Associates, Hillsdale, NJ, 1991. pp. 269-279.
- [9] S. Jensen, A. Kushniruk, C. Nøhr, Clinical simulation: A method for development and evaluation of clinical information systems, *Journal of Biomedical Informatics* **54** (2015), 65-76.
- [10] J. Simonsen, T. Robertson, *Routledge International Handbook of Participatory Design*, Routledge, New York, 2013.
- [11] E. Von Hippel, *The Sources of Innovation*, Oxford University Press, Oxford, 1988, p. 1-3.
- [12] A.M. Kanstrup, P. Bertelsen, *User Innovation Management, a Handbook*, Aalborg University Press, Aalborg, 2011.
- [13] L. Botin, The Technological Construction of the Self: Techno-Anthropological Readings and Reflections, *Techne Res Philos Technol* **19**(2) (2015), 211-232..
- [14] A.W. Kushniruk, D.R. Kaufman, V.L. Patel, Y. Levesque, P. Lottin, Assessment of a computerized patient record system: A cognitive approach to evaluating an emerging medical technology, *M.D. Computing* **13**(5) (1996), 406-415.
- [15] A.W. Kushniruk, V.L. Patel, Cognitive and usability engineering methods for the evaluation of clinical information systems, *Journal of Biomedical Informatics* **37**(1) (2004), 56-76.
- [16] T.B. Baylis, A.W. Kushniruk, E.M. Borycki, Low-cost rapid usability testing for health information systems: Is it worth the effort? *Studies in Health Technology and Informatics* **180** (2012), 363-367.
- [17] V.L. Patel, A.W. Kushniruk, S. Yang, J.F. Yale, Impact of a computer-based patient record system on data collection, knowledge organization and reasoning, *Journal of the American Medical Informatics Association* **7**(6) (2000), 569-585.
- [18] A. Li, A.W. Kushniruk, J. Kannry, D. Chrimes, T. McGinn, D. Edonyabo, D. Mann, Integrating usability testing and think-aloud protocol analysis with "near-live" clinical simulations in evaluating clinical decision support, *International Journal of Medical Informatics* **81** (2012), 761-722.
- [19] E.M. Borycki, L. Lemieux-Charles, L. Nagle, G. Eysenbach, Evaluating the impact of hybrid electronic-paper environments upon novice nurse information seeking, *Methods of Information in Medicine* **48**(2) (2009), 137-143.
- [20] P. Ehn, D. Sjøgren, *From Systems Description to Scripts for Action*, in: Greenbaum J, Kyng M, editors, Design at Work: Cooperative Desing of Computer Systems, Lawrence Earlbaum Associates, Hillsdale NJ, 1991.
- [21] M. Johansson, P. Linde, Playful collaborative exploration: New research practice in participatory design, *J Res Pract* **1**(1) (2005), M5.
- [22] A.M. Kanstrup, E. Christiansen, *Designing Games - balancing fun and seriousness*, in: Proceedings Danish HCI Resesearch Symposium DHRS 2005, Copenhagen, 2005.
- [23] G. Bateson, *Steps to an Ecology of Mind*, The University of Chicago press, Chicago, 2000.
- [24] A. M. Kanstrup, C. Nøhr, Gaming against medical errors: Methods and results from a design game on CPOE, *Studies in Health Technology and Informatics* **148** (2009), 188-96.
- [25] S. McConnell, *Rapid Development*, Microsoft Press, Redmond, WA, 1996.
- [26] L. Damodaran, User involvement in the system design process - a practical guide for users, *Behaviour & Information Technology* **15**(6) (1996), 363-377.
- [27] A. Kushniruk, P. Turner, Who users? Participation and empowerment in socio-technical approaches to health IT developments, *Studies in Health Technology and Informatics* **164** (2011), 280-5
- [28] S. Pilealm, T. Timpka, Third generation participatory design in health informatics - Making user participation applicable to large-scale information system projects, *Journal of Biomedical Informatics* **41**(2) (2008), 327-339.
- [29] F. Kensing, J. Blomberg, Participatory design: Issues and concerns, *Computer Supported Cooperative Work* **7** (1998), 167-185.
- [30] M.J. Müller, D.M. Wildman, E.A. White, Taxonomy of PD practices: A brief practitioner's guide, *Communications of the ACM* **36** (1993), 26-28.
- [31] J.J. Cimino, V.L. Patel, A.W. Kushniruk, The patient clinical information system (PatCIS): Technical solutions for and experiences with giving patients access to their electronic medical records, *International Journal of Medical Informatics* **68**(1-3) (2002), 113-127.

Clinical Simulation as an Evaluation Method in Health Informatics

Sanne JENSEN^{a,1}

^aThe Capital Region of Denmark, Copenhagen, Denmark

Abstract. Safe work processes and information systems are vital in health care. Methods for design of health IT focusing on patient safety are one of many initiatives trying to prevent adverse events. Possible patient safety hazards need to be investigated before health IT is integrated with local clinical work practice including other technology and organizational structure. Clinical simulation is ideal for proactive evaluation of new technology for clinical work practice. Clinical simulations involve real end-users as they simulate the use of technology in realistic environments performing realistic tasks. Clinical simulation study assesses effects on clinical workflow and enables identification and evaluation of patient safety hazards before implementation at a hospital. Clinical simulation also offers an opportunity to create a space in which healthcare professionals working in different locations or sectors can meet and exchange knowledge about work practices and requirement needs. This contribution will discuss benefits and challenges of using clinical simulation, and will describe how clinical simulation fits into classical usability studies, how patient safety may benefit by use of clinical simulation, and it will describe the different steps of how to conduct clinical simulation. Furthermore a case study is presented.

Keywords. Ergonomics, eHealth, qualitative evaluation, clinical simulation, risk, safety.

1. Introduction

Implementation of health IT in relation to improvement of patient safety and optimization of work flow is a paradox [1]. Even though health IT is intended and anticipated to have a positive impact on quality and efficiency of health care [2], the application of new technology in healthcare may also increase patient safety hazards [3, 4]. Studies show that adverse events are indeed often related to the use of technology [5-7].

Design of health IT focusing on protecting patient safety is one of many initiatives trying to prevent adverse events [8, 9].² Patient safety does not entirely rely on technology but is highly influenced by the interaction between users and technology in a specific context [10], and sociotechnical issues and human factors are related to many unintended consequences and patient safety hazards [7, 8, 11]. Possible patient safety hazards such as design of the IT system itself; embedding of IT system into local work

¹ Corresponding author: Sanne Jensen, The Capital Region of Denmark, Borgervanget 7, 2100 Copenhagen O, Denmark, sanne@regionh.dk.

² See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

practice in the local environment; and the introductory and training processes need to be investigated when health IT is integrated with local clinical work practice including technology and organizational structures. The substantial complexity of organizations, work practices and physical environments within the healthcare sector impacts design, evaluation and implementation of information systems [12, 13]. Healthcare environments are profoundly collaborative and rely on coordination between various health professionals [14]. They are characterized by delegated decision-making, multiple viewpoints and inconsistent and evolving knowledge bases [15]. Multiple groups with potentially divergent values and objectives work together and face many contingencies which cannot be fully anticipated [16, 17]. These matters challenge design and evaluation of health IT.

Clinical simulation tries to address this challenge. Compared with other methods, e.g. heuristic inspection and low fidelity usability evaluation³, clinical simulation takes the clinical context into account. In contrast, for example, heuristic inspection focuses on the user interface, and low fidelity usability testing focuses on technology and on the specific tasks of individual users. By including the clinical context, clinical simulation is ideal for proactive evaluation of new technology for clinical work practice [18, 19]. Clinical simulations involve real end-users as they simulate the use of technology in realistic environments performing realistic tasks [20]. Clinical simulation studies the effects on clinical workflow [21] and enables identification and evaluation of patient safety hazards before implementation at a hospital or other clinical setting [22]. Clinical simulation also offers an opportunity to create a space in which healthcare professionals working in different locations or sectors can meet and exchange knowledge about work practices and requirement needs [23, 24].

Hospital organizations and work practice are extremely complex with many different healthcare groups and many interactions and correlations [25] involved, and many acute situations are encountered during daily work practice in hospital settings [26]. This complexity affects the technology that is implemented at hospitals [27] and confronts the methodology used for design and evaluation of healthcare information systems. Failure to comprehend the nature and range of end-users has been highlighted as a key issue in many systems' failing to become accepted by healthcare professionals [28]. Furthermore, an understanding of the context in which the systems will be used must take into account not only tasks and settings [29], but also the range, competences and cognitive capacities of an increasing variety of potential end-users [30]. The risk of endangering patient security calls for careful evaluation before implementing new technology in real life settings [31].

Usability relates to how a product can be used to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.³ Usability focuses on use of technology in a specific context [32]. Context may be defined as "*users, tasks, equipment (hardware, software and materials), and the physical and social environment in which a product is used*" [32]. It does however raise several questions, e.g. 'who are the users?', 'what are their tasks?' and 'with whom, where and under what conditions are they performing these task?'. The healthcare sector poses challenges due to the larger potential numbers and classes of users, e.g. nurses, physicians and pharmacists [28]. Furthermore, the definition does not take multiple

³ See also: R. Marcilly et al., From usability engineering to evidence-based usability in health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

users and their professional interaction into account, and nor does it take parts of or a whole organization into account.

According to Hertzum [33] many views may be put on usability, dividing it into six images; 1) *universal usability*: usability in a system for everybody to use, 2) *situational usability*: quality-in-use of a system in a specified situation with its users, tasks, and wider context of use, 3) *perceived usability*: usability concerns the user's subjective experience of a system based on her or his interaction with it, 4) *hedonic usability*: usability is about joy of use rather than ease of use, task accomplishment, and freedom of discomfort, 5) *organizational usability*: usability implies groups of people collaborating in an organizational setting, and 6) *cultural usability*: usability takes on different meaning depending on the users different background. Hertzum claims that all images should be taken into account when evaluating usability.

Another aspect when designing and evaluating information systems is user involvement. User-centred design focuses on incorporating the user's perspective into the development process in order to attain a usable IT system [34].⁴ The key principles of user-centred design are 1) active involvement of users and clear understanding of user and task requirements; 2) an appropriate allocation of function between user and system; 3) iteration of design solutions; and 4) multi-disciplinary design teams. The human-centred design cycle [32] shown in Figure 1 describes five essential processes which should be undertaken in order to incorporate usability requirements into the software development process.

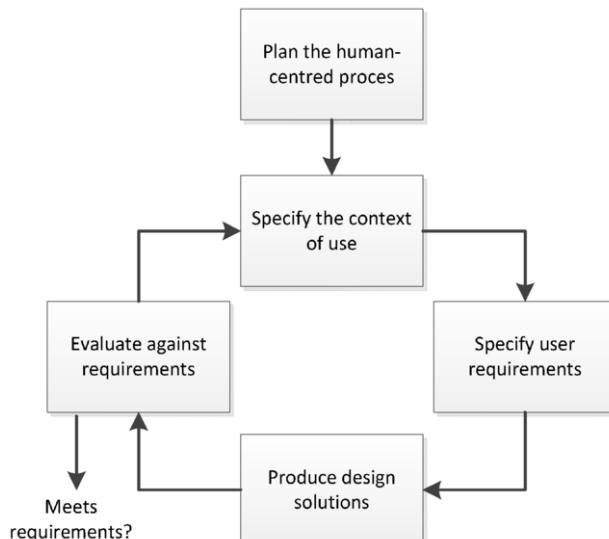


Figure 1: The human-centred design cycle.

The process is iterative with the cycle being repeated until the particular usability objectives have been obtained. Studies show that effective involvement of users may lead to 1) improved quality of the system arising from more accurate user requirements; 2) avoidance of costly system features that users do not want or cannot

⁴ See also: A. Kushniruk et al., Participatory design and health IT evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

use; 3) improved levels of acceptance of the system; 4) greater understanding of the system by the user resulting in more effective use; and 5) increased participation in decision-making in the organization [35, 36].

2. Clinical simulation

Clinical simulation supports involvement of context as well as end-users in pre-implementation design and evaluation of health IT. Clinical simulations involve real end-users as they simulate the use of technology in realistic environments performing realistic tasks [20]. As shown in Figure 2, clinical simulation can be used in different evaluation activities at various phases of the development life cycle from evaluation of work practice and user requirements, evaluation of the initial specification and early design solution so as to seek to eliminate patient risks created or perpetuated, through to application assessment in work practice and assessment of training programs.

Patient safety issues may be explored in all phases of the lifecycle by observing and analysing medical errors and work flow in a simulated situation close to a real life environment [22]. In the first phases of the lifecycle of health IT, simulation may be used for specification and evaluation of user requirements [19], as well as for obtaining knowledge and evaluate work practice [37]. This may involves observation of clinicians applying information technology under simulated conditions

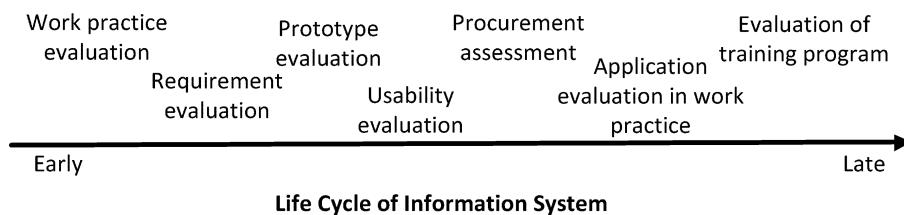


Figure 2. Simulation evaluations in information system life cycle.

Likewise in the design phase simulation is well suited as a method for user involvement in connection with evaluation of the design. Simulation studies can be designed to gain practical experience in evaluation of new technology without introducing any kind of ethical issues and without putting patients at risk [20]. In this way it is possible to test prototypical software in realistic scenarios and environments. Simulations can be performed in laboratories as well as in situ in a ward, an operating theatre or an outpatient clinic [20]. Simulation studies aim to evaluate design proposals for a new technology and combine elements of the laboratory test and the field study [22].

Particular aspects of implementation can be visualized by simulation e.g. user interaction in work practice, the need for training, and the impact of decision support [22]. Unintended consequences of new systems such as changes in work processes and patient outcome⁵ may also be detected and can provide constructive and valuable

⁵ See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

information for organizational decision makers [18]. Clinical simulation can also be used as common ground for discussion and negotiation and as an organizational learnings space, where knowledge of other parts of an organization can be acquired [37].

The realism and acceptance of the simulation depend on the degree of fidelity in the simulation set-up. Dahl and colleagues [38] have developed a simulation acceptance model with four fidelity dimensions: 1) *environment* – physical elements, such as rooms, beds and patient; 2) *equipment* – elements, such as mock-ups and electronic devices; 3) *functionality* – such as system functionalities and interactive devices; and 4) *tasks* – clinical task such as administration of drugs and ward rounds. These fidelity dimensions affect the perceived realism and thereby acceptance of the simulation made by the involved clinicians and should be considered carefully according to the purpose of the simulation.

Clinical simulations are performed in three phases; 1) introduction, 2) simulation, and 3) evaluation. Prior to the simulation, the participants are introduced to the information system and to the simulation. Simulation facilities are a dedicated facility with two rooms linked by a one-way mirror. During the simulation, a simulation facilitator is located in the simulation room. The facilitator assists the simulation and supports the participating clinician. An instructor located in the observation room instructs the patient and the simulation facilitator. A one-way mirror separates the two rooms. The simulation is observed by health informatics experts and sometimes by key stakeholders, such as colleagues from hospitals, clinical managers, quality managers and vendors [37]. The observers are located in the observation room.

An example of how simulation facilities may look like is presented in Figure 3. The simulation room is established as a bed room for two patients with bedside tables and a portable for the healthcare professional. An observation room with portables and chairs is located in the right corner. A one-way mirror is separating the two rooms.

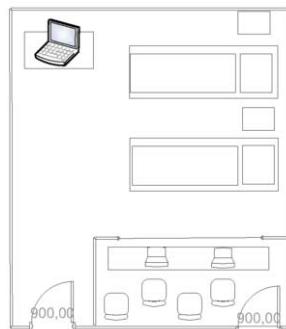


Figure 3. Overview of the physical simulation set-up.

Simulation of handover from hospital to community care by messaging technologies can also be carried out in a simulation laboratory. In such situations another simulation room may replicate a nursing office at the community care. In situations where it is not possible to replicate the location of the simulation in a laboratory, simulation *in situ* may be used. This could be scenarios where large x-ray scanners or other large equipment is involved.

The simulating clinicians are asked to “*think aloud*” so that the observers can acquire a deeper understanding of the human task-behaviour. Depending on the purpose of the clinical simulation, the clinicians are sometimes also able to observe their colleagues, when not participating in the simulation themselves [39]. The different roles and their locations are described in Table 1.

Table 1. Overview and description of different roles and their locations during clinical simulation.

Roles	Description	Location
Instructor	Overall responsible for the simulation. Instructs simulation facilitator and patient(s) during simulation by use of intercom equipment and facilitates debriefing.	Observation room
Simulation facilitator	Briefs clinicians prior to simulation and provides support during simulation. Receives instructions from and assists instructor during simulation.	Simulation room
Observers	Observes and makes notes during simulation; e.g. usability, support of work practice, patient safety	Observation room
Clinicians	Simulates scenario. Thinks aloud during simulation. Participates as interviewee in interview	Simulation room
Actor	Acts as e.g. patient, colleague during simulation and receives instructions from instructor.	Simulation room

After the simulation, the proposed information system is evaluated. Participants are asked to complete questionnaires and participate in a de-briefing interview. Additional to interview guides, observations made by the observers during the simulations are used as background for the interviews [24]. It must be clarified in advance to whom the results are to be presented and how the results and recommendations should be implemented. The same goes for the respective mandates of the participating clinicians as well as the observers.

3. Case Study: Simulation study of a clinical information system

The aim of the case study was to investigate how a newly-acquired standard clinical information system for doctors to sign for laboratory results might support clinical practice, and to identify potential patient safety hazards prior to its implementation [40]. The aim of the information system was to obtain an IT supported work flow for physicians receiving and signing laboratory test results in order to improve patient safety. In addition to implementation aspects such as training and information, the purpose was also to evaluate future work practice, the relation between technology and existing work processes, and the extent to which clinical simulation may be applied as a proactive method to identify and evaluate potential patient safety hazards prior to implementation.

The existing workflow was paper based; i.e. prints were made from digital systems and were signed by a doctor in order to document that the specific test result had been reviewed by a doctor. The laboratory tests were handled by various information systems. Some test results were on paper and others were electronic. The background for the local work flows was based on interpretations of a national guideline for handling laboratory test results. This national guideline was developed as part of a quality assurance initiative to increase patient safety. As a rule, the physicians signed to confirm that they have seen a laboratory test result. The physicians also signed to

confirm that they have handled the test results. The essential challenges about the paper based workflow were 1) lack of overview about whether a result has arrived; 2) uncertainty about whether a test result has been seen by a physician; and 3) lack of documentation about which physician has seen a test result.

The objective of purchasing the new information system was to increase quality in work practice and minimize the risk to patient safety by implementing a new standard information system. The information system collects laboratory test results and supports electronically documentation of acknowledging the results. The study was expected to be moderate and manageable because the information system was a standard off-the-shelf product and the intended work flow was supposed to be narrow and well-defined. The information system was to be implemented at two pilot departments. Both departments included patient wards and outpatient clinics. Prior to implementation, the existing work practice was analysed and future generic work flows defined. The functionality of the information system and collaborative future work practice were evaluated by means of clinical simulation. The aim of the simulation study was to assess how the information system supported clinical practice and to identify potential patient safety hazards prior to its implementation.

Initial field studies were carried out at the two pilot departments covering both patient wards and outpatient clinics in order to gain insight into existing work practice concerning receipt, handover and acknowledgement of laboratory test results. Two workshops were then held with physicians, nurses and medical secretaries from the pilot departments, health informaticians and experts from the regional quality unit. At the first workshop, future work practice and the information system were analysed and required changes were identified. At the second workshop, future work practice was determined, focusing on improved efficiency, quality, continuity and communication. Existing routines were contested and organizational changes were initiated ahead of implementation to create acceptance and a readiness to change among future end-users.

An analysis of work practice conducted prior to the clinical simulation revealed that there were significant differences between the hospitals, between the patient wards, and the outpatient clinics – and indeed also between the individual healthcare professionals. Furthermore, the design of future work practice presented a number of challenges and it was not possible to design a generic work flow to cover both patient ward and outpatient clinic. This was to some extent due to differences between local work flows but also due to the fact that the information system functionality did not provide adequate support for work practice.

Clinical simulation was conducted after the two workshops. The purpose of the clinical simulation was to evaluate patient safety issues and future work practice using the new information system before its implementation. Six healthcare professionals from the two pilot departments (two physicians, three nurses and one medical secretary) participated in the simulations. Clinical managers from the pilot sites, implementation experts and health informatics experts were observing the simulations. Figure 4 shows the simulation room seen from the observation room through a one way mirror. The simulation set-up is an outpatient clinic where a physician is preparing for a meeting with a patient.

A total of 11 scenarios were performed during the simulation; six scenarios from patient wards and five scenarios from outpatient clinics. All scenarios were related to signing and handling laboratory test results. Some of these were frequently performed work flows, e.g. ward rounds and visits to the outpatient clinic, while others were critical work flows; e.g. urgent test results, sorting test results and handover of

responsibility. The simulation set-up was very realistic. The computers used were identical with those used at the hospitals and the system was fully developed and operational. The scenarios were composed in participation with clinicians from the pilot sites and based on realistic patient cases. The simulation room was designed as either a ward bedroom or clinical office. The role of patient was enacted by a healthcare professional.

One of the purposes of using clinical simulation in relation to implementation was to investigate how the information system supported clinical practice and to determine whether the information system should be implemented at the hospitals. Therefore there was a need for high fidelity in the case study.



Figure 4. Simulation room seen from observation room.

The clinical simulation identified many uncertainties concerning work flow, handling of responsibility, and other organizational and technical challenges. High fidelity functionalities, such as integration to other information systems, revealed patient safety issues; e.g. notes related to a test result were not shown in relation to the test result in the new information system. The physician could only find the notes in the lab system. Apart from many negative findings, there were also positive findings, including improved overview of laboratory test results and no paper test results were left lying around, at the risk of disappearing.

We did not have any patient safety experts attending as observers during the simulation. Instead the simulation evaluation report was subsequently shown to the patient safety experts. Having patient safety experts observe the simulation would have improved the outcome considerably. Several organizational and technological issues, which were regarded as inconveniences by others, were detected as patient safety risks by the patient safety experts. These experts have great experience of what can go wrong and are able to focus on these matters during the simulation. They observe the interaction between the user and the interface of the technology but just as much the interaction with the technology in the clinical context. Inclusion of clinical context is one of the most powerful elements in clinical simulation. By allowing clinicians to use new technology in the way it is supposed to be used, patient safety issues become visible. Clinical simulation enables visualization of technology in connection with clinical context without endangering patients [22]. Therefore the choice of observers is very important. Each expert focuses on his or her own field. For this reason, observers must be chosen carefully and bearing in mind the purpose of the simulation.

As a result of the simulation additional new requirements of the information system were determined, e.g. new functionality for sorting the list of laboratory results according to date and time of the results. It was decided to initiate a pilot implementation despite the fact that the information system did not fully support the work flows. Some of the organizational challenges were solved and it was agreed that the remaining challenges regarding future work practice should be subject to scrutiny during the pilot implementation.

The challenges not solved prior to the pilot implementation were the transferability of work practice between patient wards and outpatient clinics, confidentiality of some test results, risk of several users handling the same test result simultaneously, missing interaction between prescription of test and signing of test results, no possibility of undoing signing of test results, comments do not stand out distinctly and integration between information system and paper-based test results from private laboratories. The issues were observed and evaluated after the system was implemented.

4. Discussion

The clinical simulation focused on formative evaluation and primarily was used as a learning process. Formative evaluation studies can facilitate system adoption and utilization [41] and aim to improve a system during its development or implementation, while summative evaluation focuses on evaluation of a system that is already up and running [42]. Formative evaluation may identify potential problems, such as patient safety issues, during the development phase and thus provide opportunities to improve a system as it develops.

In the simulation study, the results of the formative evaluation regarding patient safety issues and work practice for handling laboratory test results were presented and discussed at meetings with the various stakeholders, i.e. the patient safety unit, the quality unit and the implementation departments. Precautions were taken in relation to patient safety matters and work practice. Many of these precautions were subsequently implemented, regardless of the implementation of information system.

Unintended incidents often occur in the interaction between humans, technology and work practice [4, 10]. Clinical simulations allow visualization of the correlation between human, technology and organization. More conventional usability evaluations tend to visualize the interaction between the user and the technology but do not include work practice context [20, 43]. By including all three aspects (humans, technology and organization), patient safety challenges were revealed as well as organizational and technical challenges. New work practice in itself may also lead to unintended incidents. This was also revealed during the clinical simulation.

To expose cognitive and socio-technical issues, all fidelity dimensions described by Dahl and colleagues [23] need to be high on all four dimensions. The overall simulation configuration affects how the realism of the simulation experience is perceived [38]. Cognitive aspects of work practice relate to the clinical context and therefore depend on the degree of *environment* and *task* realism, whereas *equipment* fidelity and *functional* fidelity relate to cognitive aspects of the technical context. Socio-technical aspects and patient safety matters lie in the intersection between user, organization and technology [40]. High fidelity simulations are time-consuming [44] though and the purpose of simulation studies and the need for fidelity should therefore be planned carefully.

Traditional information systems are often designed around an idealized model of the tasks and workflow, and failures in information systems are often blamed on human, social and cultural “barriers” to technology adoption [10]. The case study revealed differences between such an idealized model of the task that needed to be accomplished and the way in which clinicians were actually working. Some of the differences were due to local interpretations of the regional guidelines and one of the conclusions reached was that the regional quality unit should develop a regional standard for signing off test results. Another issue lay in the fact that the information system was a standard system which did not provide adequate opportunities to configure the system to match the local setting. If work practice differs from department to department, local configuration is a requirement. A regional standard was introduced to resolve this issue.

Clinical simulation did not reveal all challenges related to the information system. The outcome of clinical simulations depends on the quality of the scenarios and patient cases they cover. In the case study, the scenarios during the simulations did not include unusual results or pre-ambulatory test results, but only became clear during a subsequently pilot implementation. Clinical simulation involves an inherent risk of giving an idealized picture compared to real life as it is very resource demanding to simulate the complexity of real life situations at a hospital. These matters are important to take into account when planning and designing the simulation.

Another aspect is the purpose of the evaluation and the relation between existing and future work practice. What is to be evaluated - future or existing work practice? And do the end-users comprehend and approve the new work practice? Furthermore, if the existing work practice in a department does not follow the existing guidelines, this may influence the simulation of the interaction between future work practice, end-users and technology as well as subsequent implementation.

Several muddled work flows became clear during the simulation and observers focusing on work flows agreed that a further work flow analysis was needed. This resulted in revision of the future work practice. Many of the issues found during the simulation were addressed before the pilot implementation, and those that were not solved were observed again during the pilot implementation. As such clinical simulation cannot replace a pilot implementation, but should rather be regarded as a valuable supplement.

Patient safety issues are difficult to assess due to the fact that many patient safety challenges lie in the details and are triggered by adverse events and disturbances [24]. The results of the case study showed that clinical simulation took the clinical context into account, while other methods, e.g. heuristic inspection, focus on the user interface. Low fidelity usability testing focuses on technology and specific tasks for single users. Some patient safety risks may therefore be difficult to pinpoint using these methods. Clinical simulation provides a comprehensive view on the information system taking into account the correlation between IT, work practice and adverse events, and is therefore a very suitable method for assessing patient safety issues.

The resources invested in preparing and performing simulation studies may be exhaustive, depending on the required degree of fidelity. It is essential that the resources invested in creating a realistic setting match the purposes of the simulation and the simulation set-up [43, 44]. However, the resources saved and iatrogenic effects avoided by using clinical simulation for analysis and evaluation purposes are difficult to quantify as it is difficult to put a price on the value of patients' lives. Still, clinical simulation is a beneficial evaluation method, as it takes place in a controlled environment where there is no risk of injuring real patients [20, 40].

As described clinical simulation can be used to analyse, design and evaluate user requirements and work practice and serve as common ground to help to achieve a shared understanding between various communities of practice. The primary benefits of using clinical simulation are 1) involvement of users and clinical context, 2) controlled environments for experiments and formative evaluations of user satisfaction, usefulness and patient safety, 3) environments for addressing and visualizing cross-sectorial and cross-functional topics, and 4) organizational learning space and common ground for gaining shared understanding.

The main concerns and challenges of using clinical simulation are that clinical simulation does not reflect the social-technical issues over time and does not cover all possible work practice situations and issues. The purpose and choice of scenarios determines to a great extent the outcome, and the purpose and design of clinical simulation must therefore be considered very carefully.

Recommended further readings

1. Kushniruk A, Nohr C, Jensen S, Borycki EM, From Usability Testing to Clinical Simulations: Bringing Context into the Design and Evaluation of Usable and Safe Health Information Technologies, *Yearb Med Inform* **8** (2013), 78-85.
2. Jensen S, Kushniruk A, Boundary objects in clinical simulation and design of eHealth, *Health Informatics Journal* Oct 9, 2014;1460458214551846.
3. Jensen S, Kushniruk AW, Nøhr C, Clinical simulation: A method for development and evaluation of clinical information systems, *Journal of Biomedical Informatics* **54** (2015), 65-76.
4. Ammenwerth E, Hackl WO, Binzer K, Christoffersen TE, Jensen S, Lawton K, Skjoet P, Nohr C, Simulation Studies for the evaluation of health information technologies: experiences and results, *HIM J* **41**(2012), 14-21.
5. Jensen S, Clinical Simulation: For what and how can it be used in design and evaluation of health IT, *Techno-Anthropology in Health Informatics: Methodologies for Improving Human-Technology Relations, Studies in Health Technology and Informatics* **215** (2015), pp. 217-28.

Food for thought

1. What might be the pros and cons of clinical simulation seen from an end-user perspective and how may it differ from a management and policy perspective?
2. Clinical simulation refers to simulation in a clinical set-up. How may simulation fit into other high-risk areas such as pharmacies and ambulances?
3. As healthcare technology moves into patients' homes, simulation could also be used in private settings. How would a simulation design differ when conducting simulations in patients home?
4. How may clinical simulation be used in other clinical fields, such as biomedical engineering?

References

- [1] E. Coiera, J. Aarts, C. Kulikowski, The dangerous decade, *J Am Med Inform Assoc* **19** (2012), 2-5.
- [2] D.W. Bates, M. Cohen, L.L. Leape, M.J. Overhage, M.M. Shabot, T. Sheridan, Reducing the frequency of errors in medicine using information technology, *J Am Med Inform Assoc* **8** (2001), 299-308.
- [3] E. Ammenwerth, Evidence-Based Health Informatics: How do we know what we know? *Methods Inf Med* **54**(4) (2015), 298-307.
- [4] R. Koppel, T. Wetterneck, J.L. Telles, B.T. Karsh, Workarounds to barcode medication administration systems: their occurrences, causes, and threats to patient safety, *J Am Med Inform Assoc* **15** (2008), 408-423.
- [5] F. Magrabi, M. Baker, I. Sinha, M.S. Ong, S. Harrison, M.R. Kidd, W.B. Runciman, E. Coiera, Clinical safety of England's national programme for IT: A retrospective analysis of all reported safety events 2005 to 2011, *International Journal of Medical Informatics* **84** (2015), 198-206.
- [6] Y.Y. Han, J.A. Carcillo, S.T. Venkataraman, R.S. Clark, R.S. Watson, T.C. Nguyen, H. Bayir, R.A. Orr, Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system, *Pediatrics* **116** (2005), 1506-1512.
- [7] M.I. Harrison, R. Koppel, S. Bar-Lev, Unintended Consequences of Information Technologies in Health Care – An Interactive Sociotechnical Analysis, *J Am Med Inform Assoc* **14** (2007), 542-549.
- [8] M.C. Beuscart-Zephir, C. Nøhr, Human factors engineering for computer-supported identification and prevention of adverse drug events, *Stud Health Technol Inform* **148** (2009), 14-24.
- [9] R. Koppel, D.A. Kreda, Healthcare IT usability and suitability for clinical needs: challenges of design, workflow, and contractual relations, *Stud Health Technol Inform* **157** (2010), 7-14.
- [10] E. Coiera E, Interaction design theory, *International Journal of Medical Informatics* **69** (2003), 205-222.
- [11] R. Beuscart, P. McNair, S.J. Darmoni, V. Koutkias, N. Maglaveras, M.C. Beuscart-Zephir, C. Nøhr, Patient safety: detection and prevention of adverse drug events, *Stud Health Technol Inform* **150** (2009), 968-971.
- [12] YY. Han, J.A. Carcillo, S.T. Venkataraman, R.S. Clark, R.S. Watson, T.C. Nguyen, H. Bayir, R.A. Orr, Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system, *Pediatrics* **116** (2005), 1506-1512.
- [13] M. Berg, *Health information management: Integrating information and communication technology in health care work*, Routledge; 2003.
- [14] P. Carayon, Human factors of complex sociotechnical systems, *Appl Ergon* **37** (2006), 525-535.
- [15] M. Berg, Patient care information systems and health care work: a sociotechnical approach, *International Journal of Medical Informatics* **55** (1999), 87-101.
- [16] P. Carayon, E.J. Bass, T. Bellandi, A.P. Gurses, M.S. Hallbeck, V. Mollo, Sociotechnical systems analysis in health care: a research agenda, *IIE transactions on Healthcare Systems Engineering* **1** (2011), 145-160.
- [17] T.G. Kannampallil, G.F. Schauer, T. Cohen, V.L. Patel, Considering complexity in healthcare systems, *Journal of Biomedical Informatics* **44** (2011), 943-947.
- [18] E.M. Borycki, M. Househ, A.W. Kushniruk, C. Kuziemsky, Use of qualitative methods across the software development lifecycle in health informatics, *Stud Health Technol Inform* **164** (2011), 293-297.
- [19] A.W. Kushniruk, V.L. Patel, Cognitive and usability engineering methods for the evaluation of clinical information systems, *J Biomed Inform* **37** (2004), 56-76.
- [20] A. Kushniruk, C. Nøhr, S. Jensen, E.M. Borycki, From Usability Testing to Clinical Simulations: Bringing Context into the Design and Evaluation of Usable and Safe Health Information Technologies, *Yearb Med Inform* **8** (2013), 78-85.
- [21] E. Borycki, A. Kushniruk, Identifying and preventing technology-induced error using simulations: application of usability engineering techniques, *Healthc Q* **8** Spec No (2005), 99-105.
- [22] S. Jensen, K.M. Lyng, C. Nøhr, The role of simulation in clinical information systems development, *Stud Health Technol Inform* **180** (2012), 373-377.
- [23] S. Jensen, C. Nøhr, S.L. Rasmussen, Fidelity in clinical simulation: how low can you go? *Stud Health Technol Inform* **194** (2013), 147-153.
- [24] S. Jensen, S.L. Rasmussen, K.M. Lyng, Evaluation of a Clinical Simulation-based Assessment Method for EHR-platforms, *Stud Health Technol Inform* **205** (2014), 925-929.
- [25] M. Berg, *Health Information Systems. Integrating information technology in health care work*, Routledge, 2006.
- [26] P. Carayon, *Handbook of human factors and ergonomics in health care and patient safety*, CRC Press, 2011.

- [27] J.S. Ash, M. Berg, E. Coiera, Some unintended consequences of information technology in health care: the nature of patient care information system-related errors, *J Am Med Inform Assoc* **11** (2004), 104-112.
- [28] A. Kushniruk, P. Turner, A framework for user involvement and context in the design and development of safe e-Health systems, *Stud Health Technol Inform* **180** (2012), 353-357.
- [29] E.M. Borycki, L. Lemieux-Charles, L. Nagle, G. Eysenbach, Evaluating the impact of hybrid electronic-paper environments upon novice nurse information seeking, *Methods Inf Med* **48** (2009), 137-143.
- [30] C. Pearce, A. Shachak, A. Kushniruk, S. de Lusignan, Usability: a critical dimension for assessing the quality of clinical systems, *Inform Prim Care* **17** (2009), 195-198.
- [31] M. Berg, *Health Information Technology in Health Care Work*, Routledge, 2004.
- [32] ISO 9241-210: *Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems*. International Standardization Organization (ISO), Switzerland, 2009.
- [33] M. Hertzum, Images of usability, *Intl Journal of Human-Computer Interaction* **26** (2010), 567-600.
- [34] M. Maguire, Methods to support human-centered design, *Intl Journal of Human-Computer Interaction* **55** (2001), 587-634.
- [35] D. Robey, D. Farrow, User involvement in information system development: A conflict model and empirical test, *Management Science* **28** (1982), 73-85.
- [36] L. Damodaran, User involvement in the systems design process-a practical guide for users, *Behaviour & Information Technology* **15** (1996), 363-377.
- [37] S.L. Rasmussen, K.M. Lyng, S. Jensen, Achieving IT-supported standardized nursing documentation through participatory design, *Stud Health Technol Inform* **180** (2012), 1055-1059.
- [38] Y. Dahl, O.A. Also, D. Svanæs, Fidelity Considerations for Simulation-Based Usability Assessments of Mobile ICT for Hospitals, *International Journal of Human-Computer Interaction* **26** (2010), 445-476.
- [39] S. Jensen, S. Vingtoft, C. Nøhr, Benefits of a clinical planning and coordination module: a simulation study, *Stud Health Technol Inform* **183** (2013), 220-224.
- [40] S. Jensen, *Use of clinical simulation in development and implementation of clinical information systems*, PhD thesis, University of Aalborg, 2014.
- [41] H.P. Lundsgaarde, R.M. Gardner, R.L. Menlove, *Using attitudinal questionnaires to achieve benefits optimization*, in: Proceedings of the Annual Symposium of the American Medical Informatics Association, 1989, pp. 703-708.
- [42] B. Kaplan, Addressing organizational issues into the evaluation of medical systems, *J Am Med Inform Assoc* **4** (1997), 94-101.
- [43] E. Ammenwerth, S. Graber, G. Herrmann, T. Bürkle, J. König, Evaluation of health information systems-problems and challenges, *Int J Med Inform* **71** (2003), 125-135.
- [44] E. Ammenwerth, W.O. Hackl, K. Binzer, T.E. Christoffersen, S. Jensen, K. Lawton, P. Skjoet, C. Nøhr, Simulation Studies for the evaluation of health information technologies: experiences and results, *HIM J* **41** (2012), 14-21.

Economic Evaluation of Health IT

Daniela LUZI^{a,1}, Fabrizio PECORARO^a and Oscar TAMBURIS^b

^a*Institute for Research on Population and Social Policies,*

National Research Council, Italy

^b*Department of Veterinary Medicine, University of Naples Federico II,*
Italy

Abstract. Economic evaluation in health care supports decision makers in prioritizing interventions and maximizing the available limited resources for social benefits. Health Information Technology (health IT) constitutes a promising strategy to improve the quality and delivery of health care. However, to determine whether the appropriate health IT solution has been selected in a specific health context, its impact on the clinical and organizational process, on costs, on user satisfaction as well as on patient outcomes, a rigorous and multidimensional evaluation analysis is necessary. Starting from the principles of evaluation introduced since the mid-1980s within the Health Technology Assessment (HTA) guidelines, this contribution provides an overview of the main challenging issues related to the complex task of performing an economic evaluation of health IT. A set of necessary key principles to deliver a proper design and implementation of a multidimensional economic evaluation study is described, focusing in particular on the classification of costs and outcomes as well as on the type of economic analysis to be performed. A case study is eventually described to show how the key principles introduced are applied.

Keywords. Health information technology, technology assessment, economic evaluation.

1. Introduction

The successful application and the consequent systematic adoption of a Health Information Technology (health IT) are broadly considered a promising strategy to improve the quality and delivery of health care. However, to determine whether the appropriate health IT solution has been selected in a specific health context, its impact on the clinical and organizational process, on costs, on user satisfaction as well as on patient outcomes, a rigorous and multidimensional evaluation analysis is necessary.

Since the mid-1990s an increasing number of studies have addressed this issue, and some of them also include an economic evaluation with the aim of providing decision makers with a set of analyses that can support them in prioritizing interventions and maximizing the available limited resources for social benefits [1-2]. Being the “study of choice” [3], health economic evaluation is defined as the “*comparative analysis of alternative courses of action in terms of both their costs and*

¹

Corresponding author: Dr. Daniela Luzi, Institute for Research on Population and Social Policies, National Research Council, Via Palestro 32, 00185 Roma Italy, d.luzi@irpps.cnr.it.

consequences. Therefore, the basic task of any economic evaluation is to identify, measure, value and compare the costs and consequences of alternatives being considered” [4].

Health IT has a supporting role both in the care process (diagnostic, treatment/therapy and nursing) and in the auxiliary process (for instance, appointing making, image archiving and documentation) [5]. Therefore, a causal relation [6] between the improvement of the efficacy and effectiveness of both care and auxiliary processes is usually very difficult to determine and measure. This makes the economic assessment of the health IT value – preferably in monetary terms – a challenging task, for a number of reasons e.g. the difficult identification of costs given, among other problems, the incremental development of many health IT solutions and their often locally adjusted implementation; the measurement of benefits that generally also depend on how systems are used, the organization and medical context in which they are embedded, and even on the national health system in place.

As health IT does not directly alter the states of health or disease [7], compared to other types of technologies such as drugs or medical devices, benefits have to be measured in terms of changes in the health care and management processes, for instance improvements resulting from a better sharing of patients’ information, better resource allocation, or workload definition and deployment. This also implies that the economic evaluation has to combine or privilege different methods – qualitative and quantitative – that have to be coherently selected according to the objectives and perspective driving the assessment. Moreover, the difficulty in isolating the impact of health IT may be partially solved by distinguishing the main functionalities in: 1) capturing, storing, managing and sharing data; 2) informing and supporting clinical decision-making; 3) delivering expert professional and or consumer care remotely [8].

Economic evaluation of health IT is still a research area. There are no common agreed and fixed standards that guide the performance of health economic analysis [9] considering the multiple dimensions on which health IT may have an impact. For these reasons systematic reviews generally reveal a limited number of economic evaluation compared to other types of analysis [10-11], poor use of analytical technique and documentation, partial identification of costs and benefits, or use of predictive analysis on assumptions based on limited empirical data. Moreover, although different evaluation frameworks and guidelines have been proposed (none of them specifically focused on economic evaluation of health IT), there is no uniform reporting of results, thus limiting the comparison across institutions.

This contribution intends to contribute to the discussion on the methods and approaches supporting the assessment of health IT solutions by providing key features that support a scientifically sound economic evaluation. Section 2 provides a brief overview of the economic evaluation within the health technology assessment (HTA) framework as well as in a selected number of widely diffused health IT evaluation models. Section 3 summarizes some key principles that guide to a proper design and implementation of an economic evaluation of health IT, focusing in particular on the classification of costs and outcomes as well as on the main criteria used to choose the type of analysis. These principles are applied in a case study described in section 4.

2. The framework of the economic evaluation development

Economic evaluation was the major focus of the first governmental national agencies that were constituted to develop Health Technology Assessment (HTA) round the mid 1970s. The main concern was the rising expenditure for health care, the rapid change of health technology generally associated with the ageing of population and increased population health care service demand. The establishment of the US Office of Technology Assessment (OTA) – replaced by the Agency for Health Care Policy and Research which in turn became the Agency for Health Research and Quality (AHRQ) – clearly identified its scope as provider of analyses to support decision makers in “formulating policies to ensure that research-and-development funds are invested wisely” [12].

The subsequent development of similar national agencies in Europe, even if generally motivated by rationalising health care expenditure and by cost containments, soon addressed issues more closely related to quality and safety of care as well as social and ethical implications [13]. This has led to the adoption of a more comprehensive approach to technology assessment that considers economic evaluation as part of a more complex framework of analysis that includes – at least at the level of HTA scope statements [14] – the technological, patient and organisational dimensions.² Moreover, other evaluation approaches developed within the Cochrane Collaboration and the evidence-based medicine (EBM) movement³ contributed to the consideration of economic evaluation as a specific phase of the assessment process, generally performed after safety, efficacy and effectiveness of interventions have been analysed [15].

The application of HTA differs from country to country, being influenced by the national health care system in place, and by the aim and mandate of the agencies performing the assessment. This pertains also to the economic evaluation that depending on the national agency tends to privilege certain types of analysis (for instance cost-utility instead of cost-effectiveness) and/or prefers to consider certain types of cost and/or benefit [16]. Moreover, even if most HTAs have broadened the range of technology to include drugs, medical device, procedures and organisational and support system for care provision, the majority of analysis are generally focused on pharmaceutical products. This has the consequence that traditionally applied methods to verify safety and efficacy of drugs such as RCT (Randomized Clinical Trials) have been privileged making the evaluation of the impact of health IT limited to certain aspects, such as system performance or particular changes in clinical practices that may affect patient care [17, 18].⁴

Despite differences and specificities, HTA has had the merit of providing a set of principles for the conduct of a sound evaluation defining the main steps and contents of the study design, providing guidance on types of economic analysis to be performed, on criteria and methods to be followed in the collections and analysis of data as well as in the reporting of evaluation results.

However, the need to specifically address the evaluation of health IT has led to the development of further frameworks, differently connected with HTA, that are

² See also: P. Doupi, Evolving Health IT systems evaluation: the convergence of health informatics and HTA, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

³ See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: ibid

⁴ See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: ibid.

conventionally classified as subjectivist approaches [17, 18]. These frameworks complement each other [19], as they each tend to privilege a specific perspective of the health IT evaluation, focusing on user behaviour and perception, or emphasising social/organizational relationships or software lifecycle. They are generally based on qualitative approaches that use among other methods interviews, questionnaires, or focus groups to perform their analysis (see reviews [20, 21] that use this classification of frameworks). Moreover, comprehensive and multi-dimensional frameworks have been developed to include the different aspects that influence health IT adoption and applying matrix and/or taxonomy to identify the main components to be taken into account in the evaluation.

Worth mentioning is the Information System Success model proposed by DeLone and McLean [6, 22], which provides a framework of interconnected aspects that should be considered also when performing an economic evaluation. The model is based on a taxonomy of six interrelated dimensions that measure the *system quality* (e.g. system performance and use), *information quality* (e.g. accuracy, reliability, etc.), *service quality* (e.g. the overall support delivered by the service provider), *system use* (e.g. human acceptance or resistance toward the system), *user satisfaction* (e.g. positive experiences in using the system) and *net benefit* (e.g. the combination of individual and organizational impact). The first three dimensions are to be measured singularly or jointly to evaluate how they affect the two closely interrelated variables of *system use* and *user satisfaction* so to ascertain the *net benefit*, which in the DeLone and McLean previous version of the model [6] were described as the individual and organisational impact. *Net benefit* thus summarises the outcomes of this complex interaction providing a value – a positive or negative association in DeLone and McLean terms – that can be transformed into an economic evaluation.

Further developments [23, 24] of the Information System Success model have given in more recent times a major focus on the organisational component and identify a more complex set of interactions among the dimensions identified by DeLone and McLean. The category of net benefits, common to these frameworks, helps in the identification of outcomes derived by the interaction of these dimensions and provides the basis for both qualitative and quantitative analyses on which to derive for instance cost reduction resulting from productivity and/or reduced time in performing specific tasks, error decline in terms of adverse events as well as impact on patient care and access to information.

3. Principles of economic evaluation of health IT

Guidelines on evaluation agree on the importance of the identification of a specific and clear research question that details the purpose of the analysis.⁵ The scope of the economic evaluation also defines the perspective of the analysis that has to match the need of the commissioning body that generally poses the study question. The scope and perspective of the research question determine the type of study design as well as the appropriate approach to analyse data collected during the evaluation framework. The key elements of the economic evaluation framework are shown in Figure 1 and

⁵ See also: P. Nykänen et al., Quality of health IT evaluations, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

described in the following paragraphs. They are based on the criteria described in selected HTA guidelines containing a specific part on economic analysis [1, 25-27].

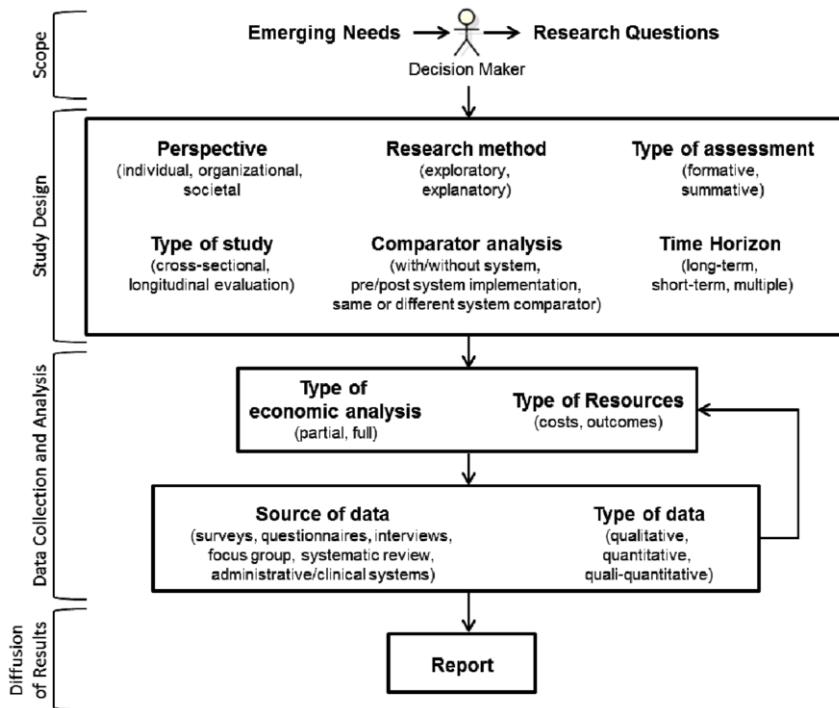


Figure 1. Principles of the economic evaluation of health IT.

3.1. Study design

Depending on the research question, the study design has to consider the key principles listed below and choose for each one the appropriate approach.⁶

The *perspective* represents the point of view from which the study is conducted (individual, organizational, societal). Clearly establishing the perspective of the economic evaluation is particularly important for the identification of costs, resources and consequences to be examined. This also ensures comparability of different analyses. The perspective can be limited to the primary stakeholders of the health care system (e.g. physicians), or it can consider impact on the organization or even on the welfare system. In the latter case, a wider range of relevant costs and consequences are considered including those that are related to other public sector agencies, patients or their carers.

The identification of the *research method* [28] is mainly based on the knowledge about the problem to be analysed. When the problem is not well defined the study is conducted using an exploratory research, for example using a case study to generate a hypothesis and find the relationships between the introduction of a new technology and

⁶ See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

its effect on the context where it is deployed. In this approach data are collected from literature reviews, databases and/or from relevant stakeholders (e.g. physicians, patients) using techniques such as informal discussions, in-depth interviews, or focus groups. Conversely, an explanatory research is adopted when the investigation is conducted analysing the relation between the cause (e.g. technology to be adopted) and the effect (e.g. costs and outcomes) derived from the introduction of the health IT. This relationship is explored using two main research methods: experimentation (e.g. randomized clinical trial performed in a hospital), and statistical research (e.g. multiple regression techniques). A clear identification of the research method is helpful to determine the best research design and data collection method as well as the selection of the target population.

The *type of assessment* indicates in which phase of the development lifecycle the health IT is analysed. Substantially, a technology can be evaluated throughout the whole development lifecycle using a formative approach, providing information on the system under development that may also lead to improvement or modifications. Once the system has been implemented, assessment of the effect/outcome is performed using a summative evaluation.

The *type of study* determines whether the relationship between costs and outcomes deriving from the introduction of a new technology is analysed at one particular time during the system deployment (i.e. cross-sectional study) or repeatedly observed over time with continuous monitoring (i.e. longitudinal study). Type of study also includes the identification of inclusion and exclusion criteria of the target population.

The identification of the *comparator* is one of the most significant activities of the economic evaluation framework. The new technology can be introduced as an improvement of existing, generally paper-based, routine care system (i.e. pre/post system implementation). In this case one or more relevant alternatives of the health IT under evaluation could be taken into account (same or different system comparator). These circumstances can involve either information systems that are classified in the same group of health IT or systems that share only a small set of functionalities. Moreover, it is also possible to evaluate a new process implemented by means of a health IT (i.e. with/without system comparison) to verify costs and benefits of the chosen solution.

The appropriate *time horizon* of the evaluation specifies the period during which all the costs and outcomes are captured (short or long-term). It strictly depends on the research questions and can vary from a few days to several years capturing changes in the patient's health status and/or impact of health IT over an expected time period. This implies the identification of outcomes and costs of alternative options measured in the specified period. It is also possible to explore multiple time horizons to verify the cost effectiveness of a health IT based on alternative scenarios.

Once the scope has been identified and the study design determined, data collection and analysis can be performed. This implies on the one hand the choice of the most fitting *type of economic analysis* (to be identified within the full and partial analysis frameworks) and the selection of related *type of resources* (in terms of both costs and outcomes). On the other hand, it implies the identification of the *types of data* (e.g. qualitative and/or quantitative) as well as the *source of data* (e.g. systematic reviews, surveys, clinical information systems already deployed). The backwards arrow in Figure 1 sets forth the mutual influence between the data and the analysis domains: the decision towards the use of a given economic analysis somehow conditions the data retrieval; conversely, the deployment of a specific type of analysis might depend on the

purpose of the economic evaluation as well as on the availability of suitable data. It should also be noted that a combination of more than one type of analysis could be useful. The next sections describe in detail the classification of costs and outcomes and the different types of methods included within the framework of full and partial economic analysis.

A structured *report of results* of the economic evaluation ensures eventually that the performed study is thoroughly presented and organized consistently to facilitate review and comparison by decision makers [29].⁷ The report has to be presented in a clear and transparent manner with enough information provided according to a consolidated schedule [30]. The Executive Summary and Conclusions should be written so that they can be understood by a non-technical reader, in order to enable the audience to critically evaluate the validity of the analysis. It is essential to explain and justify the choice of variables and methods, mention the reasons why certain data were excluded and last but not least describe in detail the organisational characteristics that may hinder or facilitate a health IT introduction or maintenance. However, it is likely that the results may not be (totally or in part) generalizable, as the key principles may differ significantly, e.g. between different jurisdictions or time periods.

3.2. Classification of costs and outcomes

The economic evaluation of health IT includes the identifications of costs to be quantified in monetary terms generally related to infrastructure (e.g. hardware, software, network), personnel (e.g. time spent for users' training), facility (e.g. space necessary to store the technology) and other materials (e.g. consumable, paper) [e.g. 31, 32]. Table 1 summarizes the different classes of costs as reported by referenced relevant literature.

Table 1. Classification of costs.

Description of costs categories (with references)	Example of costs
Tangible and intangible [e.g. 33] (level of measurability of the cost)	Tangible: tablets Intangible: stress caused to a patient due to the health IT
Direct and indirect [e.g. 34] (impact of the health IT)	Direct: information system implementation Indirect: loss of productivity
Health and non-health [e.g. 35] (cost related or not to the health sector)	Health: outpatient visits Non-health: private travel costs
One time and ongoing [e.g. 36] (cost is considered once or repeatedly)	One time: local area network installed in the health facility Ongoing: software maintenance
Average and marginal [e.g. 37] (cost is considered as a total amount or as a price per unit)	Average: software implementation Marginal: personal computers
Fixed and variable [e.g. 38] (cost remains constant or vary in proportion of the activities performed)	Fixed: initial user training Variable: telephone bills

Easily identifiable costs are generally related to the health IT implementation and maintenance as well as to the infrastructure supporting its deployment (e.g. PCs, network, printers). However, given that the introduction of a new technology impacts

⁷ See also: E. Ammenwerth et al., Publishing health IT evaluation studies, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

on the core organizational and clinical processes, identification of indirect costs such as time spent for training and/or for modifying use of the technology (loss of productivity) are difficult to measure and therefore frequently overlooked and/or subjectively attributed to different classes [36].

Similarly to costs, also outcomes can be classified as direct (e.g. investment reduction of personnel wages) and indirect (e.g. savings resulted from the decrease of adverse events) depending on whether the monetary savings are strictly related or induced by the introduction of the health IT. Moreover, outcomes can be related with the health of the patient (e.g. the reduction of medical errors) or not (e.g. time and money saved due to the reduction of patient transportation in a telemedicine program) [39]. However, the identification and classification of outcomes are even a more challenging task if compared with costs, as outputs are generally intangible and indirect measures related to the improvement of the patient's health status as well as of the organizational process. This issue is also crucial considering that a parameter can describe more than one category of benefits implying an overestimation of outcomes. For instance adverse event prevention can be measured as an improvement of both quality of care and patient safety.⁸

Many studies classify the same parameter either as a cost or an outcome of the health IT deployment. For instance, patient's length of stay can be considered either as a cost [40, 41] or as a consequence of the intervention [42, 43] depending on the point of view of the analysis. It is therefore essential that authors give in the first place, to the greatest extent possible, a clear indication of the nature (costs or benefits) of the parameters used to perform the evaluation, in order to justify the results of the evaluation as well as allow its comparison with similar studies.

Outcomes are not only a measure of the increase of revenues but also an assessment of the costs averted as a consequence of the introduction of the health IT. Their measurement implies a careful analysis as some costs may not be simply eliminated, but shifted to other hospital services or even to different components of the health care system [44]. This makes it also challenging to transform outcomes into a monetary value that is a necessary activity when the economic evaluation is performed using a Cost Benefit Analysis. For this reason analysts have often chosen other types of analysis that do not imply this conversion, such as Cost Analysis, or Cost Effectiveness Analysis [45].

The difficulty in the identification and classification of outcomes has led different authors to adopt customized classifications considering, for instance, the impact of the Electronic Health Record [46, 47] that can result in outcomes about the patient flow (e.g. reduction of patient cycle time and increasing patient capacity), resource allocation (e.g. transcription, chart management and paper consumption), coding and billing (e.g. reduced billing errors), patient safety (e.g. decrease in infection rate), caring process (e.g. high quality of care) and staff compliance (e.g. reducing the redundancy of laboratory tests). Finally, the evaluation toolkit provided by the AHRQ [48] classifies the different measures that can be used to assess a health IT project in the following categories: patient safety (e.g. hospital complication rates), effectiveness (e.g. mortality), quality of care (e.g. documentation of key clinical data elements), efficiency (e.g. length of stay), and patient centeredness (e.g. patient knowledge). The

⁸ See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

assessment of costs and outcomes included in the economic evaluation should take into account that the technology can lose its validity in a relatively short period of time. This is particularly true considering that health IT may become obsolete quite quickly, making it necessary to quantify costs to be invested to replace the technology after its use life as well as to consider the fast decline of prices (e.g. devaluation) of the technology that can be also caused by an increased value of production and a recovery of development charges. Moreover, an important aspect to be taken into account when performing an economic analysis is that the impact of a health IT often considers a broad period of time (for instance, Cost Benefit Analysis conducted on a 5-years period) that requires the correction of costs and outcomes for the effects of inflation to provide realistic resource costs.

3.3. Type of economic analysis

A health economic analysis aims to identify criteria to support decision makers in choosing between competitive alternatives the one which is most efficient and cost-effective in an environment with limited resources [4, 10, 32, 33, 48]. This comprehensive analysis is achieved within the framework of a full economic evaluation when both costs and consequences of alternative interventions (e.g. intervention X versus comparator Y) are compared to assess their efficiency. A partial economic evaluation occurs instead when costs and outcomes are separately analysed (cost analysis/cost description; efficacy or effectiveness evaluation/outcome description) and/or alternative solutions are not considered (cost outcome description). Systematic reviews [34] indicate that the majority of economic evaluation studies generally perform cost analyses that focus on cost saving of two or more alternatives.

Full economic evaluation represents a framework composed of different types of analysis, which are applied depending on the research questions, the viewpoint of the decision maker as well as data availability. Table 2 reports the most frequently adopted types of analysis giving a general description, the main objective as well as criteria that have to be fulfilled when choosing the appropriate method. What differentiates these analyses is the metric used as well as the number of parameters considered to evaluate the outcomes of the different interventions. The Cost Effectiveness Analysis (CEA) measures the health effects using a single outcome, such as the life years gained, while the Cost Utility Analysis (CUA) considers one or more outcomes aggregated in a global measure of health outcome, such as the QALY (Quality-Adjusted Life Years) or DALY (Disability Adjusted Life Years). CEA and CUA may use an incremental ratio – respectively, incremental cost-effectiveness ratio (ICER) and incremental cost-utility ratio (ICUR) – that allows comparison of the effectiveness of the intervention against an alternative solution given a fixed budget. When outcomes can be transformed in a monetary term a Cost Benefit Analysis (CBA) can be applied. However, even if this method can provide a useful indication for the right allocation of resources measuring whether gains overweight costs, its application has to face ethical issues as it means placing a value on the cost of human life.

Table 2. Types of full economic analysis.

Methodology	Description	Objective	Application Criteria
<i>Cost Effectiveness Analysis (CEA)</i>	Consequences of different health interventions are measured in natural units using a single outcome related to the objective of the program (e.g. life-years gained, adverse events avoided).	To establish whether differences in expected costs between interventions can be justified in terms of changes in expected health effects.	<ul style="list-style-type: none"> Different interventions have to be compared using an uniform measurement of a single outcome Outcomes cannot be expressed in monetary terms
<i>Cost Utility Analysis (CUA)</i>	As an extension of the CEA, it measures the strength of preference for a particular clinical outcome state. Outcomes are measured using QALY or DALY gained.	To compare the value of interventions for different health problems, in order to facilitate the allocation of resources to maximize health gains.	<ul style="list-style-type: none"> Meaningful differences in the combination of the duration of life and health-related quality of life (HRQoL) between the interventions have to be demonstrated Outcomes are not expressed in monetary terms
<i>Cost Benefit Analysis (CBA)</i>	It measures and values in monetary terms the benefits and costs of outcomes achieved from different programs or interventions.	To address the efficiency in allocating resources between sectors	<ul style="list-style-type: none"> Outcomes have to be expressed in monetary terms

Moreover, there are two additional types of analysis that are not reported in the Table 2 as they represent two specific forms of CEA: Cost Minimization Analysis (CMA) and Cost Consequence Analysis (CCA). In the CMA outcomes of alternative interventions have been proven to be identical and therefore only the least expensive option has to be determined. In the CCA multiple outcomes are analysed separately and compared with the relevant costs. This has the advantage of considering the full range of health and organisational effects of an intervention or when it is difficult or misleading to combine multiple outcomes from an intervention in a QALY for a CUA.

4. Case study

In this section we model a timely implementation of economic evaluation for health IT providing a case study based on the key principles described in section 3. Characteristics of the environment are: a mid-sized hospital (300 beds and 145 care professionals) that comprises Intensive Care Units (ICUs) hosting patients with comorbidities treated with multiple drugs. The hospital is already equipped with an Electronic Health Record (EHR) system that manages clinical and administrative patient data. The General Directorate intends to integrate the existing EHR with a Clinical Decision Support (CDS) module to overcome the current paper-based prescription procedures. The main scope is to support physicians in the choice of the appropriate medical treatment (drugs type and dosage), taking also into account the interaction with other drugs. Table 3 summarizes the key principles of the economic evaluation.

Table 3. Key principles of economic evaluation of the integration of a CDS module within an EHR system.

Principle	Description	Scope of the economic analysis
Decision maker	General Directorate of the hospital	
Emerging Needs	Reduction of Adverse Drug Events (ADEs) caused by prescription errors that derive from: <ul style="list-style-type: none"> interaction with other therapies (drug-drug interaction, DDI); dosage and/or length of the therapy; type of medicine prescribed. 	
Research questions	<ul style="list-style-type: none"> Will the integration of the existing EHR with a CDS module improve the quality of care compared with the actual paper-based prescription procedure? Is there particular evidence that the adoption of CDS modules reduce ADEs? Will outcomes derived from the CDS balance the implementation and adoption cost? 	
		Study design
Perspective	Organizational: integration of the already deployed EHR system with a CDS module to improve the quality of treatment via the implementation of e-prescription procedures.	
Research methods	A literature review is carried out to collect and analyse evidence on outcomes derived by the adoption of CDS module in other contexts (e.g. PubMed, Cochrane, AHRQ, York).	
Type of study	Cross-sectional: the evaluation is conducted considering the number of ADEs occurred in a year, in a hospital with similar environmental characteristics.	
Time horizon	The evaluation considers the costs and outcomes over a 5-years period.	
Comparator analysis	Pre-post ADE alert system implementation: manual data entry of drug prescription procedures into EHR system versus EHR system integrated with a CDS module.	
Type of assessment	Formative: the CDS module is assessed prior to its implementation.	
		Data collection and analysis
Type of economic analysis	Incremental Cost Effectiveness Analysis	
Source of data	Literature review; Open databank provided by the Ministry of Health; Budget proposal by vendors	
Type of data	Quantitative	
Costs	<pre> graph TD Costs --> Health Costs --> NonHealth[Non-health] Health --> Direct Health --> Ongoing Direct --> OneTime[One time] OneTime --> CDSimpl[CDS implementation (considering also its integration with the EHR)] Ongoing --> CDSmain[CDS maintenance] NonHealth --> Ongoing2[Ongoing] Ongoing2 --> Training[Training of users (hospital staff)] </pre>	
Outcome	Productivity loss and hardware costs are not included considering that users are already confident with the use of health IT and the hospital is already equipped with PCs and printers. Costs of process changes have not been considered as CDS module effects only a limited part of the process.	
	The number of ADEs that could be averted has been included as a unique indirect outcome	

The result of the Cost Effectiveness Analysis is reported in Table 4 highlighting costs to implement and maintain the CDS module as well as to train the physicians in its use. Costs have been measured based on the budget proposed by selected vendors and represented in US Dollars in order to pursue an as broad as possible visibility and data usability. The number of ADEs that could occur in a year have been captured from

an open data source released by the Italian Ministry of Health⁹ considering a health structure with the same environmental characteristics of the one under investigation; while the expected reduction of ADEs has been obtained from a literature review where different studies [32, 48-51] have reported that the introduction of a CDS module can reduce the number of adverse events by 40% to 80% each year. This wide range of percentage reduction makes it also necessary to perform a sensitivity analysis. Starting from the total costs and outcomes, the Incremental Cost Effectiveness Ratio (ICER) has subsequently been computed to determine the US dollars spent per ADE averted.

Table 4. Results of the cost effectiveness analysis (in US Dollars).

	0-year	1-year	2-year	3-year	4-year	5-year	Total
<i>Costs (expressed in US Dollars)</i>							
CDS implementation	500.000						500.000
CDS maintenance		50.000	50.000	50.000	50.000	50.000	250.000
<i>User training (per user)</i>	200	150					
User training (total) for 145 physicians	29.000	21.750					50.750
<i>Total costs</i>	529.000	71.750	50.000	50.000	50.000	50.000	800.750
<i>Outcomes (based on initial 183 ADEs)</i>							
# of ADEs (60% of ADEs averted)		110	110	110	110	110	550
# of ADEs (40% of ADEs averted)		73	73	73	73	73	365
# of ADEs (80% of ADEs averted)		146	146	146	146	146	725
<i>Cost-effectiveness analysis (Total expressed in US Dollar spent per ADE averted)</i>							
ICER (60% of ADEs reduced)							1455.9
ICER (40% of ADEs reduced)							2193.8
ICER (80% of ADEs reduced)							1104.5

Limitations of the study: The literature review to assess the percentage on ADEs reduced by the introduction of a CDS is based on heterogeneous studies considering specific functionalities implemented, population involved as well as the study design adopted. Moreover, ADEs are measured using different methodologies that often do not take into account non-intercepted ADEs (e.g. ADE occurred after the discharge). Another important aspect to be considered in this study is that ADEs are surrogate measures not necessarily directly related to changes in the patient-relevant medical outcomes. Moreover, the number of ADEs considered does not take into account the degree of severity of the adverse events.

Note that the proposed simplified case study has to be considered as an educational example of the application of the principles of economic evaluation of health IT described in the previous section. When the evaluation analyses the replacement of a paper-based procedure, it is necessary to assess process changes that introduce a set of specific dimensions such as savings of ceasing old processes as well as costs of new processes, equipment costs, loss of production due to the introduction of a health IT.

⁹ http://www.salute.gov.it/portale/documentazione/p6_2_8_1_1.jsp?id=6

5. Conclusions

The rapidly changing technology as well as its adoption in increasing health-related environments (suffice it to think of m-Health applications) requires the economic evaluation to become an on-going assessment that includes a multidisciplinary team of experts to comprehensively consider the benefits of health IT introduction and use.

The present contribution aimed to enrich the line of inquiry into economic evaluation approaches for the adoption and implementation of health IT, as a means to support decision makers in prioritizing interventions and maximizing the available limited resources for social benefits. The vast literature analysis conducted made clear that, though it is not possible to diverge from the principles of HTA, a sort of new interpretation (far from an adjustment) is necessary when applying the economic evaluation on health IT. This is a challenging task, as no consensus exists regarding the multiple dimensions to be considered when evaluating the indirect effects on patients' health status as well as the impact on both health care and managerial processes. To this purpose, the authors' main effort was to outline a set of guiding principles to conduct an appropriate analysis of costs and outcomes as well as to choose the proper type of economic evaluation. The case study has then applied the set of criteria emerging from the mentioned principles that can lead to a timely and consistent evaluation.

Recommended further readings

1. M.F. Drummond, M.J. Sculpher, G.W. Torrance, B.J. O'Brien, G.L. Stoddart, *Methods for the economic evaluation of health care programmes*, Oxford University Press, 2005.
2. J. Car, A. Black, C. Anandan, K. Cresswell, C. Pagliari, *The Impact of eHealth on the Quality & Safety of Healthcare. A Systematic Overview & Synthesis of the Literature*, Report for the NHS Connecting for Health Evaluation Programme March 2008. <https://www1.imperial.ac.uk/resources/32956FFC-BD76-47B7-94D2-FFAC56979B74>, last access 11 February 2016.
3. CP Friedman, J C Wyatt, *Evaluation Methods in Biomedical Informatics*, 2nd edition, Springer, New York, 2006.
4. P. Nykänen, J. Brender, J. Talmon, N. de Keizer, M. Rigby, M.C. Beuscart-Zephir, E. Ammenwerth, Guideline for good evaluation practice in health informatics (GEP-HI), *International Journal of Medical Informatics* **80** (2011), 815–827.
5. Health Information and Quality Authority, *Guidelines for the Economic evaluation of Health Technologies in Ireland*, 2014, <https://www.hiqa.ie/publication/guidelines-economic-evaluation-health-technologies-ireland>, last access 11 February 2016.

Food for thought

1. Which are pros and cons of a quantitative, objectivist research method?
2. What are the advantages and issues related to the performance of a formative economic evaluation compared to a summative one?

3. What are the criteria to be considered when choosing the type of economic analysis?
4. Think of some examples that describe the mutual relationship between the type of resources (costs and outcomes) and the availability of data to perform an economic analysis!
5. When health IT replaces a paper-based procedure, which are the difficulties in the classification of costs and benefits? Make some examples.

References

- [1] National Institute for Health and Care Excellence, *Guide to the methods of technology appraisal*, 2013, <http://publications.nice.org.uk/pgm9>, last access 11 February 2016.
- [2] M.S. Roberts, *Economic aspects of evaluation*, in: C.P. Friedman, J. Wyatt, Evaluation methods in medical informatics, New York, Springer, 2006, pp. 301-337.
- [3] A. Shiel, C. Donaldson, C. Mitton, G. Currie, Health economic evaluation, *J. Epidemiology & Community Health* **56** (2002), 85-88
- [4] M.F. Drummond, M.J. Sculpher, G.W. Torrance, B.J. O'Brien, G.L. Stoddart, *Methods for the economic evaluation of health care programmes*, Oxford University Press, 2005.
- [5] R.P. van der Loo, E.M.S.J. van Gennip, A.R. Bakker, A. Hasnianb, F.F.H. Rutten, Evaluation of automated information systems in health care: an approach to classifying evaluative studies, *Computer Methods and Programs in Biomedicine* **48** (1995), 45-52.
- [6] W.H. DeLone, E.R. McLean, Information Systems Success: The Quest for the Dependent Variable, *Information Systems Research* **3** (1992), 60-95.
- [7] C.S. Goodman, R. Ahn, Methodological approaches of health technology assessment, *International Journal of Medical Informatics* **56** (1999), 97-105.
- [8] A.D. Black, J. Car, C. Pagliari, C. Anandan, K. Cresswell, T. Bokun, B. McKinstry, R. Procter, A. Majeed, A.Sheikh, The impact of eHealth on the quality and safety of health care: a systematic overview, *PLoS medicine* **8**(1) (2011), e1000387.
- [9] J. Hjelmgren, F. Berggren, F. Andersson, Health Economic Guidelines – Similarities, Differences and some Implications, *Value in Health* **4** (2001), 225-250.
- [10] E.L. Eisenstein, D.B.A. Maqui Ortiz, K.J. Anstrom, et al., *Assessing the quality of medical information technology economic evaluation: rooms for improvements*, in: AMIA 2006 Symposium Proceedings, pp. 234-8.
- [11] T. Reardon, Research findings and strategies for assessing telemedicine costs, *Telemedicine Journal & e-Health* **11** (2005), 348-369.
- [12] Office of Technology Assessment, *Development of medical technology: opportunity for assessment*: U.S. Government Printing Office, Washington DC, 1976.
- [13] D. Banta, The development of health technology assessment, *Health Policy* **63** (2003), 121-132.
- [14] E. Draborg, D. Gyrd-Hansen, P.B. Poulsen, M. Horder, International comparison of the definition and the practical application of health technology assessment, *International Journal of Technology Assessment in Health Care* **21** (2005), 89-95.
- [15] R. Busse, J. Orvain, M. Velasco, M. Perleth, M. Drummond, F. Gürtnner, T. Jørgensen, A. Jovell, J. Malone, A. Rüther, C. Wild, Best practice in undertaking and reporting health technology assessments, *International Journal of Technology Assessment in Health Care* **18** (2002), 361-422.
- [16] J.M. Graf von der Schulenburg, C. Hoffmann, Review of European guidelines for economic evaluation of medical technologies and pharmaceuticals, *HEPAC Health Economics in Prevention and Care* **1** (2000), 2-8.
- [17] B. Kaplan, Evaluating informatics applications - some alternative approaches: theory, social interactionism, and call for methodological pluralism, *International Journal of Medical Informatics* **64** (2001), 39-56.
- [18] J.R. Moehr, Evaluation: salvation or nemesis of medical informatics? *Computers in Biology and Medicine* **32** (2002), 113-125.
- [19] M.M. Yosof, J.P. Paul, L.K. Stergioulas, *Towards a framework for health information evaluation*, in: Proceedings of 39th Hawaii International Conference on System sciences, Hawaii, IEEE 2006.
- [20] M.M. Yosof, A. Papazafeiropoulou, J.P. Paul, L.K. Stergioulas, Investigating evaluation frameworks for health information systems, *International Journal of Medical Informatics* **77** (2008), 377-385.

- [21] L.M. Currie, Evaluation frameworks for nursing informatics, *International Journal of Medical Informatics* **74** (2005), 908-916.
- [22] W.H. Delone, E.R. McLean, The DeLone and McLean model of information systems success: a ten-year update, *Journal of Management Information Systems* **19** (2003), 9-30.
- [23] M.S. Scott Morton, *The Corporation of the 1990s*, Oxford University Press, New York, 1991.
- [24] M.M. Yusof, J. Kulpis, A. Papazafeiropoulou, L.K. Stergioulas, An evaluation framework for Health Information Systems: human, organization and technology-fit factors (HOT-fit), *International Journal of Medical Informatics* **77** (2008), 386-398.
- [25] F.B. Kristensen, H. Sigmund (eds.), *Health Technology Assessment Handbook*, Danish Centre for Health Technology Assessment, National Board of Health, 2007.
- [26] Health Information and Quality Authority, *Guidelines for the Economic evaluation of Health Technologies in Ireland*, 2014.
- [27] Canadian Agency for Drugs and Technologies in Health, *Guidelines for the Economic Evaluation of Health Technologies*, Canada 3rd Edition, 2006.
- [28] E. Ammenwerth, N. de Keizer, Trends in Evaluation Research 1982-2002, *Methods of Information in Medicine* **44** (2005), 44-56.
- [29] EUnetHTA Joint Action 2. *HTA Core Model v.2*; 2013. <http://www.corehta.info/BrowseModel.aspx>, last access 11 February 2016.
- [30] J. Talmon, E. Ammenwerth, J. Brender, N. de Keizer, P. Nykänen, M. Rigby, STARE-HI - Statement on reporting of evaluation studies in Health Informatics, *International Journal of Medical Informatics* **78** (2009), 1-9.
- [31] S.P. Slight, Q. Casey, A.J. Avery, DW Bates, A. Sheikh, A qualitative study identifying the cost categories associated with electronic health record implementation in the UK, *Journal of the American Medical Informatics Association* **21** (2014), e226-e231.
- [32] K. Li, S. Naganawa, K. Wang, P. Li, K. Kato, X. Li, J. Zhang, K. Yamauchi, Study of the cost-benefit analysis of electronic medical record systems in general hospital in China, *Journal of Medical Systems* **36** (2012), 3283-3291.
- [33] S.J. Constantinos, A framework for the ex-ante evaluation of ERP software, *European Journal of Information Systems* **10** (2001), 204-215.
- [34] D. O'Reilly, J.E. Tarride, R. Goerre, C. Lokker, K.A. McKibbon, The economics of health information technology in medication management: a systematic review of economic evaluations, *JAMIA* **19** (2012), 423-438.
- [35] T.S. Bergmo, Can economic evaluation in telemedicine be trusted? A systematic review of the literature, *Cost Eff Resour Alloc* **7** (2009): 18.
- [36] J. Bassi, F. Lau, Measuring value for money: a scoping review on economic evaluation of health information systems, *JAMIA* **20** (2013), 792-801.
- [37] J. Polisena, D. Coyle, K. Coyle, S. McGill, Home telehealth for chronic disease management: a systematic review and an analysis of economic evaluations, *International Journal of Technology Assessment in Health Care* **25** (2009), 339-349.
- [38] M. Dávalos, M.T. French, A.E. Burdick, C. Scott, E. Simmons, Economic evaluation of telemedicine: review of the literature and research guidelines for benefit-cost analysis, *Telemedicine and e-Health* **15** (2009), 933-948.
- [39] H. Mistry, Systematic review of studies of the cost-effectiveness of telemedicine and telecare. Changes in the economic evidence over twenty years, *Journal of Telemedicine and Telecare* **18** (2012), 1-6.
- [40] D.J. Chisolm, A.S. McAlearney, S. Veneris, D. Fisher, M. Holtzlander, K.S. McCoy, The role of computerized order sets in pediatric inpatient asthma treatment, *Pediatr Allergy Immunol* **17** (2006), 199-206.
- [41] W.M. Tierney, M.E. Miller, J.M. Overhage, Physician inpatient order writing on microcomputer workstations: Effects on resource utilization, *JAMA* **269** (1993), 379-383.
- [42] H.S. Mekhjian, R.R. Kumar, L. Kuehn, T.D. Bentley, P. Teater, A. Thomas, B. Payne, A. Ahmad, Immediate benefits realized following implementation of physician order entry at an academic medical center, *JAMIA* **9** (2002), 529-539.
- [43] R.S. Evans, S.L. Pestotnik, D.C. Classen, A computer-assisted management program for antibiotics and other antiinfective agents, *N Engl J Med* **338** (1998), 232-238.
- [44] P.R. Orszag, Evidence on the costs and benefits of health information technology, *Testimony before Congress* **24** (2008).
- [45] S.J. Wang, B. Middleton, L.A. Prosser, C.G. Bardon, C.D. Spurr, P.J. Carchidi, A.F. Kittler, R.C. Goldszer, D.G. Fairchild, A.J. Sussman, G.J. Kuperman, D.W. Bates, A cost-benefit analysis of electronic medical records in primary care, *The American Journal of Medicine* **114** (2003), 397-403.

- [46] P. Tse, *Systematic review: the return on investment of EHR implementation and associated key factors leading to positive return-on-investment*, PhD Thesis, The University of Hong Kong (Pokfulam, Hong Kong), 2013.
- [47] C. Byrne, L.M. Mercincavage, E.C. Pan, A.G. Vincent, D.S. Johnston, B. Middleton, The value from investments in health information technology at the US Department of Veterans Affairs, *Health Affairs* **29** (2010), 629-638.
- [48] AHRQ (Agency for Healthcare Research and Quality), *Health Information Technology Evaluation Toolkit*, 2009, <https://healthit.ahrq.gov/sites/default/files/docs/page/health-information-technology-evaluation-toolkit-2009-update.pdf>, last access 11 February 2016.
- [49] R.C. Wu, A. Laporte, W.J. Ungar, Cost-effectiveness of an electronic medication ordering and administration system in reducing adverse drug events, *Journal of Evaluation in Clinical Practice* **13** (2007), 440-448.
- [50] T. Bertsche, J. Pfaff, P. Schiller, J. Kaltschmidt, M.G. Pruszydlo, W. Stremmel, I. Walter-Sack, W.E. Haefeli, J. Encke, Prevention of adverse drug reactions in intensive care patients by personal intervention based on an electronic clinical decision support system, *Intensive Care Medicine* **36** (2010), 665-672.
- [51] S.H. Forrester, Z. Hepp, J.A. Roth, H.S. Wirtz, E.B. Devine, Cost-effectiveness of a computerized provider order entry system in improving medication safety ambulatory care, *Value in Health* **17** (2014), 340-349.

Health Care Performance Indicators for Health Information Systems

Hannele HYPPÖNEN^{a,1}, Elettra RONCHI^b and Julia ADLER-MILSTEIN^c

^a*National Institute for Health and Welfare, Finland*

^b*The Organisation for Economic Co-operation and Development (OECD), France*

^c*University of Michigan, USA*

Abstract. Health Information Systems (HISs) are expected to have a positive impact on quality and efficiency of health care. Rapid investment in and diffusion of HISs has increased the importance of monitoring the adoption and impacts of them in order to learn from the initiatives, and to provide decision makers evidence on the role of HISs in improving health care. However, reliable and comparable data across initiatives in various countries are rarely available. A four-phase approach is used to compare different HIS indicator methodologies in order to move ahead in defining HIS indicators for monitoring effects of HIS on health care performance. Assessed approaches are strong on different aspects, which provide some opportunities for learning across them but also some challenges. As yet, all of the approaches do not define goals for monitoring formally. Most focus on health care structural and process indicators (HIS availability and intensity of use). However, many approaches are generic in description of HIS functionalities and context as well as their impact mechanisms on health care for HIS benchmarking. The conclusion is that, though structural and process indicators of HIS interventions are prerequisites for monitoring HIS impacts on health care outputs and outcomes, more explicit definition is needed of HIS contexts, goals, functionalities and their impact mechanisms in order to move towards common process and outcome indicators. A bottom-up-approach (participation of users) could improve development and use of context-sensitive HIS indicators.

Keywords: Quality indicators, health care, medical informatics applications, health information systems, eHealth, benchmarking.

1 Introduction

Pressures on health care systems across the world to ensure simultaneously access, quality, and affordable care are increasing with the aging population, increased demands for service equity and patient expectations, advances in medicine, and slow economic growth. Health care administrators and policy-makers are faced with major questions regarding the allocation of scarce health care resources to select interventions that support high performance of health systems and increase the quality and efficiency of care and services. [1, 2]

There are big expectations for health information or eHealth Systems (HISs) in improvement of health care system performance. In this contribution, terms HIS and

¹ Corresponding author: Dr. Hannele Hyppönen, National Institute for Health and Welfare (THL), P.O.Box 30 (Mannerheimintie 166), 00271 Helsinki, Finland, Hannele.hypponen@thl.fi.

eHealth are used as synonyms and defined according to the Medical Subject Headings (MeSH) as “*Integrated set of files, procedures, and equipment for the storage, manipulation, and retrieval of (patient) information*”. Health systems performance improvement is defined as “*positive changes in capacity, process and outcomes of public health as practiced in government, private and voluntary sector [health care] organizations*”. [3]

Adoption of HISs has grown substantially in the past years [1, 2, 4], including regional and national electronic health record (EHR) systems to capture patient health information and enable the exchange of patient information between organizations. [5] HISs have been seen as key enablers for modern, patient-centred and efficient health care services [6, 7]. Rapid technological diffusion has increased the importance of commonly agreed, reliable and valid indicators to monitor the adoption and impacts of HISs, to learn from past and current initiatives, and provide decision makers with evidence to make informed policy decisions about their HIS. Evidence-based management is a management approach adopting the ideas of evidence-based health care to management, emphasizing that common principles should apply to clinical and non-clinical investment decisions [8, 9]. An “Indicator” is “*a single summary measure, most often expressed in quantitative terms, representing a key dimension of health status, the health care system or related factors*” [10, 11]. “Monitoring” is “*a process aimed at measuring possible change in the indicator values over time (...) to provide (...) the main stakeholders of an intervention with early information on progress, or lack thereof, in the achievement of specific outcomes or objectives. (...) Monitoring must be periodic to pick up change*”. [7] For monitoring, impact mechanisms of interventions as well as mediating factors must be clearly specified [12, 13].

In defining what to measure and how for the purpose of evidence-based management of HIS, the viewpoints of different stakeholders (e.g., policy makers, administrators, researchers, citizens or others) have to be taken into account.² In addition to improving health, the wider goals of efficiency and equity of access to care are often included in health system performance monitoring. [14] The measurement of HIS-related improvements in health care system performance therefore requires taking into account these various dimensions. The need to translate these dimensions into concrete representations that can be quantified, and the need for a consensus about the most appropriate measures, complicate the definition. Many different measures of equity of access to care can, for example, be used (e.g. waiting time, availability of resources, access of costs), and some may be more sensitive to HIS than others. Measures need to be based on a sound, scientifically validated knowledge foundation, authority, or be derived from the practitioners [13] as well as have a plausible link to HIS.

2 Examples of existing health care performance indicators for HIS

An abundance of HIS indicator domains and measures has been defined for health care structural, process and outcome performance (e.g. [15]). Figure 1 presents a generic (not HIS-specific) representation of the relations between the health care structure, process

² See also: L. Lee et al., Understanding stakeholder interests and perspectives in evaluations of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

and outputs/outcomes [16]. Knowledge of structural and process elements is required if outcome impacts are to be understood and evaluated: if the structural elements – e.g. HIS functionalities – are not available, they cannot be used, and if they are not used, they cannot impact health care outputs or outcomes. In general, structural elements are measured by the care capacity using input and resource indicators. HISs and other tools and equipment used in care are one type of resource, availability of which thus form one structural indicator topic. Processes are monitored by indicators measuring care transactions (use of the capacity and resources), outputs by efficiency and volume of services provided, and outcomes by indicators measuring impacts of the care services.

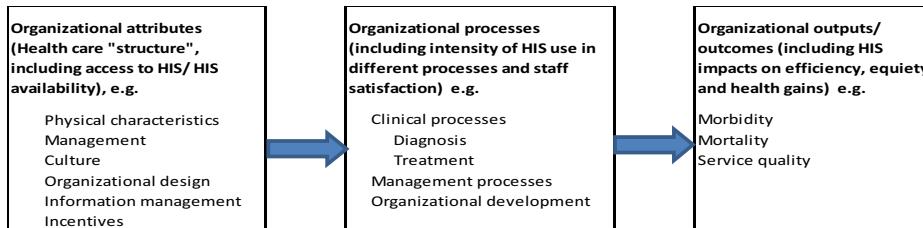


Figure 1. The relations between health care structure, process and outcome elements (modified from [16]).

Most of the existing HIS indicators focus on elements of health care structures (e.g. “HIS availability”) and processes (e.g. “HIS usage rate”), but also output and outcome indicators have been defined for some specific HIS applications or functionalities and for some health conditions. Many studies have demonstrated a positive association between HIS availability and health care outcomes, but commonly agreed outcome indicators are still rare. The HIS-outcome association is not without controversy, as the variance in outcomes can be attributed to a wide range of factors, for which there are often no ready measures. (We can directly measure EHR adoption, but many factors beyond EHR impact for example patient safety). This is another motivation for the current focus on the structural measures. There is, however, significant policy interest in evaluating the impacts of the implemented HISs and in particular changes in health system performance.

Table 1, while not intended as a comprehensive listing, depicts some examples of current HIS indicator domains and concrete HIS-related health care performance indicators as they apply to one functionality for medication management process.

Superficially, progress on developing measures for “a complete list of prescriptions made to patient” seems good as there are indicators for various aspects of health care performance, not just for structural performance. However, few of these indicators are *internationally agreed, and not all represent reliable and valid concrete measures for monitoring impacts of a complete prescription list*. There is little agreement on *the HIS functionalities that should be measured nor on the health care structural, process, output and outcome elements* that matter in order to determine the “value” of specific functionalities [7]. There is also heterogeneity in the methods used to collect data for monitoring. Main methods used are structured surveys, with no agreed definitions and scales. To understand the situation better and to move forward in the national and international HIS indicator work, we need to have a closer look at the different ways in which HIS indicators are developed and used.

Table 1. Examples of national level health care system performance indicators for a HIS functionality related to medication management.

Notes: a) Specific indicators need to be defined based on local eHealth policy and strategy priorities or stakeholder goals. b) Concrete measures need to be considered from different stakeholder viewpoints and defined for key HIS functionalities c) Many topics (e.g. usability, information quality) require more than one measure, d) There are currently many subjective measures and lack of objective data - only some register-based monitoring measures exist so far.

HIS-related health care structural performance indicators	HIS-related health care process performance indicators	HIS-related health care output indicators	HIS-related health care outcome indicators
HIS availability: Proportion (%) of public organisations where a list of prescriptions made to the patient outside own organization is available for professionals ^{1-2, (3)}	Intensity of HIS use: Proportion of viewings of prescriptions from outside own organisation by professionals/population size ^{1, (2, 3)}	Impacts on time saved: Potential for saved time with ideal system functionality ¹ ; time to take medication history/patient ³	Impacts on patient safety: No. of medication errors reported during a year/population of the country ¹ ; no. of adverse drug-related events for high risk patients / all high risk patients, for physicians with and without decision support ³
HIS technical quality: Experienced satisfaction of physicians with EHR reliability (Mean value, scale 1–5) ¹	HIS user satisfaction/attitudes: Overall satisfaction of physicians with the EHR system (scale 1–5) ^{1, 3}	Impacts on no. of contacts: No. of calls per day received by physicians for refills ³	Impacts on continuity of care: Physicians' experience on IS supporting collaboration between doctors working in different organizations (mean, scale 1–5) ¹
HIS impact on information quality: Proportion (%) of public organisations where nationally agreed information structures are available (implemented) ¹	Impacts on conformity to care guidelines: Physicians' experience of impact of HIS on conformity to care guidelines (mean, scale 1–5) ¹	Impact on Health care costs: Proportion of ICT-costs of the total budget in public organisations ¹	

Sources: 1) The Nordic eHealth Indicators [17]; 2) The OECD model survey [18]; 3) Canada Health Infoway [19].

3 Methodologies to develop and use Health Care Performance Indicators for HISs

The following sections describe some of the approaches used to develop and assess Health Care Performance Indicators for HIS. We focus on approaches taken at the country or multi-country level. We use the four phase indicator methodology depicted in environmental economics [20] as a common “standard” for comparison in the approaches to look for similarities and differences. The methodology was first adopted in HIS indicator work by the Nordic eHealth Network to make the indicator process transparent:

1. Defining the context (human and environmental) for measurement:
 - a. Identifying key stakeholders.
 - b. Defining the relevant area or system in question.
2. Defining the goals for measurement.
3. Defining methods for indicator selection and categorization.
4. Defining the data; collecting, analysing and getting feedback.

3.1 Adoption of the 4-phase methodology in the Nordic eHealth Network approach

In 2012, the eHealth group of the Nordic Council of Ministers established a network of organizations responsible for national eHealth monitoring in each of the Nordic countries, to define and test common Nordic eHealth indicators. These were required for monitoring eHealth in the Nordic countries, for use by national and international policy makers and scientific communities to support development of Nordic welfare. [21, 22] The work commenced by searching for a suitable methodology for defining eHealth indicators, and was done in close collaboration with the Nordic council of Ministers and the OECD model survey development.

A review of previous approaches for indicator development proved that a step-by-step methodology is rarely described in connection with eHealth indicator work. In the field of sustainable development, two main approaches for indicator definition have been identified [20]: Expert-led top-down and community-led bottom-up methodologies. Top-down methodology is used in indicator work that focuses on defining measures with which to monitor implementation of policies and their impact on the society level. Top-down approaches rarely define goals formally, as they are pre-determined by funding agencies or Government offices; also they may not reflect or record adequately the effect on the population served. The Bottom-up approach is used especially in the fields where the aim is to monitor policy or strategy implementation and their impacts on the micro level: the indicators are tailored to the needs and resources of the indicator users, but they still remain rooted firmly in the fundamental principles of the policy in question. The Top-down and Bottom-up approaches share four common phases. [20] These were used as a basis of eHealth indicator development in the Nordic Collaboration.

The first phase – *context definition* – calls for identification of the HIS functionalities for which indicators are needed, their contexts of use and users, and for whose viewpoint the indicators are developed. In the Nordic countries National eHealth policies were analysed to enable defining the contexts to be monitored [11]. Description of the functional architecture of HIS functionalities was found necessary especially for international comparison. Without this information it is impossible to say “which type of medicine cured the patient”. [17]

The second phase – *defining the goals* – includes identification of impact mechanisms of the selected systems in the contexts for different stakeholders, as well as the changes anticipated for health system performance. The eHealth policies in the Nordic countries provided also this information for the Nordic indicators [11, 22].

For the third phase – *indicator selection and categorization* – a longlist of available indicators from the Nordic countries was generated from existing monitoring studies, complemented with indicators from eHealth evaluation studies. The potential measures were grouped with a conceptual framework generated by Ammenwerth and de Keizer [15] under benefits for health care structural, process and outcome quality. To select and

prioritize the measures, the grouped longlist of variables was mapped against the policy goals, stakeholder priorities and the OECD model survey measures.

The fourth phase includes *defining the actual metrics and the data sources, testing the data collection and reporting the results*. For the Nordic work, this was done by comparison of questions in the existing surveys and the OECD model survey as well as national log and register data. The variables were tested by collecting and reporting the data for each of the defined variables [11, 22]. The lessons learned have been used to refine the indicators further as a basis for a permanent system for Nordic eHealth benchmarking. [17]

3.2 The Canadian approach

Canada Health Infoway is the primary lead in Canada for eHealth-related activities. In 2006 Canada Health Infoway published their Benefits Evaluation framework. A consulting team managed the Benefits Evaluation (BE) Plan development process and compiled the report. Subject matter experts developed the programme specific plans, and an Expert Advisory Panel provided guidance in the development of the BE Plan [19]. It included HIS-specific benefits indicators for six national eHealth programmes: diagnostic imaging, drug information system, laboratory information system, public health surveillance system, interoperable electronic health record and telehealth programmes. Selection of indicators was based on the following criteria:

- Importance: The indicator reflects aspects of health system functioning that matter to users and are linked to a Strategy Map priority area. Six strategy-relevant HIS functionalities are selected for monitoring: radiology, medication, laboratory, public health (immunizations), interoperable EHR (Health information exchange and personal health records), telehealth.
- Relevance: The indicator provides information that advances the understanding of population health and the health system, and can be used to monitor and measure health system performance over an extended period of time.
- Feasibility: Data required for the indicator are readily available for the areas and time periods indicated, and there are no unreasonable obstacles or constraints on access to the information collected, nor restrictions on its use.
- Reliability: The indicator produces consistent results in repeated measurements of the same condition or event.
- Validity: There is consensus on the part of users and experts that the indicator is related to the dimension it is supposed to assess (face validity), covers the whole dimension it is supposed to assess (content validity), is related to other indicators measuring the same dimension (construct validity), and has predictive power (criterion validity). [19]

The Delone & McLean Information System Success Model [23], which is based on wide review of eHealth literature, was used as a basis for conceptual grouping of the eHealth indicators. For each of the six selected HIS functionalities, separate indicators were developed.

Assessed against the 4-phase approach, the Canadian approach includes phases 1-2, even if the policy analysis is not published as part of the framework definition: the first and second bullet points anchor the definition of the HIS functionalities and goals to be monitored firmly to the national eHealth strategy. The three consequent bullet points

refer to phases 3-4. The Canadian approach uses both literature and analysis of availability of data to select the indicators, which focus on HIS-related impacts on all aspects of health care system performance, including outcomes. Most of the data are provided with individual evaluations, and indicators for continuous performance monitoring with survey, log or register data are less common [19].

3.3 The European Commission approach

An important contribution of the European Commission to improvement of public health services is benchmarking ICTs. Wide differences across health care systems at both national and regional level and the absence of commonly agreed indicators led the Commission to launch a series of eHealth benchmarking studies, e.g. [2, 4, 24, 25]. This research has also developed an indicator framework to support and guide the development of, and agreement on, a comprehensive set of key global indicators and procedures for data gathering on eHealth. [25].

The report on the methodology identifies three user groups: patients, health-service providers and payment institutions, and four types of applications: clinical information systems, telemedicine, home care and personalized health systems, integrated regional-national networks and systems and secondary use systems. The methodology included a search for eHealth monitoring and benchmarking activities in the EU, Iceland, Norway, Canada and the United States of America. The priority areas were identified using the European Union eHealth policy analysis reports. Four dimensions of indicators were defined [25]:

- Basis indicators, covering respondent demographics and basic ICT infrastructure.
- Activity-dependent indicators, covering eHealth-related health care activities.
- Attitude indicators, covering general and specific attitudes towards ICT as well as perception of ICT-related impacts.
- Horizontal issues, including IT investment, IT support, Data protection/security, Interoperability, IT skills.

Reference to several EU-level policy analysis documents is presented as sources for additional priority areas, including the eHealth action plan as a source for selecting the systems to be monitored. It is stated in the document that policies have also been used for restructuring the activity dimension based on the expertise and understanding of the eHealth domain developed in a number of projects. [25]

A matrix with stakeholders as rows and indicator categories as columns is presented in the methodology, filled from the pool of more than 4,400 indicators identified from analysis of national studies. For missing data, new indicators were generated. The European Union has conducted primary and specialized care surveys based on the defined methodology in 2009 and 2013.

Compared with the 4-phase approach, the Commission framework has identified the key stakeholders and functionalities or applications as is done in the first phase of the 4-phase approach. Also EU-level policy studies have been referred to in the methodology as is done in phase 2 of the 4-phase approach. How policy priorities (applications and goals) in each country have been mapped against the indicators to be selected remains unclear. The third phase, indicator selection and grouping, has been conducted using the existing studies. Data were collected with a survey instrument defined by the Commission projects.

3.4 The OECD approach

The OECD has led an effort to provide countries with reliable statistics to compare ICT development and policies in the health sector [6], to assist governments in understanding the barriers and incentives to ICT use and to realize the far-reaching economic and social benefits from their application. In 2010, an OECD survey of countries identified four core objectives for ICT implementation: Increase the quality and efficiency of care; reduce the operating costs of clinical services; reduce the administrative costs of running the health care system; and enable entirely new models of health care delivery [26].

In 2012, the OECD established a group of national and international experts representing seventeen OECD countries and four non-OECD countries to agree on a core set of survey indicators and an approach to measurement. Its work has been guided by three overarching principles. First, measures needed to respond to policy and information needs of countries along a continuum, starting from ICT availability, moving towards effective use, and ending with measuring outcomes and impact on population health. This helps in accommodating countries that are at different levels of maturity and progress towards achieving their e-health goals. For example, advanced countries are unlikely to devote substantial resources to collecting data on availability of ICTs if their policy needs are focused on effective use and better outcomes.

The second principle was to use the OECD “model survey” framework, which takes a staged approach in moving international measurement work forward. The model survey is composed of separate, self-contained modules for flexibility and adaptability to a rapidly changing context. Core modules can be added-on to existing national surveys or administered as a stand-alone survey while supplemental modules can be used as needed by countries.

The third principle was to use a functionality-based approach to defining key types of health ICTs to ensure that the terminology has comparable meaning across different countries. For example, while many OECD countries use the terms electronic medical record (EMR) and electronic health record (EHR) interchangeably, in Canada, EMRs refer to systems used by a healthcare professional to manage patient health information in a specific medical setting, whereas the EHR involves pooling data from multiple different clinical settings, allowing access to a more comprehensive patient record. If a core module question asked physicians about EHR use, the answers from Canada and the U.S. would, for example, mean very different things. This approach also supports technology-neutrality (i.e., the questions neither require nor assume a particular technology) and is forward looking (i.e., does not hinder the use or development of technologies in the future).

The model questionnaire was completed and published in 2013 [18] and is structured as shown in Table 2. Part I of the survey is addressed to general/primary care/family practitioners in ambulatory settings, Part II, to Chief Information Officers and administrators in the acute care settings.

Since 2013, several countries have begun piloting the model survey and/or mapping information from existing surveys and administrative data sources to indicators that would be derived from the model survey.

Assessed against the 4-step approach, phase 1 is strongly present in specification of the functionalities to be monitored. Phase 2 – how selected indicators reflect the national eHealth goals – is not explicitly reported, but becomes evident in national selection of variables to be monitored. Phases 3 and 4, where actual indicators are selected, defined and tested, are conducted nationally. Data are collected by the member states in their

national HIS surveys, where the OECD model survey variables are integrated to the extent possible.

Table 2. Structure of the OECD Model Survey.

Part I GENERAL PRACTITIONERS/ PRIMARY CARE PHYSICIANS	Part II CHIEF INFORMATION OFFICERS/ IT ADMINISTRATORS
Section A: Contextual Indicators (i.e., basic demographic data about respondents and their practice setting)	Section A: Contextual indicators (i.e., basic demographic data about respondents and their organisation)
Section B: Availability and use of electronic records and health information exchange	Section B: Availability and use of electronic records and health information exchange
Section C: Availability and use of functionalities that support patient engagement	Section C: Availability and use of functionalities that support patient engagement
Section D: Availability and use of telecommunications technologies to support health care delivery	Section D: Availability and use of telecommunications technologies to support health care delivery

3.5 The World Health Organization (WHO) framework

The eHealth strategy for the World Health Organization (WHO) was established in 2005. The WHA58.28 resolution urged Member States to plan for appropriate eHealth services in their countries. That same year WHO launched the Global Observatory for eHealth, an initiative dedicated to the study of eHealth, its evolution and impact on health in individual countries. [7] The WHO Global Observatory eHealth survey has been conducted from 2005 between 4-year intervals for three times. The latest report is a survey-based baseline review of eHealth and innovation focusing on the first four recommendations of the Commission in Information and Accountability for Women's and Children's Health (CoIA). The survey instrument was developed to monitor attainment of the CoIA recommendations. Of particular significance to the survey was Recommendation 3 on eHealth and innovation: "by 2015, all countries have integrated the use of Information and Communication Technologies in their national health information systems and health infrastructure".

The survey instrument enquires about eHealth programmes for monitoring of women's and children's health for 1) health service delivery (call centres, education, reminders, health promotion, feedback, telemedicine); 2) health and health problems monitoring and surveillance; 3) access to information for health professionals (publications, decision support systems, patient records) and 4) other eHealth programmes. There are also questions about eHealth implementation barriers, knowledge base for eHealth, internet safety, social media etc. [27]

Compared to the 4-phase approach, the WHO framework is very strong in phases 1 - 2. The survey instrument queries about availability of national eHealth policies, and lists various eHealth programmes. The selection of functionalities in the list originates from current programmes in the countries (phase 1) [27]. eHealth policy or literature analysis for identification of applications and functionalities and their impact mechanisms to match the CoIA recommendations is not reported in the documents. It is also not clearly documented how the indicators have been selected. The main data source was a survey.

3.6 The ISO health informatics and health indicators conceptual framework

The International Standardization Organisation's (ISO) standard on Health Informatics and Health Indicators Conceptual Framework (21667:2010) is intended to foster a common vocabulary and conceptual definitions for the resultant framework. The framework defines the dimensions and sub-dimensions required to describe the health of the population and performance of a health care system. The conceptual framework is broad (high-level) to accommodate a variety of health care systems, and it encapsulates all of the factors related to health outcomes and health system performance and utilization, as well as regional and national variations.

ISO 21667:2010 does not identify or describe individual indicators or specific data elements for the health indicators conceptual framework; nor does it address needs analysis, demand analysis or the range of activities that need to be supported for health system management. The framework identifies four categories or indicator domains: 1) health status (the overall health of the population served, how it compares to other regions in the jurisdiction and how it is changing over time); 2) non-medical determinants of health; 3) health system performance (the health services received by the region's residents); 4) and community and health system characteristics (characteristics of the community and the health system that provide useful contextual information). These categories have been adopted by many countries (including Canada) and have sub-domains with associated indicators.

Compared to the 4-phase model, the ISO framework focuses in phase 3, definition of the actual indicators and measures. The origins for contextual definitions (stakeholders and systems) as well as groundings to the national eHealth policy goals are not explicitly described in the framework.

4 Utilising HIS indicators

Continuous measuring of progress in HIS success in a comparable manner supports evidence-based management in order to promote successful implementation of HIS, policy learning, decision making and the on-going policy processes. It provides benchmark information of actual HIS adoption, its progress and eventually impacts in different contexts [4]. Figure 2 presents an example of the varied emphasis on eHealth policy goals in the individual Nordic countries (2010) [21], and variation in two indicators: 1) adoption rate of one key functionality in 2010-2014: proportion of public health care organisations where sending a prescription electronically to be dispensed in any pharmacy is available, and 2) number of yearly reported medication errors per resident population (2014) [17].

There are several observations that can be made from Figure 2: 1) Denmark, Sweden and Norway have all had clinical infrastructure and business support as a high policy priority (Icelandic eHealth policy was not assessed in 2010). 2) Denmark and Iceland have progressed most rapidly in deployment of one indicator impacting business processes. 3) Finland and Norway have had more focus on IT architecture, security and standards, which may explain the slower start in adoption of the functionality. 4) Medication errors can be used as one indicator measuring ePrescribing success, but we would need to have data from a longer monitoring period to show change that has happened in this indicator after implementation of the HIS functionality. Still, we could show correlation and odds ratio at most, since there may be several other parallel

interventions impacting medication errors, statistics may be defined differently in different countries and also the architectures and use settings of the functionality vary from country to country.

Combining selected health care performance indicators e.g. medication errors, with information on availability and use of specific policy-relevant HIS-interventions could thus provide one source of information that is on high demand by the decision-makers about efficiency and effects of eHealth policies and particular HIS interventions, but as the case above shows, results are far from straight-forward.

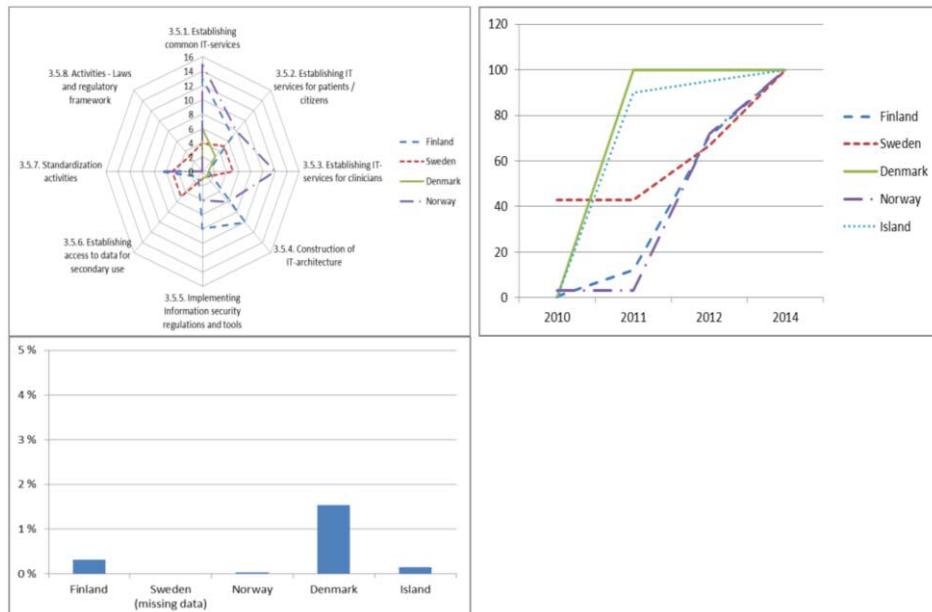


Figure 2. Examples of Benchmarking Indicators in use in the Nordic Countries: (a) Varied emphasis in eHealth policy goals in the Nordic countries in 2010 [21], (b) varied adoption rate of national ePrescribing 2010-2014 and (c) number of medication errors per resident population in 2014 [17].

5 Discussion

This contribution provides a summary overview of progress in HIS measurement by countries or multi-country coalitions – not individual health systems. The evidence presented indicates good progress in the development of internationally comparable HIS-related indicators for health care structures. A range of different approaches are in use - which creates opportunities for fruitful comparative assessment and learning. The Top-down approach appears to prevail, and goals for monitoring are rarely defined formally, as Reed et al. also found [20]. The WHO, Canadian and the Nordic approaches are among the most explicit in grounding the key HIS functionalities for HIS indicators in national health or eHealth policy goals. HIS-related indicators are targeted to policy makers, HIS managers and other stakeholders to inform decision-making related to HIS and their further development. Hence, the quality of the indicators can also be assessed against their utility for different stakeholders. Transparency of the goals and involving

stakeholders more closely in indicator development (the Bottom-up approach) is essential if the aim is to deliver information useful at different levels of decision making.

Few (international) approaches seem concerned about the comparability of ways HIS functionalities are realized across different countries and contexts. Yet, the way the HIS is structured and functions may have a strong impact in HIS usage rate, usability and thereby outcomes. Different countries are also in different stages of HIS implementation, and process and outcome measurements may vary depending on the maturity of the system, showing the importance of adequate definition, even “maturity index” of the HISs to be monitored.

The approaches differ also in indicator selection: the WHO approach is outcome- and condition-oriented, the ISO-framework is outcome- but not condition-oriented. The national Canadian approach has the widest scope with HIS-related indicators for health care inputs, processes, outputs and outcomes for six specified HIS functionalities. The OECD approach focuses on structural and process indicators. Methods of indicator selection and clustering are seldom explicitly stated.

Many methodologies (e.g. the EU, OECD and WHO) rely solely on data collection through stand-alone surveys, the Nordic and Canadian approach use also other types of data sources. The OECD and the Nordic approach use survey data collected as part of the national monitoring activities, the EU and WHO collect data themselves. The latter may be more costly but may result in more consistent and timely data collections across countries, whereas the former is a more economical option, but inevitably depends on national monitoring priorities and timelines. Moreover, achieving harmonization of the variables and data collection methodologies for cross-country benchmarking has been a challenge, which the OECD model survey and the ISO standard are helping to address. Also the organisation and funding of the national monitoring activities remain issues to be solved.

In conclusion, more explicit definition of systems, stakeholders and their goals, methods for indicator selection and categorization as well as stakeholder participation could help in moving towards stakeholder- and HIS-specific health care performance indicators for Health Information Systems that support evidence-based decision making on HIS approaches.

Recommended further readings

1. Hyppönen H, Kangas M, Reponen J, Nøhr C, Villumsen S, Koch S, et al., *Nordic eHealth Benchmarking. Status 2014*, Nordic Council of Ministers, Copenhagen, 2015, Report No. TemaNord 2015:539. <http://norden.diva-portal.org/smash/get/diva2:821230/FULLTEXT01.pdf>, last access 11 February 2016.
2. OECD, *Guide to Measuring ICTs in the Health Sector*, Organisation for Economic Co-operation and Development, COM/DELSA/DSTI(2013)3/FINAL. <http://www.oecd.org/els/health-systems/measuring-icts-in-the-health-sector.htm>, last access 11 February 2016.

Food for thought

1. In your country / in your health system, what indicators would be most important in determining the best focus for HIS investment?
2. Which terms in the following indicator can be defined in various ways in different countries, and what metrics are required to calculate the indicator value? “Proportion of public hospitals providing clinicians access to electronic storing of patient data”
3. What are the pros and cons of using the indicator “Time to take medication history per patient” for monitoring health care process improvement after implementing access to all prescriptions made to the patient from outside own organization?

References

- [1] J. Ryan, M. Doty, M. Abrams, *The Adoption and Use of Health Information Technology by Community Health Centers, 2009 - 2013*. The Commonwealth Fund Issue Brief 14, 2014, http://www.commonwealthfund.org/-/media/Files/Publications/Issue Brief/2014/May/1746_Ryan_adoption_use_hlt_IT_CHCs_rb.pdf, last access 11 February 2016.
- [2] C. Codagnone, F. Lupiáñez-Villanueva, *Benchmarking Deployment of eHealth among General Practitioners (2013)*, Final report, European Commission, 2014. <http://www.open-evidence.com/final-report-benchmarking-deployment-of-ehealth-among-general-practitioners>, last access 11 February 2016.
- [3] *National Public Health Performance Standards (NPHPS) programme*, Atlanta, Georgia, US: Centres for Disease Control and Prevention, 2015, <http://www.cdc.gov/nphps/index.html>, last access 11 February 2016.
- [4] European Hospital Survey, *Benchmarking Deployment of eHealth Services* (2012-2013), Final report. Luxembourg: Publications Office of the European Union, 2014, <http://ec.europa.eu/digital-agenda/en/news/european-hospital-survey-benchmarking-deployment-ehealth-services-2012-2013>, last access 11 February 2016.
- [5] OECD. *Strengthening Health Information Infrastructure for Health Care Quality Governance. Good Practices, New Opportunities and Data Privacy Protection Challenges*. Organisation for Economic Co-operation and Development, 2013. OECD Health Policy Studies, <http://www.oecd-ilibrary.org/content/book/9789264193505-en>, last access 11 February 2016.
- [6] J. Adler-Milstein, E. Ronchi, G. Cohen, W.L. Pannella, A. Jha, Benchmarking Health IT among OECD countries: better data for better policy, *J Med Inform Assoc* **21**(1) (2014):1-6.
- [7] R.E. Scott, A Saeed, *Global eHealth – Measuring Outcomes: Why, What, and How*, A Report Commissioned by the World Health Organization’s Global Observatory for eHealth, Bellagio, Italy, Global Observatory for eHealth, 2008.
- [8] K. Walshe, D.G. Rundall, Evidence-based Management: from theory to practice in Health Care, *The Millbank Quarterly* **79**(3) (2001), 429-457.
- [9] R. Pawson, T. Greenhalgh, G. Harvey, K. Walshe, Realist review – a new method of systematic review designed for complex policy interventions, *Journal of Health Services and Research Policy* **10**(suppl 1) (2005), 21-34.
- [10] *Health informatics - Health indicators conceptual framework*, ISO 216670:2010, International Standardisation Organisation, 2010.
- [11] H. Hyppönen, E. Ammenwerth, N. de Keizer, Exploring a methodology for eHealth indicator development, *Stud Health Technol Inform* **180** (2012), 338-42.
- [12] A. Donabedian, The Quality of Care How Can It Be Assessed? *JAMA* **260** (1988), 1743-1748.
- [13] L. Keele, D. Tingley, T. Yamamoto, Identifying mechanisms behind Policy Interventions via Causal Mediation Analysis, *Journal of Policy Analysis and Management* **34**(3) (2015), 1-27.
- [14] O. Arah, G. Westert, J. Hurst, S. Klazinga, A conceptual framework for the OECD Health Care Quality Indicators Project, *International Journal for Quality in Health Care* **18** (Suppl 1) (2016) 5–13.
- [15] E. Ammenwerth, N. de Keizer, An inventory of evaluation studies of information technology in health care: Trends in evaluation research 1982 - 2002, *Methods of Information in Medicine* **44** (2005), 44-56.
- [16] S. Glickman, K. Baggett, C. Krubert, E. Peterson, K. Schulman, Promoting Quality: The Health-Care Organization from a Management Perspective, *Int J Qual Health Care* **19**(6) (2007), 341-348.

- [17] H. Hyppönen, M. Kangas, J. Reponen, C. Nøhr, S. Villumsen, S. Koch, G.A. Hardardottir, H. Gilstad, L. Jerlvall, T. Pehrsson, A. Faxvaag, H. Andreassen, B. Brattheim, V. Vimarlund and J. Kaipio, *Nordic eHealth Benchmarking - Status 2014*, TemaNord **539** (2015). Nordic Council of Ministers, Copenhagen, 2015.
- [18] OECD: *Guide to Measuring ICTs in the Health Sector*, Organisation for Economic Co-operation and Development, COM/DELSA/DSTI(2013)3/FINAL. <http://www.oecd.org/els/health-systems/measuring-icts-in-the-health-sector.htm>, last access 11 February 2016.
- [19] Canada Health Infoway, *Benefits Evaluation Indicators Technical Report*, Toronto, Canada, 2006.
- [20] M.S. Reed, E.D.G. Fraser, A.J. Dougill, An adaptive learning process for developing and applying sustainability indicators with local communities, *Ecological Economics* **59** (2006), 406–418.
- [21] H. Hyppönen, A. Faxvaag, H. Gilstad, G.A. Hardardottir, L. Jerlvall, M. Kangas, S. Koch, C. Nøhr, T. Pehrsson, J. Reponen, A. Walldius, V. Vimarlund, Nordic eHealth Indicators: Organization of Research, First Results and Plan for the Future, *Stud Health Technol Inform* **192** (2013), 273-7.
- [22] H. Hyppönen, A. Faxvaag, H. Gilstad, G.A. Hardardottir, L. Jerlvall, M. Kangas, S. Koch, C. Nøhr, T. Pehrsson, J. Reponen, A. Walldius, V. Vimarlund, *Nordic eHealth Indicators: Organisation of research, first results and plan for the future*, TemaNord, Nordic Council of Ministers, Oslo, 2013.
- [23] W. DeLone and E. McLean, The DeLone and McLean Model of Information Systems Success: a ten-year update, *Journal of Management Information Systems* **19**(4) (2003), 9-30.
- [24] European Hospital Survey: *Benchmarking Deployment of e-Health Services (2012-2013)*, Final Report, European Commission, PwC, Joint Research Center, Publications Office of the European Union, Luxembourg, 2014.
- [25] I. Meyer, T. Hüsing, M. Didero, W. Korte, *eHealth Benchmarking - Final Report*, European Commission, Bonn, 2009, http://www.ehealth-benchmarking.eu/fileadmin/ehealth-benchmarking/results/documents/ehealthbenchmarking_final-report_2009.pdf, last access 11 February 2016.
- [26] OECD, *Improving Health Sector Efficiency, The Role of Information and Communication Technologies*. OECD Health Policy Studies, Paris, 2010.
- [27] WHO, Global Observatory for eHealth, “*eHealth for women’s and children’s health*” 2013 survey, World Health Organisation, Geneva, 2013.

Evaluating the Impact of Health IT on Medication Safety

Hanna M. SEIDLING^{a,1} and David W. BATES^b

^aDepartment of Clinical Pharmacology and Pharmacoepidemiology, Cooperation Unit Clinical Pharmacy, University of Heidelberg, Heidelberg, Germany

^bDepartment of Medicine, Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, and Harvard Medical School, Boston MA, USA

Abstract. Health IT is becoming an increasingly powerful tool for improving medication safety. While errors may happen at all stages of the medication process, different tools have been developed to support the prescribing process (e.g. computerized prescribing with decision support), the dispensing process (e.g. barcoding or automated dispensing and unit-dose systems), or the administration process (e.g. electronic medication administration records and smart pumps). Health IT can reduce medication error and preventable adverse drug event rates by increasing documentation quality and transparency, enhancing accuracy and correctness of the medication process, and supporting information exchange and interlinking different stages of the medication process. Typical evaluated endpoints comprise process-related outcomes such as number of medication errors, harm-related outcomes such as adverse drug events, or cost-related outcomes. Typical study design to measure effectiveness of health IT in medication safety comprises before-after studies and randomized controlled trials. However, implementation is challenging; it often has a major impact on the overall workflow and such technologies must be carefully introduced and their effects must be closely monitored in order to achieve the desired reductions, as in addition to preventing errors they nearly always introduce new ones. As complex interventions, their impact depends crucially on the real world setting and the implementation details and thus, transferability of study results is variable.

Keywords. Medication safety, medication error, computerized physician order entry, clinical decision support, complex intervention.

1. What is medication safety?

Medication safety can be defined as the attempt to safeguard the medication process ensuring that the risk for medication errors is minimized [1]. One definition of a medication error which has been widely used in research is that they are errors “in the process of ordering, dispensing, or administering a medication, regardless of whether an injury occurred or whether the potential for injury was present.” [2].

Every sub-step of the medication process is error-prone and errors may happen at all stages, though they are much more common at some stages than others. Most errors do not result in patient harm because errors especially during early stages of the medication process can be caught and corrected (i.e. near misses) and even errors that reach the patient may not necessarily result in actual patient harm. While the risk of

¹ Corresponding author: Dr. Hanna Seidling, Department of Clinical Pharmacology and Pharmacoepidemiology, Cooperation Unit Clinical Pharmacy, University of Heidelberg, INF 410, 69120 Heidelberg, Germany, Hanna.seidling@med.uni-heidelberg.de.

whether an error reaching a patient results in harm depends predominantly on the (dose-dependent) toxicity of a drug, evidence regarding which errors are likely to harm the patient is scarce.

Thus, medication safety describes a safety net of routine drug prescription and treatment, ideally in which well-trained personnel or responsible patients handle medical products which were designed to prevent faulty administration. To safeguard their actions, processes are optimized to minimize human errors, reduce information loss and anticipate future challenges in an intended treatment course.

However, today's routine drug treatment does not always meet these expectations, and therefore, errors arise and some of them result in harm. The risk of error at each sub-step depends on the complexity of the respective process and is therefore particularly high during the prescription process when the provider must consider the patient's history, his current clinical situation and the risk-benefit ratio of the intended treatment [3].

2. In what ways can health IT influence medication safety?

Health IT in the context of medication safety may support an individual sub-step of the medication process as well as their interlinkage (Figure 1).

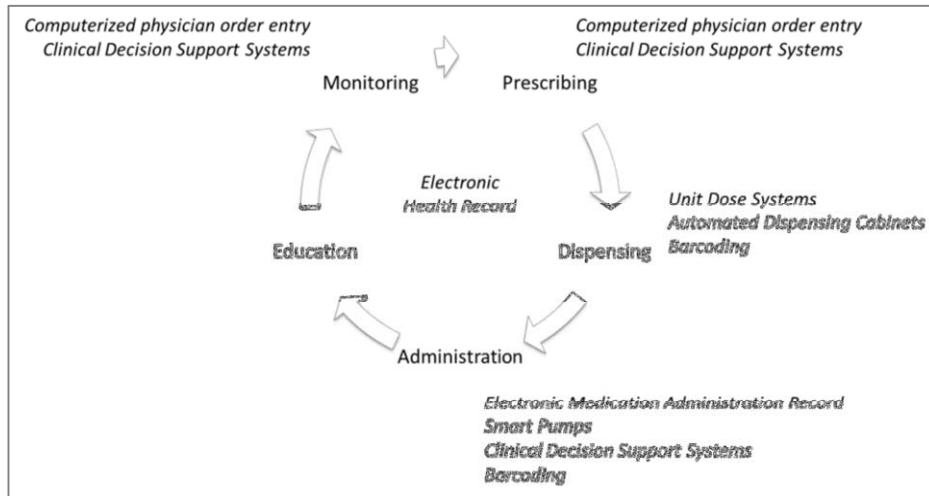


Figure 1. Display of health IT solutions along the medication process.

In general, health IT has the potential to (1) increase documentation quality and transparency including structure, standardization, readability and retrievability of information, (2) increase accuracy and correctness of clinical decisions or single tasks, and (3) improve information exchange and interlinkage of single sub-steps of the medication process.

2.1. Increase documentation quality and transparency

Compared to handwritten documentation, health IT can increase the process safety and documentation quality throughout the entire medication process. The most prominent examples of such health IT solutions are electronic prescribing systems (computerized physician order entry, CPOE). These systems offer the possibility to chart prescriptions and indicate dosage schemes. Thus, their benefit strongly depends on their design and usability. If the CPOE is basically a typing machine allowing free-text entries only, readability of orders will be increased in comparison to handwritten prescriptions, but prescriptions will not necessarily be more accurate. On the other hand, if the CPOE provides a catalogue referencing the prescribable drugs including their characteristics such as dosage forms and strengths etc., prescriptions can be more easily structured and prepared for basic plausibility checks, and if a default dose is suggested based on the patient's characteristics such as age and level of renal function that adds substantial additional benefit.

CPOE systems typically offer the possibility to pre-enter order templates or order sets, enabling standardization of prescriptions. There are a number of (national) recommendations for which functionalities CPOE systems should have [4-6]. Often CPOE systems are linked to a medication administration record (eMAR) which translates the provider's order into a request for administration. Thus, in the eMAR, nurses can seamlessly document whenever a drug was actually administered and thereby eliminate transcription errors [7].

2.2. Increase accuracy and correctness of the medication process

Health IT can increase accuracy and correctness of the medication process by redefining processes prone to human errors. For instance, during the drug distribution process in hospitals, drugs are typically ordered in the hospital pharmacy, packaged, sent to the ward, stocked on the ward and the dispensed to the patient. During each of those steps, confusions or look-alike errors may happen. The introduction of consequent barcoding [8] or automated dispensing [9] as well as unit-dose dispensing [10] can reduce these errors by automatizing the single steps and reducing interfaces.

In addition, many errors particularly during drug prescribing result from a lack of knowledge or information at the time of decision making. These errors are harder to address than for instance dispensing errors, because prescribing is typically the first step in the treatment process and not referring to executing a planned action. To reduce such errors, clinical decision support systems (CDSS) have been developed. Typically, CDSS are linked with CPOE and include a knowledge-base including the respective prescription-related information, an algorithm that links the prescription-related information with the actual information on the clinical context and the clinical case and a graphical user interface to display the resulting advice [11].

Depending on their scope, CDSS may support the selection (considering contraindications), the dosage (considering indication and patient characteristics) or the combination (considering drug-drug interactions) of drugs. CDSS can either lead providers in the correct direction, or redirect them using warnings. In contrast to health IT supporting the dispensing process, CDSS will only be effective if the provider considers the displayed information and changes his behavior. In many systems, as many as 95% of displayed warnings are neglected [12]. Thereby, the major challenges include the specificity of warnings and the integration in the workflow. Hence, we

know today that it is possible to refine the generation and display of warnings so that fewer are shown [13] and most are accepted [14]. Thus, while CDSS are clearly beneficial in certain systems and improve prescribing performance [15], they are insufficiently integrated in other systems and hence ignored, which leaves their actual impact on the overall healthcare marketplace unclear.

2.3. Increase information exchange and interlinkage of the medication process

A major challenge of medication safety is discontinuities in the medication process with changes of responsibilities, involved persons, and media. Thus, any health IT platform that enables seamless care might support medication safety. However, exchange of medication-related information is complex and can result in errors in dosage or route, for example. Many vendor applications do not routinely support reconciliation of medication lists from different electronic health records. Thus, a personal electronic health record supporting drug treatment throughout different health care sectors might be a way forward to seamless care.

A crucial prerequisite enhancing or limiting the effect of health IT is the fact how it is implemented into existing practice. Often, health IT influences existing workflows and forces the staff to potentially alter their routines. If the impact of health IT on existing workflows is not closely monitored and encountered difficulties solved this can both lead to workarounds (i.e. sometimes the health IT is not even used and hence cannot positively influence processes of care) or – which is even worse – can cause new iatrogenic errors, i.e. new errors that are actually caused by the health IT solution [17]. It is thus essential to prepare the implementation of health IT by depicting the existing workflows, assessing the potential influence of the planned health IT intervention and potentially adopting both existing workflows or the health IT solution before putting them together. It is also critical to monitor after the introduction of health IT for new errors, and to make changes that reduce their likelihood.

3. What are typical outcomes to measure effectiveness of health IT in medication safety?

Medication safety can be measured using several approaches, depending on the stakeholder's perspective. Typically, the most frequent approach is to assess process-related outcomes including the number of medication errors that occur. However, process-related outcomes are only a proxy for actual quality in care and indeed, not every medication error translates into actual patient harm. Hence, the rate of preventable adverse drug events or a number of higher level outcomes also assessing patient harm including (re)hospitalization and mortality can be used as measure for medication safety. Further approaches include impact on patient-related endpoints such as quality of life and patient satisfaction with care as well as cost-related endpoints that combine both savings resulting from prevented adverse events and spending on measures to improve medication safety.

Typically, the most preliminary endpoints applied for the assessment of health IT are those directly related to the purpose of the respective solution. For instance, if a clinical decision support system is designed to support the choice of a specific antibiotic treatment in the emergency department, the ratio of correctly chosen antibiotics before and after implementation could be assessed. However, these highly

specific outcome measures are of limited usefulness for comparison and to overall judge the benefit of any health IT solution. Hence, general outcome measures are applied and discussed in more detail:

3.1 Process-related outcomes

Medication errors are the most commonly used outcomes used to assess the effectiveness of health IT. The number of overall medication errors as well as of predefined subgroups (e.g. drug prescribing errors, or drug dosing errors) is generally assessed, typically as a ratio that gives some sense of overall potential for errors, i.e. for example patient-days or the overall number of drug prescriptions. The definitions used for medication errors in different studies vary [16], making it difficult to compare study results when studies use different definitions. Nevertheless, the impact on the medication error rate has been assessed for the majority of health IT solutions for medication safety [18].

3.2. Harm-related outcomes

Harm-related outcomes are frequently applied to estimate the potential benefit of medication safety strategies. Adverse drug events have been defined as “an injury resulting from medical intervention related to a drug” [19]. Thereby, a fraction of adverse drug events results from medication errors and is thus classified as preventable whereas inherited risks with a certain drug are classified as non-preventable adverse drug events. Only a minority of medication errors actually cause adverse drug events, with one estimate being one in 10 medication errors. [20]

More distal harm-related outcomes include (re)hospitalization and mortality, however, only a few studies have actually evaluated impact of health IT on these higher level outcomes and results are inconsistent [21]. Moreover, when assessing these higher level outcomes, it becomes more and more difficult to assess the influence of health IT, probably both because the events are infrequent and the health IT solution is just one intervention amongst many other influencing factors in a complex setting.

3.3. Cost-related outcomes

Cost-related outcomes include assessing the costs of adverse drug events, cost-minimization, cost-utility, cost-benefit and cost-effectiveness analysis.² Such assessments have been done for a minority of health IT solutions. However, for example, bar-coding in the pharmacy appears highly cost-effective [22]. Cost-effectiveness of CPOE potentially is modulated by the fact whether it is linked to a CDSS or not [23], and moreover, even if CPOE and CDSS might prevent adverse drug events and medication errors, hospitals might need to invest for this improvement of medication safety [24]. Since many health IT solutions are complex interventions that are implemented over a longer period of time and that might affect the medication process in several ways cost-related assessments remain challenging.

² See also: D. Luzzi et al., Economic evaluation of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

4. What are typical study designs to measure effectiveness of health IT in medication safety?

The assessment of effectiveness of health IT on medication safety generally falls into the category of quality improvement studies, so that study planning and reporting should consider the SQUIRE guidelines [25]. Typically quality improvement studies comprise complex interventions and therefore, meticulous descriptions of the setting, the intervention and the implementation are required to ensure a high study quality.³ This approach takes into account the fact that the medication process often is highly tailored in a specific setting which affects the generalizability of the results. Indeed, the success of a distinct quality improvement strategy is difficult to predict [26] and a quality improvement strategy proven successful in one setting might fail in another. The following section presents two typical study designs to measure effectiveness of health IT in medication safety.

4.1. *Before-after designs*

Given the uniqueness of a specific care setting and because many health IT interventions affect the medication process of an entire care setting, many studies are performed with a before-after design in the respective setting. This has the advantage of allowing the setting to serve as a control for itself, and the disadvantage that it is hard to assess the impact of other temporal considerations.

In a before-after study, baseline assessment is followed by an implementation period and a follow-up phase. Typically, data from the baseline assessment are then compared with the follow-up phase, however, there is no standardized rule on what the time span should be between baseline assessment and follow-up phase. Since the majority of health IT interventions also affect the processes and process changes are typically not easily implemented, the full benefit of the health IT intervention often becomes obvious only after a certain period of time. Indeed, immediately after implementation the risk of errors might even be higher, so that it is common to exclude that period, and only to conduct the “post” evaluation after stabilization in order to assess a net effect. However, particularly the phase during or immediately after implementation is crucial to assess the potential risk of health IT and its potency to introduce iatrogenic errors into the care process (on the risks of health IT, compare.⁴

While before-after designs allow for a very detailed look at a specific health IT intervention in a specific setting, the transferability of study results may be limited. Part of this restriction can be mitigated by the thorough description of the implementation and the intervention, however, the quality of healthcare over time might always be affected by other factors of influence than the implemented intervention.

4.2. *(Randomized) controlled designs*

To account for time effects and overall changes in a respective setting, (randomized) controlled designs can also be applied.³ Typically, the level on which the study is

³ See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁴ See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: *ibid.*

controlled depends on the level of the intervention. For instance, if a CPOE is introduced in an intensive care unit, a suitable control would be another similar intensive care unit. However, whether this control unit is an appropriate control depends on whether the two wards are indeed comparable with regard to baseline error rate, case mix and other factors. Hence, even in a controlled design, typically a baseline assessment is performed. In case of randomization, this baseline assessment can be used for a stratified randomization.

If the health IT intervention does not affect the overall medication process but rather supports a distinct sub-step (e.g. smart pumps or CDSS), randomization can also be performed on the individual patient level, however, in these cases, carry over effects are frequent, and typically cluster randomization is preferred.

5. What are pitfalls of today's methods to evaluate the impact of health IT on medication safety and how can they be overcome?

5.1. Real world settings

A major pitfall of today's methods to evaluate the impact of health IT on medication safety is the fact that most studies are typically performed in routine care, and hence processes are often not standardized. Indeed, the implementation of health IT often provokes the standardization and redesign of routine medication processes and hence it is not possible to separate the benefit of the health IT intervention from the additional benefits from the redesign of the medication process.

Moreover, health IT interventions are designed to support or improve a specific medication process and hence interventions might be deliberately adapted to a specific setting. While this approach might limit the comparability of several implementations of a distinct intervention [27], it will likely increase the success for a specific setting – which is, after all, the first and most urgent aim of the implementing institution. Indeed, the adaption rather than the unmodified adoption of interventions is a core element of quality improvement strategies. To account for resulting differences, the SQUIRE guidelines recommend describing in detail which adaptions were performed and for what reason.

5.2. Limited implementation details

In the past, most studies on health IT interventions often lacked implementation details, and for instance report on a “CDSS” that was introduced in an “intensive care unit” warning against potential “drug interactions”. Any result reported on the potential benefit of such system depends on how the CDSS is designed, what alerts it contained, how it was integrated into the routine care, when and how the alerts were displayed, how the provider was encouraged to interact with the system, etc. The simple description that such a system reduced the number of drug-drug interactions by half is hard to interpret, because it remains unclear how these results might apply to a different CDSS, a different drug-drug interaction database or a different setting and how reproducible they might be.

5.3. Limited comparability of studies

Indeed, the most common sentence in today's reviews trying to gather information on health IT intervention is probably the limited comparability of studies making meta-analysis difficult.⁵ However, to assess the impact of health IT on major endpoints such as hospitalization and mortality it will be essential to have larger datasets. One positive is that it is becoming increasingly easy to extract large quantities of data from electronic health records, and also to organize and share clinical decision support enabling very large implementations, so that it is likely to be possible to assess the impact of certain rule sets, for example, at scales that have not previously been possible.

One development which could be helpful would be to develop an adaptation of SQUIRE guidelines for specific health IT interventions, including some suggestions about which details on the health IT intervention or their implementation should be reported in order to allow for accounting for these details in meta-analysis. It will also be helpful to perform large-scale analyses across populations to get better assessment of the net impact of medication safety-related interventions on populations.

6. Case study

One early study which was a landmark in medication safety was a study that evaluated the impact of computerized physician order entry linked with clinical decision support on the serious medication error rate in two academic hospitals [15]. Units were divided into intervention and control and matched by patient type.

Key results were that the serious medication error rate fell by 55% in the intervention units, and that the decline occurred for all stages of the medication process. The preventable adverse drug event rate also fell 17%, but that decrease was not statistically significant. A team intervention was also evaluated, but that conferred no additional benefit over CPOE.

The generalizability of these results was uncertain, because the study was conducted in only two hospitals using an internally developed system, but many other studies have subsequently confirmed that the medication error rate falls with computerization of prescribing in the inpatient setting [28]. These results helped justify implementation of the HITECH Act in the U.S., which provided approximately \$30 billion in financial incentives to providers and hospitals which adopt health information technology and has resulted in broad adoption of electronic health records in both the inpatient and outpatient settings in the U.S. [29].

7. Conclusions

Health IT has now been shown to improve medication safety in a number of ways. It can have an impact at all major stages of the medication process in the hospital setting

⁵ See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

that are known to be error prone: prescribing – by structuring prescriptions and checking them for errors, dispensing – through bar-coding and automation of dispensing, and administration – through electronic medication administration records and smart pumps. The evidence for benefit is stronger for some of these stages than for others. Most studies have used process-related outcomes such as medication error rates, but some use harm-related outcomes such as adverse drug events, and a few studies have evaluated costs. The most frequent types of study design are before-after studies and randomized controlled trials. Implementation has a major effect on whether or not any particular intervention will be successful or not, and transferability has been variable. Any intervention can introduce or create new problems, and organizations as well as evaluators health IT should track these and attempt to minimize them.

Recommended further readings

1. E. Ammenwerth, P. Schnell-Inderst, C. Machan, U. Siebert, The effect of electronic prescribing on medication errors and adverse drug events: a systematic review, *J Am Med Inform Assoc* **15** (2008), 585-600.
2. E. G. Poon, C. A. Keohane, C. S. Yoon, M. Ditmore, A. Bane, O. Levzion-Korach, T. Moniz, J. M. Rothschild, A. B. Kachalia, J. Hayes, W. W. Churchill, S. Lipsitz, A. D. Whittemore, D.W. Bates, T. K. Gandhi, Effect of bar-code technology on the safety of medication administration, *N Engl J Med* **362** (2010), 1698–1707.
3. *SQUIRE Guidelines*, <http://www.squire-statement.org>. Last access: 11.2.2016.
4. P. G. Shekelle, P. J. Pronovost, R. M. Wachter, S. L. Taylor, S. M. Dy, R. Foy, S. Hempel, K. M. McDonald, J. Ovretveit, L. V. Rubenstein, A. S. Adams, P. B. Angood, D.W. Bates, L. Bickman, P. Carayon, L. Donaldson, N. Duan, D. O. Farley, T. Greenhalgh, J. Haughom, E. T. Lake, R. Lilford, K. N. Lohr, G. S. Meyer, M. R. Miller, D. V. Neuhauser, G. Ryan, S. Saint, K. G. Shojania, S. M. Shortell, D. P. Stevens, K. Walshe, Advancing the science of patient safety, *Ann Intern Med* **154** (2011), 693-696.

Food for thought

1. What health IT interventions do you think would most improve medication safety in your setting?
2. If you were designing a study to assess this health IT intervention, what design would you use?
3. What are the biggest risks related to medication safety in the main setting that you work in?

References

- [1] A.-F. Aly, Arzneimitteltherapie: Ein „Was ist Was“ der Sicherheit, *Dtsch Arztbl* **111** (2014), A-1892 / B-1618 / C-1550.
- [2] D.W. Bates, J. M. Teich, J. Lee, D. Seger G.J. Kuperman, N. Ma'Luf, D. Boyle, L. Leape, The impact of computerized physician order entry on medication error prevention, *J Am Med Inform Assoc* **6** (1999), 313-321.
- [3] H. M. Seidling, J. Kaltschmidt, E. Ammenwerth, W. E. Haefeli, Medication safety through e-health technology: can we close the gaps? *Br J Clin Pharmacol* **76** Suppl 1 (2013), i-iv.
- [4] M. Sweidan, M. Williamson, J. Reeve, K. Harvey, J. A. O'Neill, P. Schattner, T. Snowdon, Identification of features of electronic prescribing systems to support quality and safety in primary care using a modified Delphi process, *BMC Med Inform Decis Mak* **10** (2010), 21.
- [5] NHS, *Guidelines for Hazard Review of ePrescribing Systems*. http://webarchive.nationalarchives.gov.uk/20130502102046/http://www.connectingforhealth.nhs.uk/systemsandservices/eprescribing/hazard_framework.pdf, last access 11 February 2016.
- [6] NHS, *ePrescribing functional specification for NHS trusts*. <http://webarchive.nationalarchives.gov.uk/20130502102046/http://www.connectingforhealth.nhs.uk/systemsandservices/eprescribing/baselinefunctional.pdf>, last access 11 February 2016.
- [7] E. G. Poon, C. A. Keohane, C. S. Yoon, M. Ditmore, A. Bane, O. Levzion-Korach, T. Moniz, J. M. Rothschild, A. B. Kachalia, J. Hayes, W. W. Churchill, S. Lipsitz, A. D. Whittemore, D.W. Bates, T. K. Gandhi, Effect of bar-code technology on the safety of medication administration, *N Engl J Med* **362** (2010), 1698-1707.
- [8] E. G. Poon, J. L. Cina, W. Churchill, N. Patel, E. Featherstone, J. M. Rothschild, C. A. Keohane, A. D. Whittemore, D.W. Bates, T. K. Gandhi, Medication dispensing errors and potential adverse drug events before and after implementing bar code technology in the pharmacy, *Ann Intern Med* **145** (2006), 426-434.
- [9] K. N. Barker, R. E. Pearson, C. D. Hepler, W. E. Smith, C. A. Pappas, Effect of an automated bedside dispensing machine on medication errors, *Am J Hosp Pharm* **41** (1984), 1352-1358.
- [10] E. Cousein, J. Mareville, A. Lerooy, A. Caillau, J. Labreuche, D. Dambre, P. Odou, J. P. Bonte, F. Puisieux, B. Decaudin, P. Coupé, Effect of automated drug distribution systems on medication error rates in a short-stay geriatric unit, *J Eval Clin Pract* **20** (2014), 678-684.
- [11] H. M. Seidling, W. E. Haefeli, *Erhöhung der Medikationssicherheit im Krankenhaus durch elektronische Verordnungsunterstützung*, in: Burgard G, Baberg HAT, Popken G, ed., *Patientensicherheit. GemeinsamSicher*. Medizinisch Wissenschaftliche Verlagsgesellschaft Berlin; 2014.
- [12] H. van der Sijs, J. Aarts, A. Vulto, M. Berg, Overriding of drug safety alerts in computerized physician order entry, *J Am Med Inform Assoc* **13** (2006), 138-147.
- [13] H. M. Seidling, U. Klein, M. Schaier, D. Czock, D. Theile, M. G. Pruszydlo, J. Kaltschmidt, G. Mikus, W. E. Haefeli, What, if all alerts were specific - estimating the potential impact on drug interaction alert burden, *Int J Med Inform* **83** (2014), 285-291.
- [14] N. R. Shah, A. C. Seger, D. L. Seger, J. M. Fiskio, G. J. Kuperman, B. Blumenfeld, E. G. Recklet, D.W. Bates, T. K. Gandhi, Improving acceptance of computerized prescribing alerts in ambulatory care, *J Am Med Inform Assoc* **13** (2006), 5-11.
- [15] D.W. Bates, L. L. Leape, D. J. Cullen, N. Laird, L. A. Petersen, J. M. Teich, E. Burdick, M. Hickey, S. Kleefield, B. Shea, M. Vander Vliet, D. L. Seger, Effect of computerized physician order entry and a team intervention on prevention of serious medication errors, *JAMA* **280** (1998), 1311-1316.
- [16] M. Lisby, L. P. Nielsen, B. Brock, J. Mainz, How are medication errors defined? A systematic literature review of definitions and characteristics, *Int J Qual Health Care* **22** (2010), 507-518.
- [17] R. Koppel, J. P. Metlay, A. Cohen, B. Abaluck, A. R. Localio, S. E. Kimmel, B. L. Strom, Role of computerized physician order entry systems in facilitating medication errors, *JAMA* **293** (2005), 1197-1203
- [18] A. Agrawal, Medication errors: prevention using information technology systems, *Br J Clin Pharmacol* **37** (2009), 681-686.
- [19] D.W. Bates, D. J. Cullen, N. Laird, L. A. Petersen, S. D. Small, D. Servi, G. Laffel, B. J. Sweitzer, B. F. Shea, R. Hallisey, Incidence of adverse drug events and potential adverse drug events. Implications for prevention, ADE Prevention Study Group, *JAMA* **274** (1995), 29-34.
- [20] A. Krähenbühl-Melcher, R. Schlienger, M. Lampert, M. Haschke, J. Drewe, S. Krähenbühl, Drug-related problems in hospitals: a review of the recent literature, *Drug Saf* **30** (2007), 379-407.
- [21] S. Eslami, N. F. de Keizer, A. Abu-Hanna, The impact of computerized physician medication order entry in hospitalized patients - a systematic review, *Int J Med Inform* **77** (2008), 365-376.

- [22] S. M. Maviglia, J. Y. Yoo, C. Franz, E. Featherstone, W. Churchill, D.W. Bates, T. K. Gandhi, E. G. Poon, Cost-benefit analysis of a hospital pharmacy bar code solution, *Arch Intern Med* **167** (2007), 788-794.
- [23] E. Zimlichman, C. A. Keohane, C. Franz, W. L. Everett, D. L. Seger, C. Yoon, A. A. Leung, B. Cadet, M. Coffey, N. E. Kaufman, D.W. Bates, Return on investment for vendor computerized physician order entry in four community hospitals: the importance of decision support, *Jt Comm J Qual Patient Saf* **39** (2013), 312-318.
- [24] K. M. Vermeulen, J. E. van Doormaal, R. J. Zaal, P. G. Mol, A. W. Lenderink, F. M. Haaijer-Ruskamp, J. G. Kosterink, P. M. van den Bemt, Cost-effectiveness of an electronic medication ordering system (CPOE/CDSS) in hospitalized patients, *Int J Med Inform* **83** (2014), 572-580.
- [25] *SQUIRE Guidelines*, <http://www.squire-statement.org>, last access 11 February 2016.
- [26] P. G. Shekelle, P. J. Pronovost, R. M. Wachter, S. L. Taylor, S. M. Dy, R. Foy, S. Hempel, K. M. McDonald, J. Ovretveit, L. V. Rubenstein, A. S. Adams, P. B. Angood, D.W. Bates, L. Bickman, P. Carayon, L. Donaldson, N. Duan, D. O. Farley, T. Greenhalgh, J. Haughom, E. T. Lake, R. Lilford, K. N. Lohr, G. S. Meyer, M. R. Miller, D. V. Neuhauser, G. Ryan, S. Saint, K. G. Shojania, S. M. Shortell, D. P. Stevens, K. Walshe, Advancing the science of patient safety, *Ann Intern Med* **154** (2011), 693-696.
- [27] J. Metzger, E. Welebob, D.W. Bates, S. Lipsitz, D. C. Classen, Mixed results in the safety performance of computerized physician order entry, *Health Aff (Millwood)* **29** (2010), 655-663.
- [28] T. A. Shamsiyan, S. Duval, J. Du, K. L. Kane, Just what the doctor ordered. Review of the evidence of the impact of computerized physician order entry system on medication errors, *Health Serv Res* **43** (2008), 32-53.
- [29] ONC. *Adoption of Computerized Physician Order Entry in the US*. <http://dashboard.healthit.gov/index.php>, last access 11 February 2016.

Evaluation of Implementation of Health IT

Kathrin CRESSWELL^{a,1}

^a *eHealth Research Group, Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, UK*

Abstract. Information systems can only reach their full potential if their implementation is effective, and there is much to be learned as to what makes an “effective” implementation. In light of the substantial investments in Health Information Technology internationally, implementation evaluations are a powerful tool to ensure that technologies are enabled to fulfil their potential in improving care, reducing cost and increasing efficiency. The most salient characteristics of such evaluations are outlined, considering how they can help to assess adoption processes and outcomes through a continuous cycle of scientific enquiry and learning. A brief description surrounding potential theoretical lenses that may be drawn on is given. Issues discussed will be illustrated with the help of a case study on the implementation and adoption of Electronic Health Records in English hospitals. Practical challenges encountered and potential ways to address these during the conduct of health IT implementation evaluations illustrated include: 1) ways to cope with the shifting nature of reality (e.g. changes in local implementation strategies need to be reflected in the methods), 2) the need to examine processes as well as outcomes, 3) researching implementation in context with attention to both local processes and wider (e.g. political) developments, and 4) the pragmatic use of theoretical lenses where different approaches can shed light on different aspects of the implementation and adoption process.

Keywords. Health information technology, medical informatics, implementation, evaluation, evaluation research.

1. Introduction

Substantial investments into ever more complex Health Information Technologies (health IT) are currently actively being made in many countries across the world [1,2,3]. Despite significant monetary investments and some noteworthy implementation progress, particularly in relation to health IT infrastructures and bespoke technologies for discreet areas of care, a lot still remains to be desired [4-6]. Underlying difficulties tend to relate not only to technical issues, but also stem from the social consequences of system implementation. These include difficulties of integrating new technology with existing working practices of users and with organisational processes, potentially impacting negatively on patient safety [4-6]. Consequently, whatever the system, its use and thus its impact are only as good as its fit to the workplace.

¹ Corresponding author: Dr. Kathrin Cresswell, Centre for Population Health Sciences, The University of Edinburgh, Teviot Place, EH8 9AG, Edinburgh, UK, kathrin.beyer@ed.ac.uk.

Health IT implementation evaluations are a powerful tool to increase understanding of what makes for “successful” implementation, so as to optimise current and promote better future implementation experiences. Thus, implementation evaluations can help to ensure that technologies and their applications realize their maximum potential in improving care and work for all concerned, whilst keeping potentially adverse effects to a minimum [7]. Evaluation activities should therefore be at the heart of any health IT implementation. However, they are not straightforward activities that are easy to learn and apply, particularly when immediate implementation related pressures take priority. They require a certain amount of scientific ability and flexibility in order to be tailored to the environment in which they are conducted.

This contribution will outline some prevalent existing health IT implementation evaluation frameworks and methods. In doing so, it will re-visit some of the previously discussed aspects in context, particularly those surrounding theoretical basis², stakeholder perspectives³, and study design⁴. A case study outlining common challenges encountered during health IT implementation evaluation of a complex Electronic Health Record (EHR) system and potential ways to tackle these will conclude the contribution. It is hoped that this discussion will help readers to develop an appreciation of the necessary components of such activity, but also to learn to appreciate its complexity and application in context, so as to be able to apply it in their own setting.

2. What is health IT implementation evaluation?

To frame the discussion, it is important to first define what is meant by health IT implementation evaluation. Evaluation has been described as “*the process of determining merit, worth, or significance*” [8], but implementation evaluation focuses not only on overall assessments of the value of an intervention. It also involves examining processes surrounding implementation activities in order to investigate how the intervention produces its effects. By examining processes in context, it can also help to steer implementations and inform strategic decisions [9]. However, such activity is particularly difficult in the context of health IT implementation, because effects of IT are hard to trace and to attribute, resulting in issues in identifying the focus and defining the methods of the evaluation [10]. This challenge will be re-visited in Section 5 of this contribution.

Nevertheless, there are some crucial evaluation components that can help to focus health IT implementation evaluation activities. These include examining the following dimensions [11,12]:

- (1) Content: the characteristics of the technology that is implemented.
- (2) Context: the social, organizational, and wider environment in which the implementation takes place.
- (3) Process: how the technology is introduced and the implementation is conducted.
- (4) Outcomes: what has changed as a result of the implementation.

² See also: J. Brender, Theoretical basis of health IT evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

³ See also: L. Lee et al., Understanding stakeholder interests and perspectives in evaluations of health IT, in: *ibid.*

⁴ See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: *ibid.*

Health IT implementation evaluation should explore not only processes but also outcomes in context [11]. The former are often associated with qualitative and the latter with quantitative methods, making mixed-methods designs a popular choice.⁵ Insights obtained should be both formative and summative, where evaluation activities inform ongoing implementations (formative), and also give an overall *summary* of merit (summative) [10,13].

Ideally, these components are used together (Figure 1). This means that evaluation activities (both qualitative and quantitative) should be exploring technological change before, during and after implementations throughout the whole technology lifecycle - not just after an implementation, as is currently often the case [7]. *Before* measurements are important to establish a basis for comparison, whilst *during* measurements can help to assess changes resulting from immediate changes and during periods of upheaval. *After* measurements can help to assess medium- to longer-term consequences resulting from the implementation.

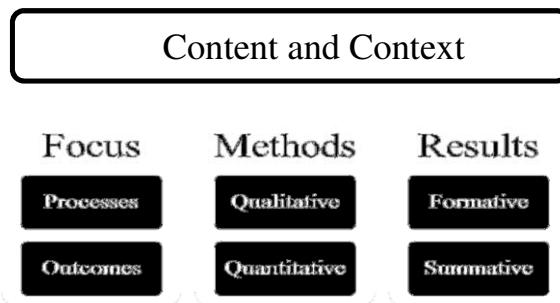


Figure 1. Essential components of health IT implementation evaluation.

3. Why is it so important to evaluate health IT implementations?

Implementation evaluations can help to measure and predict outcomes of health IT implementation, such as financial returns on investment and improvements in patient safety. They can also help to identify risks and unintended consequences, such as the inadvertent introduction of new threats to patient safety introduced by changes in work practices or technological features [4]. These insights can then be used to adjust implementation strategies, as is found necessary [7,9]. For instance, if users do not use technology as intended by management, there may be underlying issues with training or technological design that may warrant a more comprehensive user engagement strategy in the organization. Health IT implementation evaluations can also help to plan for future implementation strategies of other technologies in the same organization and other implementations in different organizations. For example, extracting lessons learnt can help to inform designing training plans and user engagement strategies, and

⁵ See also: P.J. Scott et al., Mixed methods: a paradigm for holistic evaluation of health information system, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

dissemination of experienced risks can help other implementers yet to introduce systems to mitigate for these in advance.

4. What theoretical lenses may usefully be drawn on to inform health IT implementation evaluation?

An in-depth discussion of the overall theoretical basis for health IT evaluations has been given otherwise.⁶ There are many theoretical lenses from various disciplines that may be drawn on to inform health IT implementation evaluations, and a comprehensive overview is beyond the scope of this contribution. On a basic level, existing theoretical stances draw on a variety of disciplines including, amongst others, areas of organizational change, information systems, psychology, human factors and informatics. Each existing approach attempts to explain different aspects of innovation implementation processes. Drawing on such theories is important as it can help to disentangle what processes can be expected to result in certain outcomes, thereby facilitating learning across implementations.

Theoretical lenses can broadly be divided into those that focus on technical and social “micro-factors” (e.g. those relating to technical characteristics and the humans using technologies), organizational aspects (e.g. change management, care delivery and business processes and implementation strategies), and wider “macro-environmental” factors (e.g. economic and political considerations). Amongst the most commonly used are sociotechnical perspectives [14].⁷ To reiterate, these focus on exploring both structural technological factors and associated social processes as well as the interrelationship between the two during implementation. They have been applied widely to understanding, for example, how technological change can result in changes to work practices of healthcare professionals, and vice versa, how users can shape technical designs. A related perspective, also focusing on the micro-environment surrounding the technology and related social factors, is Normalisation Process Theory [15]. This can help to investigate how health IT becomes routinely used over time and what inhibits and/or facilitates this process. For example, it is likely that technologies that are perceived to be associated with time savings for users are more readily “normalized” than those that are associated with increased workloads.

A wider organizational perspective is taken by the theory surrounding the Diffusion of Innovations [16]. This attempts to explain how health IT is adopted and how adoption spreads within and across organizations. This perspective can, for instance, help to explain how certain change management strategies surrounding user engagement can facilitate or inhibit adoption. Other perspectives draw on an even wider horizon, taking into account factors surrounding implementing organizations. For example, the Social Shaping of Technology can help to explain how health IT is shaped by historical, cultural and economic factors [17]; whilst work surrounding Information Infrastructures can help to explore how technologies are linked to support information sharing between care settings [18].

⁶ See also: J. Brender, Theoretical basis of health IT evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁷ See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: *ibid.*

Overall, it is probably best to draw on a variety of theoretical perspectives, depending on the focus and methods of the evaluation, as no one theory can explain the variety of factors involved in the adoption of complex health IT [12]. This issue will be explored in more depth in Section 5.4 below.

5. A case study: Evaluation of the introduction of EHRs in English hospitals

As both health IT implementations and hence evaluations of implementations vary significantly across contexts, it is best to illustrate with a concrete example the issues discussed above and associated challenges commonly encountered during evaluation activity. This will consist of an implementation evaluation case study surrounding the English quest to achieve “true” interoperability across the various settings of its nationally funded healthcare system through the large-scale procurement of EHRs (see Box 1 for the originally envisaged methods of the evaluation) [19].

To place the evaluation activities in context, it is necessary to discuss briefly the evaluation characteristics (see Section 2 above). The technology to be evaluated consisted of national EHRs that were to be implemented in hospitals across England. Anticipated functionality included sharing of electronic clinical data across different teams and settings. The implementation context was unusual in that hospitals were introducing technologies that were neither chosen, nor procured, by them but instead commissioned by the government through a national information technology strategy. Qualitative investigation of implementation processes revealed that this “top-down” procurement resulted in the rapid implementation of often immature technologies that were in many ways ill-suited to individual organizational needs. This in turn led to unanticipated outcomes, such as some new areas of risks to patient safety and financial threats to implementing organizations and system suppliers. Exploring implementation processes and contexts was an important aspect of the evaluation, as this could help to explain outcomes.

During the conduct of the work, evaluators liaised closely with individual organizations, feeding back how the new system was received by users and patients, but also with policy makers, in order to help inform strategic considerations and ongoing policy making (formative component). An overall summative statement was made, once evaluation activities had concluded. This ultimately led to a change in strategic direction from a “top-down” implementation model to increased involvement of local organizations in decision making [4]. Although longitudinal design was a defining feature of the work, restrictions in research funding meant that longer-term processes and outcomes (i.e. throughout the whole technology lifecycle) could not be assessed.

Throughout the work, evaluators encountered a number of conceptual and practical challenges which were not initially anticipated. Many of these stemmed from a lack of appreciation of the rapidly changing political, economic and organizational environment in which the evaluation was undertaken (Box 2) [20]. Although in many ways unique, challenges encountered may to some degree be transferable to other health IT implementation evaluations, as they reflect broader issues surrounding methodology and context (see Sections 2 to 4 above).

5.1. Coping with the shifting nature of reality

A challenge that pervaded all aspects of the work was the impact of shifting implementation landscapes and timelines, perhaps reflecting the on-going negotiations between relevant groups of stakeholders, resulting in the originally envisaged methodology (Box 1) having to be revised throughout the conduct of the evaluation [21]. Although the example was unique in relation to scale and ambition, this challenge is likely to be present in most health IT implementation evaluations, as implementation plans are rarely translated into reality without complications. For example, delays in implementation are commonly experienced. In the present case, these meant that the team was not, as originally anticipated, able to investigate changes over time. This, in turn, ensued in a lack of understanding surrounding the assessment of software systems once they had had an opportunity to embed within hospitals. Consequently, the insights were based on early implementation experiences only, which in the light of the relatively long time it takes to realize returns of investments of technologies for individuals and organizations, led to a lack of appreciation of longer-term impacts. These longer-term impacts are, however, particularly important in order to be able to make valid summative assessments on overall merit of an intervention (see Section 2).

Similarly, changes in implementation context are to be expected and require methodological flexibility. In the present example, political strategies changed throughout the conduct of the evaluation, influenced by implementation progress (which was overall slower than anticipated), an economic recession (which resulted in reduced funding for some functionalities), tensions and struggles within the central senior management teams (leading to a lack of integrated direction and leadership), and a change in government (which ultimately led to a modification of the overall strategy allowing increased local involvement in implementation approaches and software customization) [21,22]. Due to these changes in policy context, local strategies and technologies, the evaluation team had to adapt both the aims of the work and the methodologies originally envisaged [21,23]. In doing so, the focus shifted from the planned evaluation of embedded national systems, towards studying the various early local consequences and experiences of the implementation of changing systems with relatively limited functionality. Changes in context also meant that the evaluation team had to move from the original plan of assessing impacts with quantitative means (through measuring costs and safety indicators) towards studying processes using a predominantly qualitative case study-based approach [21, 24].

These experiences highlight the need for longitudinal evaluative work over extended periods of time in order to be able to assess the longer-term impact of embedded systems and allow evaluations to be adaptable to contextual changes. Naturally, such activity requires sufficient time and funding, which is often under-estimated. It is, for instance, recommended that program evaluations of innovative interventions should be allocated up to 10% of overall budgets [25], a figure which is far higher than what is currently commonly allocated to such activity. Admittedly, feasibility is an issue, given pressure on resources even in relation to financing full implementations, but considering evaluation budgets in advance is of primary importance as investments at this stage can help to prevent potential enduring operational inefficiencies.

5.2. Researching implementation and deployment in context

As we have seen in Sections 2 and 4, implementation context can be conceptualized ranging from micro factors (e.g. social context of immediate use), over meso-context (e.g. organizational environments), to macro-context (e.g. political and economic factors). These contextual factors can impact significantly on implementation processes and outcomes and vice versa. Exploration of such dimensions therefore needs to constitute an essential part of health IT implementation evaluation activity. In the example case study, for instance, broader political and commercial developments took a significant role in shaping local experiences [5,26]. Users of the nationally procured technology (which was designed to ensure coordinated efforts towards realizing data sharing across care settings and thereby maximize interoperability) experienced increased workloads due to a perceived lack of customizability of the application [5,26]. This, in turn, impacted on adoption rates as some users resisted use. Similarly, the exit of major commercial companies that had originally agreed to supply relevant applications not only impacted on profits, reputations and commercial relationships; but also resulted in some originally anticipated functionality not being implemented at all [27].

The importance of these contextual factors was not fully appreciated when the evaluation study was designed, and the emphasis of the work therefore had to be changed accordingly, with a greater focus on exploring the perspectives of a wider range of stakeholders than initially expected.⁸ Practically, this included shifting the focus of the evaluation from interviewing hospital staff towards consulting a greater number of political and commercial representatives to examine wider strategic decisions and underlying perceptions. In doing so, the evaluation team conducted additional anonymized interviews with individuals to gain an insight into personal experiences and examined a range of publicly available documents to understand historic developments. These included strategic reports, press releases, official statements, and media coverage.

5.3. Researching outcomes and consequences

Central to these efforts of placing evaluations within their appropriate contexts should also be a cyclical and iterative relationship between formative evaluation activities and strategic decisions (see Sections 2 and 3) [7]. Although summative statements about overall merit of an implementation are an essential component of health IT implementation evaluations, it is important to keep in mind that these can promote a somewhat unhelpful distinction between retrospective judgments of “success” and “failure” at a certain point in time. More applied and immediate usefulness of health IT implementation evaluation results can be promoted by fostering a mutually shaping formative relationship between strategic decisions and evaluative research, helping to promote pro-active decision making by identifying potential difficulties early and helping to devise risk mitigation strategies [28]. This emphasis on formative evaluation was, although mentioned (Box 1), not sufficiently appreciated when the evaluation study was designed, resulting in a lack of attention being paid to mechanisms surrounding feedback loops to individual organizations and policy makers during the conduct of the work.

⁸ See also: L. Lee et al., Understanding stakeholder interests and perspectives in evaluations of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

With the above in mind, researching of temporary (and some potentially permanent) consequences of and processes involved in technological change for a variety of stakeholders (as circumstances and effects may change over time and differ across professional groups) can be seen as one of the most important aspects of health IT implementation evaluation activities. An originally planned emphasis on *a priori* defined outcomes is in many cases difficult, as these are often dependent on these wider developments and almost impossible to anticipate. For example, the limited deployment of clinical functionality within the timeframe of the work impacted on the ability to study the proposed safety and quality indicators, prompting the evaluation team to choose indicative process measures surrounding data availability instead. Similarly, stakeholders were frequently unable to provide documents containing cost information thereby limiting access to relevant quantitative data. This led the evaluation team to develop costing categories qualitatively based on data obtained through interviews [21]. These experiences indicate that the most meaningful outcome measures therefore tend to emerge during the conduct of the research and are informed by qualitative findings. This can also help to ensure that unanticipated benefits and risks are fully explored.

Again, this highlights the need for longitudinal designs over extended timeframes in order to assess short-, medium-, and longer-term consequences (see Section 2). In the example study, an exploration of the longer-term impact of technological systems on the safety and quality of care as well as assessments of cost implications was planned, but impossible to conduct, as fully functioning systems had not been deployed. Consequently, the initial focus shifted towards assessing short- to medium-term processes qualitatively. Longer timeframes for the conduct of evaluations to allow measuring of quantitative outcomes are desirable, but if this is not possible, evaluators should focus on collecting robust quantitative baseline data to establish a basis for comparison to be used in future work, instead of attempting to quantitatively assess limited functionality in organizational periods of transition.

5.4. Theoretical pragmatism

As alluded to in Section 4, learning across implementations can be promoted by identifying causal mechanisms underlying observed outcomes and hypothesizing how these may translate to other contexts [29]. It is now commonly recognized that this can be achieved by drawing on theory when evaluating complex processes involving humans and technological systems.[30]

However, when considering the large-scale implementation of complex systems such as EHRs, there are some major theoretical challenges. Existing theoretical work tends to have restrictive foci relating mainly to micro-contexts surrounding implementations, potentially resulting in a lack of appreciation of contextual factors [31]. Therefore, extracting implications for overall strategic (organizational or political) decision making can be difficult. This can be addressed by drawing on a range of perspectives, an approach that is also apparent in the example study, where the protocol (Box 1) draws on principles of two very different theoretical lenses. Both approaches offered potentially helpful angles to the topic of investigation: realistic evaluation (a lens that pays attention to social and political context) and Cornford and colleagues' evaluation framework (a sociotechnical perspective) [32,33].

Moreover, diverse approaches conceptualize technologies, processes and stakeholders in significantly different ways [34]. These fundamental variations in philosophical underpinnings may result in limited opportunities for drawing on a

combination of theoretical lenses, and evaluators therefore have to select approaches carefully when designing health IT implementation evaluations. For example, the notion of “success” can be viewed very differently. Some research traditions, commonly associated with positivist approaches, view a “successful” implementation as something that can be prospectively defined and objectively measured. Through other theoretical lenses, “success” may be seen as emergent and context, viewpoint and time-dependent – that is, different stakeholders may view the system differently in different contexts and at different points in time.

Another question relates to the pragmatic use of theory. For instance, during the course of the evaluation activities (Box 1), it became apparent that none of the theoretical frameworks could integrate the complexity of findings obtained. This was particularly noticeable when attempting to integrate qualitative and quantitative discoveries, as the evaluation team was in many ways assessing temporary effects as opposed to fully functioning systems, whilst at the same time being faced with the challenge to provide immediately useful formative feedback (see Section 5.3 above). The existing body of theoretical literature, however, either emphasized complexity (potentially resulting in a lack of tangible recommendations) or over-simplified the implementation process (potentially jeopardizing a realistic assessment of the status quo). To address these tensions, the evaluation team decided to draw on theory but focus primarily on influencing policy making and helping to disseminate lessons learned to inform future strategies [35]. The choice surrounding degrees of theoretical integration will inevitably depend on who has funded the work and why, but more tangible outputs may be promoted by fostering closer relationships between evaluators and policy makers. This can be facilitated by identifying strategic priority areas for focusing investigations in advance, through for example conducting risk assessments in collaboration.

This is not to say that theory building is not important, but rather that it should primarily be retrospective in nature, with the benefit of hindsight informing these important deliberations, as a sufficiently rounded understanding of developments over time can only be obtained on reflection. A major theoretical challenge in this respect will be the attempt to extrapolate from studies of local contingencies to implementation policies. Such work will need to involve disentangling how and through what mechanisms structural developments result in local effects and vice versa.

6. Conclusions

As can be seen, evaluations of health IT implementations are important to determine impacts of technologies, identify risks and unintended consequences, and to promote inter- and intra-organizational learning. This is particularly important in light of the immaturity of many health IT systems that are now implemented internationally and the significant investments associated with these.

However, although health IT implementation evaluation is not a straightforward activity, due to variations in context and the transformative changes often brought about by technological change, some general desirable characteristics can be specified. These include a focus on content, context, process and outcomes; mixed methods longitudinal designs; formative and summative components; and effectively drawing on theoretical lenses to facilitate learning across implementations.

Commonly experienced conceptual and practical challenges during health IT implementation evaluation and potential ways to tackle these have been discussed. This

highlights the need to adapt methodologies in line with shifting implementation landscapes; the importance of formative evaluation activities; the necessity for evaluators to form close relationships with policy makers in order to situate local developments within their broader contexts; the initial need to focus on consequences and processes as opposed to outcomes accompanied by longer evaluation timelines; the need to collect meaningful baseline data; and the need for theoretical pragmatism to allow evaluations to be strategically useful.

It is hoped that these reflections will contribute to developing rigorous, yet flexible and empirically grounded, approaches to health IT implementation evaluation. In doing so, the work can help evaluators to anticipate problems that might arise in advance during this complex activity, and to address these in the most pragmatic and immediately useful manner possible.

Appendix: Box 1 + 2

Box 1. Summary of the originally envisaged protocol of an evaluation of the implementation and adoption of national electronic health records (EHRs) across hospital providers throughout England

Background: We will conduct both a formative and summative evaluation of the implementation and adoption of national electronic health records (EHRs) across hospital providers throughout England. In doing so, we will inform implementation and adoption, and to generate insights to inform future local and national strategic implementation decisions.

Aims: The main aims of this work are to inform the roll-out of EHRs in English hospitals with a view to ensuring that systems are successfully used and have the maximum chances of introducing benefits whilst minimizing harm. In doing so, our main aims are to:

1. Identify benefits and negative impacts of the new system across a variety of dimensions that were reflected in six interrelated work packages (WPs)
2. Liaise with policy makers throughout the project in order to inform both local implementation and national roll-out of the system.

Methods: We will conduct a prospective multi-faceted mixed methods evaluation of the implementation and adoption of EHRs in English hospitals in order to generate insights that can support the implementation (formative assessment) and the future roll-out to other settings (summative assessment). In doing so, we are planning to use the principles of a stepped wedge design to select hospitals, with data generation and analysis being informed by realistic evaluation and sociotechnical considerations drawing on Cornford and colleagues' evaluation framework.

Sampling of organizations and individual participants

We plan to recruit up to five hospitals from different geographical locations, demographics, specialties, system suppliers, predispositions to and history of information technology (IT) implementation. These will be conceptualized as in-depth longitudinal case studies in order to allow us to explore local processes and, based on this, make inferences relating to wider implementation-related developments through generating potentially transferable lessons. Data will be collected longitudinally, tracking changes over time in order to generate insights into the unfolding activities of implementation.

Individual participants at each site will be selected using purposive sampling to recruit a diverse range of interviewees. In doing so, we will recruit relevant individuals within hospitals (managers, implementation team members and IT staff, doctors, nurses, allied health professionals, administrative staff, patients and carers), and also stakeholders outside the immediate hospital setting (e.g. policy makers, system developers).

Individual work packages

WP 1 - Implementation, deployment and organizational learning: This WP is designed to study the implementation strategies and experiences (e.g. technical, clinical and organizational issues). Here, we plan to explore the interrelationship of different contexts (including the macro and local context), and the degree of organizational learning.

WP 2 - Attitudes, expectations and experiences of stakeholders: In this WP, we are seeking to study the attitudes, experiences and expectations of the various stakeholder groups in order to gain an insight into how systems are received.

WP 3 - Organizational consequences: The focus of this WP is on the organizational consequences expected as a result of implementation including, for example, changes in data quality, workflows, organizational roles and responsibilities.

WP 4 - Assessment of costs of implementation: This WP is designed to focus on the formative assessment of implementation costs and the development of a framework for costing that can be rolled-out to other hospitals that are yet to implement EHRs.

WP 5 - Assessing error, safety and quality of care: This WP is designed to consider key quantifiable benefits in relation to improving quality and/or safety of care, with a focus on those outcomes that are most likely to be influenced by the software in question. Four measures are planned: medication errors; medicines reconciliation on hospital admission; completeness of information provided at hospital discharge; and availability of key information in medical records in hospital outpatient clinics.

WP 6 - Organizational consequences and implications for future IT deployments and evaluations: This final WP is designed to integrate findings to provide the overall summative element of the evaluation, and to make recommendations for implementation and evaluation of future large-scale IT deployments in healthcare.

Integration of work packages

WPs 1, 2, 3 and 6 are qualitative in nature consisting of collecting hospital documents, conducting interviews, and on-site observations. Whilst data collection in most WPs will primarily be based on selected case study sites, WP 6 will extend the work to include other stakeholders (such as policy makers and system developers). The first three WPs are designed to allow data collection activities to be closely coordinated using, as far as is possible, the same researchers and the same respondents.

Data obtained from the qualitative work will feed into the development of quantitative measures of the impact of the EHR systems on the safety and quality of care (WP 5) and also help to develop the costing model (WP 4). Data obtained from all WPs will be integrated in WP 6, as this is designed to provide the overall summative element of this work.

Box 2. Common challenges of conducting health IT implementation evaluations and suggested ways of addressing these

Coping with the shifting nature of reality

Shifting implementation landscapes and timelines may need to result in changes in aims and methodologies

Although desirable in order to investigate changes over time, before-during-after assessments are often difficult because of inevitable delays in implementation. This is particularly true in relation to assessing software systems once they had had an opportunity to embed.

Changes in policy context may impact on timelines, technologies and strategies – may need to reconsider certain aspects of original plans (e.g. may need to move to a predominantly qualitative case study-based approach).

The focus of the work – researching aspirations or facts or both: need to determine throughout work which functionality is aspirational and which actually reflects the reality on the ground.

Evaluation framework needs to be discussed with investigators at contracting stage.

Researching implementation and deployment paying attention to both local and macro contexts

Evaluation in context: need to focus on data collection activities, whilst also paying attention to broader developments.

The need to explore and link micro and macro levels when evaluating technology and policy interventions that underlie its introduction.

Delineating stakeholders and organisations from which data will need to be collected.

Researching outcomes, researching consequences

The need to research temporary consequences of, and processes involved in, technological change for a variety of stakeholders.

The initial focus should be on consequences and processes as opposed to outcomes.

Flexibility in outcome measures as these are dependent on broader developments (e.g. implementation timelines).

Focus on formative evaluation activities and collection of meaningful baseline data.

Theoretical pragmatism

Some theories can be too complex and detract from the practical focus of the work, which is likely to be influencing policy making and helping to disseminate lessons learned.

A practically useful framework for evaluation needs to be devised in collaboration with policy makers and results need to be fed directly into policy making.

Recommended further readings

1. K. Cresswell, M. Ali, A. Avery, N. Barber, T. Cornford, et al., *The Long and Winding Road... An Independent Evaluation of the Implementation and Adoption of the National Health Service Care Records Service (NHS CRS) in Secondary Care in England*, <http://www.haps.bham.ac.uk/publichealth/cfhep/005.shtml>, last access 11 February, 2016.
2. Sheikh, T. Cornford, N. Barber, A. Avery, A. Takian, et al., Implementation and adoption of nationwide electronic health records in secondary care in England: final qualitative results from a prospective national evaluation in "early adopter" hospitals, *British Medical Journal* **343** (2011), d6054.
3. Takian, D. Petrakaki, T. Cornford, A. Sheikh, N. Barber, Building a house on shifting sand: methodological considerations when evaluating the implementation and adoption of national electronic health record systems, *BMC Health Services Research* **12** (2012), 105-129.
4. Robertson, K. Cresswell, A. Takian, D. Petrakaki, S. Crowe, et al., Implementation and adoption of nationwide electronic health records in secondary care in England: qualitative analysis of interim results from a prospective national evaluation, *British Medical Journal* **341** (2010), c4564.
5. T. Greenhalgh, H. Potts, G. Wong, P. Bark, D. Swinglehurst, Tensions and paradoxes in electronic patient record research: A systematic literature review using the meta-narrative method, *The Milbank Quarterly* **87** (2009), 729-788.

Food for thought

1. Does the "perfect" health IT implementation evaluation exist? If, yes, what does it look like?
2. How long after should one evaluate beyond implementation?
3. How can theory contribute to health IT implementation evaluation?
4. How can closer relationships between evaluators and policy makers be promoted whilst maintaining impartiality?

References

- [1] Z. Morrison, A. Robertson, K. Cresswell, S. Crowe, A. Sheikh, Understanding contrasting approaches to nationwide implementations of electronic health record systems: England, the USA and Australia, *Journal of Healthcare Engineering* **2** (2011), 25-41.
- [2] E. Coiera, Building a National Health IT System from the middle out, *Journal of the American Medical Informatics Association* **16** (2009), 271-273.
- [3] World Health Organisation, *Atlas – eHealth Country Profiles*, Global Observatory for eHealth, WHO, Geneva, 2010.

- [4] K. Cresswell, R. Williams, Z. Morrison, A. Slee, J. Coleman, et al., Evaluation of medium-term consequences of implementing commercial computerized physician order entry and clinical decision support prescribing systems in two 'early adopter' hospitals, *Journal of the American Medical Informatics Association* **21** (2014), e194-202.
- [5] A. Sheikh, T. Cornford, N. Barber, A. Avery, A. Takian, et al., Implementation and adoption of nationwide electronic health records in secondary care in England: final qualitative results from a prospective national evaluation in "early adopter" hospitals, *British Medical Journal* **343** (2011), d6054.
- [6] K. Cresswell, A. Worth, A. Sheikh, Integration of a nationally procured electronic health record system into user work practices, *BMC Medical Informatics and Decision Making* **12** (2012), 15.
- [7] L. Catwell, A. Sheikh, Evaluating eHealth interventions: the need for continuous systemic evaluation, *PLoS Medicine* **6** (2009), e1000126.
- [8] M. Scriven, *The logic of evaluation*, in: H.V. Hansen, et. al. (eds), *Dissensus and the Search for Common Ground*, CD-ROM (pp. 1-16), OSSA, Windsor, ON, 2007.
- [9] J.J. McGowan, C.M. Cusack, E.G. Poon, Formative Evaluation: A Critical Component in EHR Implementation, *Journal of the American Medical Informatics Association* **15** (2008), 297-301.
- [10] E. Ammenwerth, S. Gräber, G. Herrmann, T. Bürkle, J. König, Evaluation of health information systems—problems and challenges, *International Journal of Medical Informatics* **71** (2003), 125-135.
- [11] V. Symons, A review of information systems evaluation: content, context and process, *European Journal of Information Systems* **1** (1991), 205-212.
- [12] A. Donabedian, Evaluating the quality of medical care, *The Milbank Quarterly* **83** (2005), 691-729.
- [13] V. Viswanath, S.A. Brown, H. Bala, Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems, *MIS quarterly* **37** (2013), 21-54.
- [14] B. Latour, *Reassembling the social: an introduction to Actor-Network Theory*, Oxford University Press, Oxford, 2005.
- [15] C. May, T. Finch, F. Mair, L. Ballini, C. Dowrick, et al., Understanding the implementation of complex interventions in health care: the normalization process model, *BMC Health Services Research* **7** (2007), 148.
- [16] E. Rogers, *Diffusion of innovations*, Free Press, New York, 1983.
- [17] R. Williams, D. Edge, The social shaping of technology, *Research Policy* **25** (1996), 865-899.
- [18] P.N. Edwards, G.C. Bowker, S.J. Jackson et al., Introduction: an agenda for infrastructure studies, *Special Issue on e-Infrastructure of the Journal of the Association for Information Systems* **10** (2009), 364-374.
- [19] A. Black, J. Car, C. Pagliari, C. Anandan, K. Cresswell, et al., The impact of eHealth on the quality and safety of health care: A systematic overview, *PLoS Medicine* **8** (2011), e1000387.
- [20] K. Cresswell, A. Sheikh A, Undertaking sociotechnical evaluations of health information technologies, *Informatics in Primary Care* **21** (2014), 78-83.
- [21] K. Cresswell, M. Ali, A. Avery, N. Barber, T. Cornford, et al., *The Long and Winding Road...An Independent Evaluation of the Implementation and Adoption of the National Health Service Care Records Service (NHS CRS) in Secondary Care in England*, <http://www.haps.bham.ac.uk/publichealth/cfhep/005.shtml>, last access 11 February, 2016.
- [22] A. Robertson, D.W. Bates, A. Sheikh, The rise and fall of England's National Programme for IT, *Journal of the Royal Society of Medicine* **104** (2011), 434-435.
- [23] A. Takian, D. Petrakaki, T. Cornford, A. Sheikh, N. Barber, Building a house on shifting sand: methodological considerations when evaluating the implementation and adoption of national electronic health record systems, *BMC Health Services Research* **12** (2012), 105-129.
- [24] R.E. Stake, *The Art of Case Study Research*, SAGE Publications, London, 1995.
- [25] Practical Handbook for Ongoing Evaluation, How to establish a good budget for an evaluation? <http://wiki.interact-eu.net/pages/viewpage.action?pageId=23756932>, last access 11 February 2016.
- [26] K. Cresswell, A. Worth, A. Sheikh, Integration of a nationally procured electronic health record system into user work practices, *BMC Medical Informatics and Decision Making* **12** (2012), 15-27.
- [27] eHealth Insider, *Political row over NPfIT: London on hold*, <http://www.digitalhealth.net/news/24212/political-row-over-npfit-london-on-hold>, last accessed: April-14, 2015.
- [28] M. Berg, J. Aarts, J. van der Lei, ICT in health care: sociotechnical approaches, *Methods of Information in Medicine* **42** (2003), 297-301.
- [29] J. Mason, *Qualitative Researching*, Sage, London, 2002.
- [30] N. Thompson, *Theory and Practice in Health and Social Welfare*, Open University Press, Buckingham, Philadelphia, 1995.
- [31] J. Aarts, H. Doorewaard, M. Berg, Understanding implementation: The case of a computerized physician order entry system in a large Dutch university medical center, *Journal of the American Medical Informatics Association* **11** (2004), 207-216.

- [32] T. Cornford, G. Doukidis, D. Forster, Experience with a structure, process and outcome framework for evaluating an information system, *Omega - International Journal of Management Science* **22** (1994), 491-504.
- [33] R. Pawson, N. Tilley, *Realistic Evaluation*, Sage, London.
- [34] T. Greenhalgh, H. Potts, G. Wong, P. Bark, D. Swinglehurst, Tensions and paradoxes in electronic patient record research: A systematic literature review using the meta-narrative method, *The Milbank Quarterly* **87** (2009), 729-788.
- [35] A. Robertson, K. Cresswell, A. Takian, D. Petrakaki, S. Crowe, et al., Implementation and adoption of nationwide electronic health records in secondary care in England: qualitative analysis of interim results from a prospective national evaluation, *British Medical Journal* **341** (2010), c4564.

Evolving Health IT Systems Evaluation: The Convergence of Health Informatics and HTA

Persephone DOUPI^{a,1}

^aTHL, National Institute for Health and Welfare, Finland

Abstract. The credibility and reliability of health IT systems as a means of achieving changes towards safer and cost-effective care have been questioned for over two decades due to the lack of methodologically strong evidence. As national level adoption and implementation of health IT are becoming widespread across the EU and globally, but are also being offset by adverse reports, the demands for evidence become more pronounced and the stakes higher. The adaptation of HTA (health technology assessment) methodology as a means to address gaps in health IT evidence production has been proposed repeatedly and tested in the field of telemedicine services. HTA has in many ways run a course parallel to that of health IT, while in certain respects attaining more clear achievements. This contribution investigates aspects of a bilaterally beneficial relation between the two disciplines using three lines of exploration: the methodological goodness-of-fit between health IT evaluation and HTA; the solutions each has proposed to the problem of producing high quality evidence in reduced amount of time; and the way each has addressed the strengthened role and position of patients. The analysis demonstrates areas of convergence between health IT and HTA. It also highlights topics which would need to be jointly addressed in the process, such as innovative and high quality data collection and analysis, inclusion of patient reported outcomes and patient safety, and transferability and generalizability of findings. In closing, it takes a glimpse of the challenges emerging as a result of the progress at the cross-roads of medicine, science and technology.

Keywords. Medical informatics, biomedical technology assessment, outcome and process assessment, healthcare, policy.

1. Introduction

“Information technologies (IT) are often put forward as important instruments to improve quality and efficiency in health care. However, the evidence is lacking of the specific contribution of these technologies to outcome and efficiency improvement. ... A major cause for lack of evidence of effectiveness is the methodological difficulties in establishing this evidence. ... What is needed in this area is consensus on methods and criteria to be applied in assessment, similar as in e.g. evaluation of drugs or diagnostic devices. This consensus is needed both for the industry, as well as for the IT users, at various levels” [1].

¹ Corresponding author: Dr. Persephone Doupi, Senior Researcher, THL – National Institute for Health and Welfare. PL 30, 00271 Helsinki, Finland, persephone.doupi@thl.fi.

These words could have been written today, but surprisingly they are already twenty years old. High hopes and expectations have been placed on health IT as a major driver of changes that would make healthcare practices and systems better and safer, in an efficient and cost-effective manner. With financial pressures rising, populations and work forces both aging but also becoming increasingly mobile, the need for validated, proven and transferrable health IT solutions becomes an imperative, particularly from the perspective of decision makers. Meanwhile, Health Technology Assessment (HTA), although relatively young as a discipline, has established its position as a methodology for high quality, research-based evidence generation. In turn, such evidence forms the basis supporting the decisions necessary for ensuring health system sustainability. Could HTA hold the key to improving the quality, reliability and cost-effectiveness of health IT?

Until now, the subject has been approached on the basis of what health IT could and should learn from HTA, particularly in order to be able to produce robust and convincing evidence of its worth. The additional dimension this contribution aims to bring to the existing discourse is the complementary side of the image, i.e. what HTA can gain from health IT and what are the requirements for this to be achieved. In other words, the contribution approaches the subject of discipline synergy as a mutually beneficial process, with a particular emphasis on methodological issues.

After a short introduction covering the evolution of the health IT domain (section 2), we will explore the following three questions:

- *How good is the fit between the two disciplines? (section 3)*

We explore the question of degree of fitness through a review of selected HTA studies of key health IT applications, in order to demonstrate the messages which have emerged through the application of HTA methodology. Conversely, the experiences gained and lessons learned through the large scale testing of the Model for the Assessment of Telemedicine (MAST model) [2] in several EU-funded projects provide indications on health IT's experiences with HTA approaches – MAST being the first concrete instrument to result from adaptation of HTA methodology instruments to telemedicine evaluation.

- *How have the two disciplines addressed the challenge of quick production of high quality evidence? (section 4)*

Technology and change are deeply intertwined and as choices amongst interventions increase and resources diminish policy makers demand reliable evidence within a shorter turn-around cycle. We go further into methodological perspectives by examining examples of the instruments each of the two fields has developed and proposed in order to address the problem of rapid delivery of quality evidence.

- *How have HTA and health IT related to the role of patients? (section 5)*

Making the patient a (potentially equal) partner in the production of evidence goes hand in hand with the trend of patients taking a central role in the definition and implementation of their care plans. The efforts undertaken earlier in each of the two fields are raised to a different dimension of prospects and challenges brought about by the possibilities of own data production and analysis facilitated by latest technologies such as health apps and sensors.

Finally, we will sum up the conclusions to be drawn from the parallel and intersecting course of the two fields so far and attempt to look at future developments, including drawing up a list of proposed work topics for those interested and motivated to explore further the synergies between health IT and HTA.

2. Setting the scene

2.1. Definitions of the health IT domains and their evolutions

Health IT, just as any area of activity relying heavily on technology, is in a constant state of change, which is also reflected in the abundance of definitions over time. The scientific discipline and corresponding term of Medical Informatics dates back to the 1970 and was seen as belonging to the area of applied informatics research [3]. According to van Bemmel, “*medical informatics comprises the theoretical and practical aspects of information processing and communication, based on knowledge and experience derived from processes in medicine and health care*”. [4]

The expansion to the more general term ‘Health Informatics’ reflects the need to capture the increasingly multidisciplinary practice of medicine, as well as the growing interest in a universal approach to health and well-being. The concept of eHealth – a newer arrival – was if not born, at least strongly supported by policy making such as the European Commission’s hallmark eHealth Action Plan of 2004 [5]. The scientific community also invested considerable energy in demarcating the field, with a series of articles published on eHealth definitions about a decade ago and the exploration still ongoing [6].

Following the drive of technology evolution through mobile devices, we are living already in the mHealth era [7]², which presents some interesting new features, namely the increased practical ability to focus on individuals rather than organisations, decentralization and ubiquity [8].

It is important to also take a look at the definitions related to telemedicine, since it has partly been a distinct field all along, as well as the first concrete test-bed for HTA methodology application in IT. Taking once again the view of the policy makers, we see the European Commission understanding the element of distance between patients and healthcare professionals as the defining feature of telemedicine and telemonitoring [9]. From the UK then we have the emergence of two related, but still distinct terms: “Telehealth” and “Telecare” [10]. What differentiates telehealth from telemedicine is its inclusion of preventive, promotive and curative aspects. Telecare on the other hand is rather close to telemonitoring if viewed specifically in the context of home care, but including wider daily living aspects.

2.2. The imperative of evidence

“*The major finding from reviewing the empirical evidence – which is of variable quality ... is that there is very limited rigorous evidence demonstrating that these technologies actually improve either the quality or safety of healthcare.*” [11]

² See also: B. Vallespin et al., Ensuring evidence-based safe and effective mHealth applications, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

Other than the fact that it is a persistently recurrent finding, why is the lack of evidence on health IT a major problem? A large part of the answer is to be found in the implications of health policy making, particularly on the EU-level, where eHealth uptake has been consistently promoted. The 2004 eHealth Action Plan has been followed by a series of ministerial conferences and decisions (e.g. Communication on Telemedicine & large scale pilots, eHealth Task force report Redesigning Health 2020 etc.) and more recently the new eHealth Action Plan until 2020 [5, 9, 12, 13].

The general and increasingly stronger trend is the transition from hospital-based care to extramural care with shortened periods of hospitalization, combined with the increased transfer of responsibility to the patient/citizen. The degree of collaborative work and the variety of actors involved to realize these plans essentially make health IT an indispensable tool [14]. Moreover, the advent of cross-border care brings along new demands and challenges or expands old ones to a new scale [15]; either way the solutions are unimaginable without health IT.

However, the fact that these policy and practice shifts towards a dependency on health IT are being espoused without evidence of its supporting quality or safety of healthcare is alarming – hope-based policy is no substitute for evidence-based policy and would not be tolerated in other aspects of health practice. Promoting the use and integration of technologies cannot be done without simultaneously preventing or mitigating the accompanying risks, such as diversion of valuable resources or causing actual harm. Reflecting these concerns, the WHO saw the ‘rigorous evaluation of eHealth’ as a necessary requirement and among the recommendations and identified action items [16] listed the following priorities:

- Identify and adapt, where necessary, robust and relevant tools for the evaluation of eHealth;
- Develop simple and pragmatic tools to enable decision makers to review and select eHealth systems, based on appropriate evaluation-generated evidence of impact, and potential for scalability and sustainability;
- Develop principles and recommended practices to evaluate and assess eHealth, with a view to increasing transparency, accountability and integrity.

A similar approach can be seen in the eHealth Stakeholder Group statement concerning telemedicine: “*Benefit and added value of telemedicine services should be systematically monitored and evaluated to allow for justified inclusion into guideline supported clinical practice.*” [17]

In the quest for robust evidence and methods to obtain it, it is easy to see why HTA is a good candidate source. According to the HTA Glossary [18], Health Technology Assessment has been defined as follows:

“*The systematic evaluation of the properties and effects of a health technology, addressing the direct and intended effects of this technology, as well as its indirect and unintended consequences, and aimed mainly at informing decision making regarding health technologies.*”

The glossary further notes that “*HTA is conducted by interdisciplinary groups that use explicit analytical frameworks drawing on a variety of methods*”.

The discipline focuses on the assessment of individual health technologies in a manner useful to decision makers, combined with the adoption of a global perspective: ensuring that the medical, social, ethical, and economic implications of the development, diffusion, and integrated use of technologies are addressed. On the other

hand, the technologies usually being the subject of HTA analysis are different in nature from health IT. In the view of many, eHealth/health IT applications are socio-technical systems.³ There is a constant interplay between the technology and human/social factors in the environment of implementation, which brings about changes in all involved (systems, humans, organizations, services). Further, health IT systems are usually a combination of technologies and services, or a means supporting innovative service provision. Essentially, before progressing with proclaiming HTA as the method of preference in health-IT evaluation we need to address the first of our questions.

3. How good is the fit between the two disciplines?

We shall begin with exploring how well the HTA approach transfers to a domain with the features of health IT. Through utilizing Ohtanen [19], an online database of HTA reports maintained by THL [20], we have selected among the featured health IT/eHealth related HTA studies examples focusing on popular and promising health IT applications, such as: early warning and handover systems, Computerized Physician Order Entry (CPOE), medication management, and treatment of psychiatric disorders. The main focus is on the findings and conclusions of the studies, in order to gain insight on what they have revealed for health IT applications and their developers and how well they have addressed socio-technical matters.

3.1. *Health technology assessment of the use of information technology for early warning and clinical handover systems*

The study conducted by the Irish Health Information and Quality Authority (HIQA) [21] examined clinical and cost-effectiveness of IT for early warning and clinical handover systems. It used the methodology of a systematic literature review. In addition, benefits and investment requirements were estimated and key themes for effective robust implementation were outlined. The results indicated the presence of some evidence that implementation of electronic early warning systems has contributed to reduced mortality rates. However, the quality of studies on the clinical effectiveness of these systems was hampered by poor study design, small sample size and unspecified follow-up, while cost-effectiveness data was minimal. Due to the significant differences in the models of healthcare provision between the US and Ireland the ability to generalize return-on-investment findings to the Irish context was deemed rather uncertain.

On the socio-technical aspect, the review found clinicians' perception of improved patient safety to be positive, due to better handover communication processes. However, a face-to-face element to clinical handover was identified as an important part of patient care. The review also underlined the importance of strong leadership and adequate staff training levels⁴ and pointed out the significant capital investment

³ See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁴ See also: E. Hovenga et al., Learning, training and teaching of health IT and its evidence for informaticians and clinical practice, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

required for implementation. The study recommended that in order to maximize the effectiveness of implementation, the employment of human factors analysis would be instrumental in creating work environments supportive of productivity, while minimizing risks to patient safety.⁵

3.2. Medication management and IT

The objective of the study by McKibbon et al. [22] was to review the evidence on the impact of health IT on all phases of the medication management process (prescribing and ordering, order communication, dispensing, administration and monitoring as well as education and reconciliation), to identify the gaps in the literature and to make recommendations for future research.

Most included studies evaluated changes in process and outcomes of use, usability, and knowledge, skills, and attitudes. Most showed moderate to substantial improvement with implementation of IT-enabled medication management. Although the field of IT-enabled medication management is well-studied, a closer examination of the literature showed that the evidence is not uniform across phases of medication management, groups of people involved, or types of medication management. The application of health IT to medication management was assessed as having the potential to improve processes; however, shortage of clinical and economics studies and limited understanding of sustainability issues were also identified.

With regard to socio-technical parameters, the study showed that physicians were more often the subject of evaluation than other participants. Even though other health care professionals, patients, and families have an important role to play they are not studied as thoroughly as physicians. These non-physicians groups often value different aspects of IT-enabled medication management, have diverse needs, and use systems differently.

3.3. Computerized Physician Order Entry – effectiveness and efficiency of electronic medication ordering with decision support systems

Prescription is an important step within medication management, and health IT in the form of Computerized Physician Order Entry/Clinical Decision Support Systems (CPOE/CDSS) has been specifically designed to support it. The study [23] examined the effects CPOE/CDSS on medication errors. The study found that CPOE/CDSS systems are able to reduce the rate of errors when ordering medications.⁶ However, using the data available, it could not be assessed conclusively to what extent CPOE systems or the reduction of medication errors has an impact on the Adverse Drug Event (ADE) rate – a clinically more relevant element - or on mortality. Regarding the cost-benefit-ratio from the hospital perspective, the two qualitatively best economic studies arrived at contradictory conclusions. A positive cost-benefit-ratio for any individual hospital cannot therefore be assumed, particularly as the results cannot be generalized.

Prospective, systematic multi-centre evaluation studies with clear methodology which include an analysis of the user-friendliness and of social and technical aspects of

⁵ See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: *ibid.*

⁶ See also: H. Seidling et al., Evaluating the impact of health IT for medication safety, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

the system are needed. A detailed description of the system used and of the hospital evaluated is essential. If possible, costs and cost effects should be surveyed and documented transparently. The authors noted that a quantitative evaluation of the economic effects of implementing a CPOE/CDSS system in (all) hospitals in a large country would be too far-fetched: the reliability of study results regarding relevant endpoints was found to be still limited. Conclusions in regard to another context are only possible when data presentation is highly transparent, enabling local assessment of transferability of evidence. Structured interviews at selected hospitals with and without CPOE/CDSS systems would generate important input and help to assess the need for further research.

3.4. *Telehealth Services for the Treatment of Psychiatric Issues: Clinical Effectiveness, Safety, and Guidelines*

The study conducted by the Canadian Agency for Drugs and Technologies in Health [24] analyzed the reported clinical effectiveness and safety of tele-psychiatry. It accepted for inclusion seven studies in total, including two systematic literature reviews, two randomized controlled trials and three clinical practice guidelines. One of the two analyzed systematic reviews reported that very few studies have examined the effectiveness of tele-psychiatry in improving the outcomes for patients or clients. Therefore, although positive outcomes have been reported for the management of depression, post-traumatic stress disorder, bulimia nervosa and psychosis, the evidence was not enough to support strong conclusions about effectiveness. Tele-psychiatry may be comparably effective and safe; and may be a feasible alternative for making tele-mental health services available in resource constrained settings.

On the socio-technical dimensions, the study places emphasis on two issues, in order to promote effective management of emergency situations when providing tele-mental healthcare services: a) ensuring that both patients and staff at the point of care are familiar with emergency protocols and procedures specific to each of the tele-psychiatry services and environments in which care is provided, and b) that staff have had appropriate training on the procedures and techniques.

3.5. *MAST model: adapting HTA methodology to health IT*

Having explored the view of health IT from the HTA perspective, now we reverse our observation angle. We review the experiences from the application of the MAST model [25] – the first tangible tool born out of an effort to adapt HTA methodology to the needs of a health IT area. MAST was developed in the context of an EU-funded project (MethoTelemed) [26] that aimed at developing an evaluation framework for telemedicine applications, based on the principles of the HTA Core Model [27]. Its creators have described it as “*A multidisciplinary process that summarizes and evaluates information about the clinical, economic, organizational and socio-ethical issues related to the use of telemedicine, in a systematic, unbiased and robust manner*” [27].

The first large scale testing and validation of the MAST model took place in the context of the Renewing Health project, a study of telemedicine services covering almost 7,000 patients across nine EU regions [28]. The study also utilized know-how of the UK National Health Service Whole System Demonstrator project, in assessing patients’ perceptions with the same 22-question questionnaire [29]. From the

experiences reported on the methodological aspects of assessment [30] it is important to note some key points:

- Economic aspects are difficult to assess in a transferrable manner because of their dependence on a deep knowledge of the organisation of healthcare and reimbursement systems at the location of implementation.
- Organizational aspects are affected by the availability of technology or skills to master the technology which in turn might differ across countries;
- The main problem within transferability of assessments is the general lack of interest within the field, leading to limited dedicated local resources. There is no strong tradition within the field of transferability or generalisability on methodology or reporting.
- Even though MAST can produce useful information for assessing a telemedicine application, its applicability is limited when it comes to evaluating a new, immature application [31].

The overall impression of applying MAST, based on the questionnaire answers of the 11 cluster or pilot leaders in the Renewing Health project, was that it is a valuable framework. Challenges included problems in obtaining scientific and rigorous knowledge from local sites, as well as assessment of ethical and legal aspects. Both the legal framework as well as the perception on the delimitation between legal and ethical issues, appeared to vary across countries/regions. A practical proposal to address the latter problem is explicitness and transparency, by means e.g. of a description of the legislative references and how they were met in the context of each specific project.

Additional value generated by the Renewing Health application of MAST has been the accompanying guidance documentation generated (concerning data collection, analysis and reporting of results) as well as the expertise attained locally [32]. Following Renewing Health, various national and international telemedicine studies in Europe, some of them still ongoing, have proceeded to use MAST as the framework for assessment. Of particular interest is the extension and adaptation of the model to cover aspects of social and healthcare integration (such as in the projects SmartCare and CareWell) [33, 34], as well as the application of the MAST model in accordance with methodologies of the HTA field (recommendations of the International Society for Pharmacoeconomics and Outcomes Research Good Research Practice Task Force on Prospective Observational Studies and the STROBE statement) [35].

4. How have the two disciplines addressed the challenge of quick production of high quality evidence?

Change is a typical feature of the evolution phase of a health IT, as defined by Brender [36], according to whom the evolution phase starts when "*the complete IT solution has achieved reasonable stability (with regard to faults and adaptation) of operations and when actual new developments or major revisions are being started*". Generally in the field of health IT, innovation and speed are desirable properties, giving rise to continuously changing technologies.

As a result, any form of evaluation and assessment activity has often been perceived as a hindering factor. Even more so, HTA methodology traditionally comes into play only after a technology has matured enough and evidence has been gathered

that can inform decision making. How feasible is it to align these needs and traditions in the best way in order to achieve the desired development and implementation of evidence-based health IT? Can the newly developed Rapid Assessment methodology of HTA provide a solution to the challenge of evolving systems evaluations? What respective trends can be detected on the side of health informatics?

Let us begin by taking a look at what Rapid Assessment HTA methodology is about. Rapid Relative Effectiveness Assessments (REAs) are assessments of a specific technology within a limited time frame in comparison with one or more relevant alternative interventions. (Relative) effectiveness focuses on events occurring under the usual circumstances of health care practice, as opposed to (relative) efficacy, where observations are produced under ideal circumstances [37]. A REA covers generic research questions (i.e. issues) considered most relevant for four different applications each focusing on the assessment of specific types or uses of health technologies: pharmaceuticals, diagnostic technologies, medical and surgical interventions and screening technologies

The first published version (V3.0) of the HTA Core Model® for Rapid Relative Effectiveness Assessment (abbreviated as 'Model for Rapid REA') was developed for pharmaceuticals only with the intention to produce a rapid assessment within a limited timeline [38]. The driver was EU countries' legal obligations in assessment of technologies or for the purposes of pending coverage decisions. The latest version of the Model (V4.1) (public consultation of which has been recently completed) has been extended to cover also applications for medical and surgical interventions, and for screening and diagnostic technologies. The aims are three-fold: to improve the applicability of HTA information in other (e.g. national or regional) HTA projects, to enable actual collaboration between HTA agencies by providing a common framework for the production of rapid Relativeness Effectiveness Assessments, and to avoid duplication of work.

The 'Model for Rapid REA' is based on the HTA Core Model® [27]. But where the Core Model® organizes study-relevant information into nine domains, the Model for rapid REA – in search of time savings – covers only the first four domains⁷ and within these domains only a subset of issues. In addition, and because the objective is to share commonly required elements of information, only information that is considered both important and transferable is collected.

The remaining five domains (i.e. on costs, ethical, legal, social and organizational issues) are excluded as highly context-dependent topics and hence areas of limited transferability. Instead, a checklist is supplied for a quick assessment of possible relevant issues emerging in these domains which would be justifiable to address. Relevant assessment elements from these four domains may be selected from the HTA Core Model®. Pre-established problems/issues, with regard to ethical, organisational, social and legal aspects, which are common to the technology to be assessed and its comparator(s) will, as a rule, not be addressed, as it is not to be expected that the addition of a new technology will lead to changes. To date, several studies have been published based on the application of the REA Model [39, 40].

The attempt to speed up the evaluation process is visible also on the side of health informatics, although the context and driving forces partly differ. Both demonstrative examples come from the United States, where political will and its translation to

⁷ Description and technical characteristics of technology (TEC), health problem and current use of the technology (CUR), clinical effectiveness (EFF) and safety (SAF).

legislation have lately been boosting the wide-scale implementation of EHR (Electronic Health Record) systems. Time pressures, combined with the high-speed change of the implementation environment, generate the need for new approaches to evaluation. It is also perceived that more rapid evaluation will also increase relevance and thus make the translation of results into policy and practice more likely.

Drawing on the tradition of anthropology and its mixed-method approaches to evaluation⁸, McMullen et al. [41] developed a rapid assessment methodology for clinical information systems, intended to be flexible enough as to address the needs and characteristics of different healthcare practice settings. Application of the method in the assessment of clinical decision support systems has provided already interesting results, by making explicit critical areas of unanimity and even more so difference of opinions among vendors and users (on usability, training, metrics, interoperability, product use, and legal issues) [42].

Glasgow and colleagues on the other hand [43] have investigated the possibilities offered by implementation science approaches, selecting four candidate models (the Evidence Integration Triangle; the Expanded Chronic Care Model; the Health Literate Care Model; and Reach Effectiveness Adoption Implementation Maintenance model RE-AIM) and applied their methodology in several studies [44, 45]. The authors have acknowledged the potential for applying HTA approaches to rapid evaluation of health information systems and have also drawn attention to the need for a stronger role of patients in health information systems evaluations, by making concrete proposals for the incorporation of patient-reported outcomes in electronic health records [46]. The latter brings us to our next topic, the role of the patient in HTA and health IT evaluation approaches.

5. How have HTA and health IT related to the role of patients?

Health IT developers and scientists as well as the HTA community have each in their own ways approached the subject of the patient taking up a different, more defining and determining role in modern healthcare delivery. As mentioned earlier, increased responsibility of patients for their care has also become an essential component of many healthcare policies. It is therefore relevant to explore whether there is a shared view and vision of the role of the patient between the two communities, identify what has been done in practice to achieve it and what could be the next steps.

Health informatics has had a long and difficult road in establishing successful relationships with health IT end users who, in most cases have been healthcare professionals, rather than patients [47]. The attention to patients and their needs, as well as the consideration of possible means for better engaging them has been constantly on the rise for over a decade⁹, nevertheless the focus is still more on the design rather than the evaluation side of health IT. And when it does come to

⁸ See also: P.J. Scott et al., Mixed methods: a paradigm for holistic evaluation of health information system, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁹ See also: L.L. Novak et al., Realizing the potential of patient engagement: Designing IT to support health in everyday life, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

evaluation, there is still some way to go in order to see results of health informatics applications use on the level of patient (relevant) outcomes [48].

The trend is however there [49] and this could be yet another input area from HTA to health IT: the engagement of patients through patient reported outcomes. Systems built by eHealth researchers and developers should take into account the accumulated experience and available tools on the side of the HTA community, as well as existing information on types of data required [50].

The complexity and challenge of accepting and engaging patients as equal stakeholders, essentially a major paradigm shift, is demonstrated by the fact that, in spite of the consistent and long-term commitment of HTA researchers to ensure the representation of patient priorities, a lot remains to be done. A 2010 report [51] targeted at clarifying the views of patients and their organisations towards HTA, as well as the understanding of various stakeholders, concluded that "*patients are ignorant about HTA, and regard the process as complex, and often beyond their comprehension*". Greater involvement, more transparency and heavier patient influence are among the desired changes.

6. Working towards synergy

Having come full circle, we can now review the key messages that have emerged in answer to the three questions which guided us in this journey.

6.1 Goodness of fit between the two disciplines

The concern as to whether HTA can address successfully the socio-technical dimension of health IT at least at first instance would not seem completely justified; HTA studies have explicitly brought forward socio-technical dimensions in which health IT implementations could and should invest more and have also proposed the means for doing that. Nevertheless, that is not to say that there are no differences in the type of technologies which have traditionally formed the bulk of work for each discipline. In that respect, the disambiguation of each field's targeted technology characteristics, the identification of its particularities and a harmonized means of describing it, also constitute an area where further work will be required in order to bring HTA and (evidence-based) health informatics closer together.

The application of HTA to health IT systems keeps returning to the need for improvement of quality and reporting of studies. The good news is that tools to support the task are available [52, 53] and the challenge remains to ensure their uptake and implementation among health IT researchers.

6.2 Addressing the balance of speed and quality in production of high quality evidence

Both disciplines have explored methods for rapid evidence generation, while preserving high quality standards. Experiences from their application have pinpointed transferability and generalizability of findings as challenges shared by both domains. Transferability was found to be one of the weak areas in the validation of MAST and it is an area that HTA REAs have addressed by trying to avoid the inclusion, if possible, of the most controversial assessment domains such as those of ethical and legal aspects.

Transferrable and more generally applicable results constitute such key potential benefits of coordinated evaluation that they still deserve thorough analysis and investigation. In addition to being tested, the REA proposed approach of selective inclusion can also be considered. Ethical principles do also have universal dimensions, codes of conduct have been internationally agreed at least for some areas of activity, legal frameworks can be developed and managed on the international level and organizational issues present also internationally. Moreover, flagging up of key 'hot spot' issues which will need to be addressed within the local context of health system, law, and ethics of health care delivery will be helpful, since the solutions can only be found locally.

Overall, the areas of comparative efficacy and effectiveness assessment, in their traditional but even more so in their rapid form, also constitute a meeting point for HTA and health IT – the shared interest being the collection of relevant data. At the core of the data collection process are EHRs – Electronic Health Record Systems. Relative efficacy and the 'ideal' use circumstances of clinical trials have constituted the focus of many research and development health informatics projects [54-56]. The other major source of observational data is patient registries appropriately termed as the "goldmine" of healthcare [57] – a recognition of the value engrained in extensive collections of curated longitudinal patient data. Paradoxically perhaps, but largely due to historical reasons, a major limitation of registries is the currently low uptake and utilization of IT in standard operations. In the EU context the problem is being addressed on several complimentary levels though a collaborative effort of the European Commission and several Members States (the PARENT Joint Action) [58], aiming at ensuring high quality and interoperable electronic registry data. Moreover, collaboration with the EU HTA community has begun, in an effort to clarify and address the data needs of HTA from electronic patient data collection systems [59].

6.3 The role of patients in evaluation

Focusing attention on the preferences, experiences and perceptions of patients is one of the important pathways in the future development of both disciplines, including among others the incorporation and utilization of patient reported data as part of regular data collection and analysis processes both in healthcare services, as well as in statistics and research [50, 60, 61]. In addition to the action points raised earlier, yet another area of particular interest from the perspective of patients is that of safety, where HTA and health IT should also seek synergies and build jointly on their existing achievements.¹⁰

The Safety (SAF) domain of the HTA Core Model covers "...the direct and indirect harms of a technology for patients and staff and how to reduce the risk of harms. There is usually a spectrum of known and unknown harms, which can be intended or unintended, of different seriousness, and dose or time dependent". In the context of health IT, the element of safety can be viewed from at least two different sides:

- a. development of systems for recognition and identification of patient harm;

¹⁰ See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

- b. follow up of the safety profile of health IT applications themselves, an area that presents its own set of challenges. Nevertheless, steps are being taken such as the Health IT Safety Roadmap in the US and the focus on engagement of patient as partners of patient safety initiatives [62, 63].

7. Emerging Challenges

As technology keeps evolving, new challenges lie ahead when seeking synergies between health IT and HTA. Two such upcoming areas are direct to consumer digital health technologies, and modelling tools for personalized medicine.

The former, part of the mHealth area, are suffering from the same shortage of meaningful outcomes and means for quality control as many other health IT applications before them [64] (see also *Chapter 20: Ensuring evidence-based safe and effective mHealth applications*). In addition, technologies such as health and well-being apps have a very rapid development and evolution cycle, combined with very localized implementation setting [65]. Applying or extending HTA methodology to cover this domain will not be an easy task. On the other hand, there are once more gains expected for both sides: it is in seeking clinically meaningful endpoints that HTA turns to direct data collection from the patient rather than relying on the interpretation of clinicians regarding the degree of success in treatment outcomes. mHealth opens new routes precisely for that sort of data gathering.

While the science and practice of medicine move towards personalized, preventive and integrative solutions, also the methodology of assessing health technologies will have to keep abreast with these developments [66]. As researchers involved in the Virtual Physiological Human project have discussed, the application of HTA in a field based on predictive computer modeling will need to take into account two additional elements: a) the technology's possibility to revolutionize currently applied clinical guidelines; b) the extension of the assessment to Research and Technology Development (RTD) policymaking, i.e. decisions made during the development of the technology itself. The proposed solution is based on the introduction of technology readiness levels throughout its lifecycle.

Taking stock of the broader landscape, health IT evaluation stands to benefit from a consolidation and through it possibly better uptake of robust methodology. HTA, even though an equally multi-disciplinary field, has successfully aligned the various approaches to the area of study and worked consistently towards a unified view and application of methodology on an international scale, albeit respecting national differences. HTA agencies and units operating world-wide on regional and national levels already provide a more concrete structure and operational framework for coordination of activities. Health informatics on the other hand is still very much either a business or an academic undertaking. The move into policy and established organizational forms such as national competence eHealth centers in several EU countries is a relatively recent development. The successful coupling between the academic and the policy world is still a challenge to be answered. Academic organisations and communities need to focus their efforts and initiatives on establishing communication and collaboration channels to policy makers. Promoting the objective of robust quality and policy-relevant evidence for health IT is a step in the right direction.

Recommended further readings

1. Geissbühler A, How Can eHealth Help Fix Broken Health Systems? *Methods Inf Med* **50** (2011), 297–298.
2. Sorenson C, Drummond M, Kanavos P, *Ensuring Value for Money in Health Care. The role of health technology assessment in the European Union*, Observatory Studies Series No 11, 2008. European Observatory on Health Systems and Policies, <http://www.euro.who.int/en/about-us/partners/observatory/publications/studies/ensuring-value-for-money-in-health-care-the-role-of-health-technology-assessment-in-the-european-union>, last access 11 February 2016.
3. May N, Godfrey M, *Who has the say in Health Technology Assessment?* Report, Hill & Knowlton, London, 2010, <http://vortal.htai.org/?q=node/173>, last access 11 February 2016.
4. IOM (Institute of Medicine), *Best care at lower cost: The path to continuously learning health care in America*, The National Academies Press, Washington, DC, 2013, <http://www.nap.edu/read/13444/chapter/1>, last access 11 February 2016.

Food for thought

1. What means would you propose for enhancing the uptake of methodological standards to improve the quality of the available health IT evidence such as GEP-HI¹¹ and STARE-HI¹²?
2. What opportunities and challenges do you see emerging for health IT and HTA through their application to integrated care environments?
3. What is your view on the HTA REA Model approach to evaluation items of limited transferability (see sections 4 & 6.2. of this contribution)? Do you find the argumentations in favor of exclusion valid or not and why? What other ways of addressing transferability challenges can you come up with or have already encountered?
4. How do you understand patient engagement in health IT evaluation? How are you addressing it in the project(s) you are currently involved with?

References

- [1] E.M.S.J. van Gennip, J.L. Talmon, editors, *Assessment and Evaluation of Information Technologies in Medicine*, Studies in Health Technology and Informatics **17** (1995), Amsterdam, IOS Press.
- [2] K. Kidholm, A.G. Ekeland, L.K. Jensen, J. Rasmussen, A model for assessment of telemedicine applications: MAST, *International Journal of Technology Assessment in Health Care* **28** (1) (2012), 44-51.
- [3] J.H. van Bemmel, M.A. Musen, editors, *Handbook of Medical Informatics*, Bohn Stafleu van Loghum, The Netherlands, 1997.
- [4] J.H. van Bemmel, The structures of medical informatics, *Med Inform* **9** (1984), 175-80.

¹¹ See also: P. Nykänen et al., Quality of health IT evaluations, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

¹² See also: E. Ammenwerth et al., Publishing health IT evaluation studies, in: *ibid*.

- [5] European Commission. *e-Health - making healthcare better for European citizens: An action plan for a European e-Health Area*. COM 0356 final, Communication from the Commission, April 2004. <http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52004DC0356>, last access 11 February 2016.
- [6] E.A. Boogerd, T. Arts, L.J. Engelen, T.H. van de Belt, "What Is eHealth": Time for An Update? *JMIR Res Protoc* **4**(1) (2015), e29.
- [7] European Commission, *Digital Agenda for Europe: mHealth*. <http://ec.europa.eu/digital-agenda/en/mhealth>, last access 11 February 2016.
- [8] J.C. Briggs, C. Adams, S. Fallahkhair, A. Iluyemi, D. Prytherch, *M-health review: Joining up healthcare in a wireless world*, Report, Univ. of Portsmouth, 2012, <http://eprints.port.ac.uk/8941>, last access 11 February 2016.
- [9] European Commission, *Communication Telemedicine for the benefit of patients, healthcare systems and society*, COM **289** (2008), <http://eur-lex.europa.eu/legal-content/GA/TXT/?uri=celex:52008DC0689>, last access 11 February 2016.
- [10] EHTEL, *Sustainable Telemedicine: paradigms for future-proof healthcare*, A Briefing paper, prepared for the Sustainable Telemedicine Task Force, February 2008. <https://www.ehtel.eu/publications/ehtel-briefing-papers/ehtel-briefing-paper-sustainable-telemedicine>, last access 11 February 2016.
- [11] A.D. Black, J. Car, C. Pagliari, et al., The impact of eHealth on the quality and safety of health care: a systematic overview, *PLoS Med* **8** (2011), e1000387.
- [12] eHealth, *Task Force report – Redesigning Health in Europe for 2020*, May 2012, <https://joinup.ec.europa.eu/node/125731>, last access 11 February 2016.
- [13] European Commission, *eHealth Action Plan 2012-2020*, December 2012, <https://ec.europa.eu/digital-agenda/en/news/ehealth-action-plan-2012-2020-innovative-healthcare-21st-century>, last access 11 February 2016.
- [14] A. Geissbühler, How Can eHealth Help Fix Broken Health Systems? *Methods Inf Med* **50** (2011), 297–298.
- [15] D. Horgan, D. Byrne, A. Brand, EU directive on patients' rights to cross border healthcare, *BMJ* **347** (2013), f7694.
- [16] WHO, *The WHO Call to Action on Global eHealth Evaluation - Consensus Statement of the WHO Global eHealth Evaluation Meeting, Bellagio, September 2011*, <https://imianews.wordpress.com/2011/10/31/e-health-evaluation-bellagio-call-to-action-released>, last access 11 February 2016.
- [17] European Commission, *Report of the EU eHealth Stakeholder Group on the widespread deployment of telemedicine services in Europe*, March 2014. http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=5167, last access 11 February 2016.
- [18] INAHTA (International Network of Agencies for Health Technology Assessment), *Health Technology Assessment glossary*, http://inahta.episerverhotell.net/upload/HTA_resources/Edu_INAHTA_glossary_July_2006_final.pdf AND [http://htaglossary.net/health+technology+assessment+\(HTA\)](http://htaglossary.net/health+technology+assessment+(HTA)), 2006, last access 11 February 2016.
- [19] FinOHTA (Finnish Office for Health Technology Assessment), *HEN Sources of Evidence database*, <http://data.euro.who.int/HEN/Search/SearchResult.aspx?SourceEvidenceId=17>, last access 11 February 2016.
- [20] *Ohtanen database*, Finnish homepage: <https://meka.thl.fi/ohtanen/Default.aspx>, last access 11 February 2016.
- [21] The Health Information and Quality Authority (HIQA), *Health technology assessment of the use of information technology for early warning and clinical handover systems*, Safer Better Care, March 2015, <http://www.hiqa.ie/publications/health-technology-assessment-use-information-technology-early-warning-and-clinical-hand>, last access 11 February 2016.
- [22] K.A. McKibbin, C. Lokker, S.M. Handler, L.R. Dolovich, A.M. Holbrook, D. O'Reilly et al., *Enabling Medication Management Through Health Information Technology*, April 2011, Evidence Report/Technology Assessment No. 201, prepared by the McMaster University Evidence-based Practice Center under Contract HHS-A-290-2007-10060-I. AHRQ Publication No. 11-E008-EF, Rockville MD: Agency for Healthcare Research and Quality, <http://www.ahrq.gov/research/findings/evidence-based-reports/er201-abstract.html>, last access 11 February 2016.
- [23] H. Stürzlinger, C. Hiebinger, D. Pertl, P. Traurig, *Computerized Physician Order Entry - Effectiveness and efficiency of electronic medication ordering with decision support systems*, Wirksamkeit und Effizienz elektronischer Arzneimittelverordnung mit Entscheidungsunterstützungssystemen, Köln, Deutsche Agentur Für Health Technology Assessments des Deutschen Instituts für Medizinische Dokumentation und Information (DAHTA @ DIMDI); HTA-Bericht, 2009, English Abstract: http://portal.dimdi.de/de/hta/berichte/hta228_summary_en.pdf, last access 11 February 2016.

- [24] Canadian Agency for Drugs and Technologies in Health, *Telehealth services for the treatment of psychiatric issues: clinical effectiveness, safety, and guidelines*, Rapid response report: Summary with critical appraisal. Ottawa: Canadian Agency for Drugs and Technologies in Health (CADTH), 2015, <https://www.cadth.ca/telehealth-services-treatment-psychiatric-issues-clinical-effectiveness-safety-and-guidelines>, last access 11 February 2016.
- [25] K. Kidholm, A.G. Ekeland, L.K. Jensen, J. Rasmussen et al., A model for assessment of telemedicine applications: MAST, *International Journal of Technology Assessment in Health Care* **28**(1) (2012), 44–51.
- [26] Short version of MethoTelemed project, *Methodology to assess telemedicine applications*, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=2887, last access 11 February 2016.
- [27] K. Lampe, M. Makela, M.V. Garrido, et al., The HTA Core Model: A novel method for producing and reporting health technology assessments, *International Journal of Technology Assessment in Health Care* **25** (2009), 9-20.
- [28] *Regions of Europe Working together for HEALTH*, Project homepage: <http://www.renewinghealth.eu/en/home>, last access 11 February 2016.
- [29] M. d'Angelantonio, K. Kidholm, *Learnings from EU funded projects*, Presentation at the 3rd Meeting of the HTA Network, October 29 2014, Rome, Italy, http://ec.europa.eu/health/technology_assessment/events/ev_20141029_en.htm, last access 11 February 2016.
- [30] K. Kidholm, P. Stafylas, A. Kotzeva, C.D. Pedersen et al., July 2014, D1.12 v1.5 *Renewing Health Final Project Report*, <http://www.renewinghealth.eu/documents/28946/555381f3-9686-4955-8547-76b58be34a04>, last access 11 February 2016.
- [31] E. Narne E, *How will decision makers of Azienda Ospedaliera of Padova/ Padova University Hospital use MAST in practise?* 5th May 2010, Berlin, MAST meeting, <http://docsslide.us/documents/narne-e-mast-5th-may-2010-berlin-how-will-decision-makers-of-azienda-ospedaliera-of-padova-padova-university-hospital-use-mast-in-practise.html>, last access 11 February 2016.
- [32] Renewing health, *Assessment methods*, <http://www.renewinghealth.eu/en/assessment-method>, last access 11 February 2016.
- [33] *SMARTCARE, Delivering integrated eCare*, Project homepage <http://www.pilotsmartcare.eu/home>, last access 11 February 2016.
- [34] Carewell, *Delivering integrated healthcare to frail patients through ICT*, <http://www.carewell-project.eu/home/>, last access 11 February 2016.
- [35] S. Daugbjerg, P. Stafylas on behalf of CareWell project investigators, *D7.1 Evaluation Framework*, WP7 Evidence gathering and evaluation report, Version 2.1, 9th February 2015, http://carewell-project.eu/fileadmin/carewell/_images/regions/d7.1_carewell_evaluation_framework.pdf, last access 11 February 2016.
- [36] J. Brender, *Handbook of Evaluation Methods for Health Informatics*, Elsevier Academic Press, Amsterdam, Boston, 2006.
- [37] H.G. Eichler, B. Bloechl-Daum, E. Abadie, D. Barnett et al., Relative efficacy of drugs: an emerging issue between regulatory agencies and third-party payers, *Nature Reviews - Drug Discovery* **9** (2010), 277-291.
- [38] EUneHTA Joint Action 2 (2012–2015), *The HTA Core Model® for Rapid Relative Effectiveness Assessments*, developed by Work Package 5, Version 4.1, July 2015.
- [39] EUneHTA WP5 Joint Action 2 Strand B, *Rapid assessment of other health technologies such as medical devices, surgical interventions or diagnostics*, renal denervation systems for treatment-resistant hypertension, Pilot ID: WP5-SB-12, Version 1.3, 18 December 2013.
- [40] S. Kleijnen, I. Pasternack, P. Rannanheimo, J.M. Vuola et al., Piloting international production of rapid relative effectiveness assessments of pharmaceuticals, *International Journal of Technology Assessment in Health Care* **30**(5) (2015), 521-529.
- [41] C.K. McMullen, J.S. Ash, D.F. Sittig, A. Bunce, K. Guappone, R. Dykstra, J. Carpenter, J. Richardson, A. Wright, Rapid Assessment of Clinical Information Systems in the Healthcare Setting: An Efficient Method for Time-Pressed Evaluation, *Methods Inf Med* **50**(4) (2011), 299–307.
- [42] J.S. Ash, D.F. Sittig, C.K. McMullen, A. Wright, A. Bunce, V. Mohan, D.J. Cohen, B. Middleton, Multiple perspectives on clinical decision support: a qualitative study of fifteen clinical and vendor organizations, *BMC Med Inform Decis Mak* **15**(35) (2015).
- [43] R.E. Glasgow, S.M. Phillips, M.A. Sanchez, Implementation science approaches for integrating eHealth research into practice and policy, *Int J Med Inform* **83**(7) (2014), e1-11.
- [44] R.E. Glasgow, R.S. Kessler, M.G. Ory, D. Roby et al., Conducting rapid, relevant research: lessons learned from the My Own Health Report project, *Am J Prev Med* **47**(2) (2014), 212-9.
- [45] R.E. Glasgow, R.M. Kaplan, J.K. Ockene, E.B. Fisher et al., Patient-reported measures of psychosocial issues and health behavior should be added to electronic health records, *Health Aff (Millwood)* **31**(3) (2012), 497-504.

- [46] R.E. Glasgow, R.C. Brownson, R.S. Kessler, Thinking about health-related outcomes: what do we need evidence about? *Clin Transl Sci* **6**(4) (2013), 286-91.
- [47] M. Rigby, J. Talmon, J. Brender McNair, E. Ammenwerth et al., Linking Informaticians and End Users - Using the STARE-HI Evaluation Reporting Framework as a Unifying Design Approach, *Studies in health technology and informatics* **150** (2009), 66-70.
- [48] A.R. Hardisty, S.C. Peirce, A. Preece, C.H. Bolton et al., Bridging two translation gaps: A new informatics research agenda for telemonitoring of chronic disease, *Int J Med Inform* **80**(10) (2011), 734-744.
- [49] S. Newman, *The Whole System Demonstrator Project*, Presentation at the International Congress on Telehealth and Telecare 2011, <http://www.kingsfund.org.uk/sites/files/kf/Professor%20Stan%20Newman.pdf>, last access 11 February 2016.
- [50] HTAi Patient & Citizen Involvement interest sub-group, *Patient Submission Templates for HTA – Call for Piloting/ Consultation*, July 2014, <http://www.inahta.org/2014/07/patient-templates>, last access 11 February 2016.
- [51] N. May, M. Godfrey, *Who has the say in Health Technology Assessment?* Report, Hill & Knowlton, London, 2010, <http://vortal.hbai.org/?q=node/173>, last access 11 February 2016.
- [52] J. Talmon, E. Ammenwerth, J. Brender, N. de Keizer et al., STARE-HI–Statement on reporting of evaluation studies in Health Informatics, *Int J Med Inform* **78**(1) (2009), 1-9.
- [53] P. Nykänen, J. Brender, J. Talmon, N. de Keizer et al., Guideline for good evaluation practice in health informatics (GEP-HI), *Int J Med Inform* **80**(12) (2011), 815-27.
- [54] SALUS project - *Scalable, Standard based Interoperability Framework for Sustainable Proactive Post Market Safety Studies*, <http://www.salusproject.eu>, last access 11 February 2016.
- [55] TRANSFoRm project, *Developing a 'rapid learning healthcare system'*, <http://www.transformproject.eu>, last access 11 February 2016.
- [56] *Electronic Health Records for Clinical Research – EHR4CR project*, <http://www.ehr4cr.eu>, last access 11 February 2016.
- [57] NordForsk, *Joint Nordic Registers and Biobanks - A goldmine for health and welfare research*, NordForsk Policy Paper 5, 2014, http://www.nordforsk.org/en/publications/publications_container/joint-nordic-registers-and-biobanks-a-goldmine-for-health-and-welfare-research, last access 11 February 2016.
- [58] M. Meglić, P. Doupi, I. Pristăs, Y. Skalkidis et al., PARENT Joint Action: increasing the added value of patient registries in a cross-border setting, *Stud Health Technol Inform* **192** (2013), 1161.
- [59] P. Doupi, W. Goettsch, M. Klemp, I. Pristas, *Patient registries as HTA tools in economic outcomes research: requirements, barriers, and the way forward*, Workshop overview, ISPOR 17th Annual European Congress, Amsterdam, The Netherlands, 2014, <http://www.ispor.org/Event/GetReleasedPresentation/124>, last access 11 February 2016.
- [60] University of Oxford, *Patient Reported Outcome Measurement Group*, http://phi.uhce.ox.ac.uk/inst_types.php, last access 11 February 2016.
- [61] T.H. Weldring, S.M.S. Smith, Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs), *Health Serv Insights* **6** (2013), 61-68.
- [62] Office of the National Coordinator for Health Information Technology, *Health IT Safety Center Roadmap*, prepared by RTI International, July 2015, <http://www.healthitsafety.org/uploads/4/3/6/4/43647387/roadmap.pdf>, last access 11 February 2016.
- [63] WHO, *Patients for Patient Safety programmes*, http://www.who.int/patientsafety/patients_for_patient/en, last access 11 February 2016.
- [64] F. Muench, The Promises and Pitfalls of Digital Technology in Its Application to Alcohol Treatment, *Alcohol Res* **36**(1) (2014), 131-142.
- [65] mHealth Competence Center, *mHealth Assessment Workshop Barcelona*, 25th February 2015, Workshop Discussion Summary, <http://www.mobilehealthglobal.com/media-resources/mhealth-cc-working-papers/2/mhealth-assessment-workshop>, last access 11 February 2016.
- [66] R. Thiel, M. Viceconti, K. Stroetmann, Assessing Biocomputational Modelling in Transforming Clinical Guidelines for Osteoporosis Management, in: A. Moen et al. (Eds.), *User Centred Networked Health Care*, Proceedings of MIE Europe Conference 2011, IOS Press, 2011. pp. 432-436.

Realizing the Potential of Patient Engagement: Designing IT to Support Health in Everyday Life

Laurie L. NOVAK^{a,1}, Kim M. UNERTL^a and Richard J. HOLDEN^b

^a*Department of Biomedical Informatics, Vanderbilt University Medical Center, USA*

^b*Department of BioHealth, Indiana University School of Informatics and Computing, Indianapolis, USA*

Abstract. Maintaining health or managing a chronic condition involves performing and coordinating potentially new and complex tasks in the context of everyday life. Tools such as reminder apps and online health communities are being created to support patients in carrying out these tasks. Research has documented mixed effectiveness and problems with continued use of these tools, and suggests that more widespread adoption may be aided by design approaches that facilitate integration of eHealth technologies into patients' and family members' daily routines. Given the need to augment existing methods of design and implementation of eHealth tools, this contribution discusses frameworks and associated methods that engage patients and explore contexts of use in ways that can produce insights for eHealth designers.

Keywords. eHealth, community based participatory research, qualitative research, patient engagement.

1. Introduction

Maintaining health or managing a chronic condition involves performing and coordinating potentially new and complex tasks in the context of everyday life. Activities such as medication management, exercise, implementing new dietary recommendations, and monitoring health indicators must be conducted in addition to the pre-existing activities of everyday life, e.g. preparing meals, caring for children, and working. Tools such as reminder apps and online health communities are being created to support patients in carrying out these activities [1, 2]. These tools, referred to as eHealth, have been defined by the World Health Organization in this way [3]:

E-health is the transfer of health resources and health care by electronic means. It encompasses three main areas:

1. *The delivery of health information, for health professionals and health consumers, through the Internet and telecommunications.*
2. *Using the power of IT and e-commerce to improve public health services, e.g. through the education and training of health workers.*
3. *The use of e-commerce and e-business practices in health systems management.*

¹ Corresponding author: Assistant Professor Dr. Laurie Lovett Novak, Department of Biomedical Informatics, Vanderbilt University Medical Center, 2525 West End Avenue, Nashville, TN 37203, United States, Laurie.l.novak@vanderbilt.edu.

Previous studies have documented mixed effectiveness and problems with continued use of eHealth tools [4-9]. Jimison and colleagues [10] showed that more widespread adoption may be aided by design approaches that facilitate integration of eHealth technologies into patients' and family members' daily routines. Given the need to augment existing methods of design and implementation of eHealth tools [11, 12], methods and frameworks are needed that engage patients and explore contexts of use with the goal of producing insights for eHealth designers. Key questions include: Who is acting? What are the activities? How are the activities structured? Which information infrastructures can contribute to informatics solutions?

This contribution reviews two research frameworks and a design method that may be used together or separately to increase the fidelity of design specifications to the actual needs of patients, family members and other participants as they attempt to integrate health-related activities into everyday life.

2. Patient Work: A Focus on Activity

A persistent challenge in patient engagement research methodology is understanding health and chronic illness management *in context*, i.e., as it actually occurs in the patient's home and community. Patients' homes and communities are the "frontlines" of health and illness-related activities. However, due to the difficulty of doing research in these contexts and a lack of methods for doing so effectively, with a few exceptions [13, 14], prior work has taken research on everyday self-management out of context, focusing primarily on the skills and capacities of the individual patient. Thus, our current understanding of self-management inadequately accounts for the full complexity and dynamics of the context in which it is carried out.

One approach for addressing this challenge is to adapt existing methods for studying activity that have been developed and used in the study of work practices in industrial settings. The Patient Work concept [15, 16] has roots in social science [17] and industrial engineering [18], and holds that the health-related activities of patients, family members, and other lay caregivers constitute a type of work, defined as "exertion of effort and investment of time on the part of patients or family members to produce or accomplish something." [19]. Patient work can be similar or analogous to the work of health care professionals. For example, tracking medications and arranging "handovers" to another caregiver are tasks that parents of children with chronic illness routinely perform. Patient work occurs within a *context* (or "work system") that comprises interacting structural components such as task, technology, environment, and community factors. These factors act as constraints, facilitators, or both, with respect to patient work activity.

Methods for Patient Work research include general approaches such as interviews and observation, methods from work sciences such as cognitive task analysis [20] or rapid ethnography [21], and emerging, technologically-mediated methods such as Ecological Momentary Assessment (EMA), a method that involves repeated sampling of participant's activities in real time [22]. Table 1 provides examples of key considerations for application of these methods in Patient Work research.

Table 1. Methods amenable to investigating Patient Work and considerations for their use in the patient work domain.

Methods	Considerations for the patient work domain
Traditional Research Methods	
Interviews	If multiple actors are involved, how should the perceptions of each be captured - separate or joint interviews?
Observation: structured, semi-structured, and ethnographic	Conducting observations of private or infrequent activities.
Standardized surveys	Appropriate tailoring of language.
Focus group / group interview	Maintaining comfort and privacy about personal medical issues.
Document analysis	Are documents available and legible?
Experimental trial	Isolating behavioral interventions to one group when experimental and control groups are socially connected
Work Study Methods	
Cognitive task analysis	Determining level of informant expertise.
Incident/accident analysis	Will self-reported causes be inaccurate or incomplete?
Process mapping	How to portray complex processes crossing boundaries of health and everyday life?
Critical incident technique	How will memory of events be preserved in old or young individuals?
Cognitive work analysis	Where is domain expertise found? Clinicians, patients, lay caregivers, or all?
Macroergonomic work analyses	Identifying appropriate work system models for application to unpaid, community-based work.
Simulation modeling	How to handle outside sources of variability, e.g., personal life changes?
Assessment of workload and situation awareness	How to measure without affecting workload and situation awareness themselves?
Participatory design	Balancing participants' input and expectations with final design elements.
Emerging Patient-oriented Methods	
Ecological momentary assessment	How to implement when lacking internet access?
Diary methods	Overtaxing participants while ensuring participation. Data management of paper and electronic diaries.
Online group / social network analysis	Maintaining privacy and confidentiality.
Sensor-based monitoring	Maintaining sensor networks when hardware problems arise. Ethics in research design.

Methods from the work sciences tend to be too generic or else rooted in their domain of origin (e.g., aviation), requiring adaptation to the patient work domain. For example, interviews may need to be focused based on a theoretic lens such as illness trajectory [23] or illness narrative [24] in order to capture the full relevant experience of a patient and his or her family and friends. Cognitive task analysis for patient work may need to accommodate the possibility that expertise on performing a health-related process such as medication management is distributed across actors and artifacts; thus, a complete task analysis requires observations of patients, informal caregivers, clinicians, and various paper and electronic tools, across many settings [25]. Methods that are being developed specifically for collecting patient data also require further development and adaptation to specific contexts. For instance, using portable accelerometry devices to track the activities and step counts of older adults with physical disabilities is complicated by the relative inactivity of these individuals; the potential disuse of wearable devices due to

loss or forgetting; and variability in gait and pace that may render inaccurate step-counting algorithms. A different set of issues may affect the use of the same devices to study children, for example, concerns about privacy or disabling of the device due to battery use.

Patient Work research seeks to describe and theorize the activities engaged in by patients and other participants in health or chronic illness management, and in this way is similar to Burden of Treatment Theory [26]. Descriptions and frameworks that emerge from this research can point to a range of eHealth development opportunities, including infrastructural requirements, specific information needs, and interface design needs. For example, designers may use the research to identify data sources to enable offering clinic appointments linked to transportation schedules. This need has been recognized in international research policy development. The 2011 OECD-NSF Workshop on Building a Smarter Health and Wellness Future called for investigators to “look at data outside the health domain and link population data from different sources to better understand environmental determinants of nutritional illness, stress, mental health.” [27]

Investigations of Patient Work delineate the roles of specific actors in the overall system. This enables moving beyond traditional classifications e.g. “family” or “spouse” toward more functional roles e.g. “medication administration” or “transportation to clinic appointments.” Delineation of specific roles can provide useful input to the design of roles for privacy protection. For example, the person who drives the patient to the clinic may benefit from having access to the patient’s appointment times, but may *not* need to know information such as diagnoses or medication prescriptions. These information needs, along with those of formal caregivers and health care providers, produce what the European Science Foundation has referred to as “overlapping domains of confidentiality” [28]. In order for eHealth applications to be regarded and used with trust, these domains must be described and built into systems.

Finally, rich studies of Patient Work can provide insight into relationships among illness activities and structures of everyday life. These relationships can be temporal e.g. the timing of the school day structuring the medication management of a teen with asthma. The relationships can also be spatial, e.g. neighborhood design and physical activity, or functional, e.g. the role of material artifacts (backpacks, medication organizers, inhalers, glucometers) in facilitating everyday adherence.

3. Community-Based Participatory Research: Reaching Under-Represented Individuals

Building a base of robust evidence to support innovative developments in eHealth presents challenges, particularly in access to data and participants. Unlike work practice research conducted in hospitals and ambulatory clinics, researchers studying patient activity do not typically have uncomplicated access to patient homes, schools, workplaces and other community settings. In some cases, prior negative interactions with research institutions can lead to feelings of distrust between community members and academic institutions seeking to conduct research. In addition, individuals from under-represented groups, such as racial/ethnic minority populations or low socioeconomic status, are often under-represented in research activities, leaving researchers with an incomplete perspective on depth and breadth of patient work [29]. The current result is inadequate published evidence; hence innovators (service developers or technical developers) need to study the requirements – in a way which is robust and unbiased.

Studying patient work poses inherent challenges. Community-based participatory research (CBPR) seeks to address issues of access, trust, and representation through building relationships between researchers and the community throughout all stages of research [30]. While the CPBR model has room for the use of multiple methodologies, the core argument of the approach is that community members and partner community organizations need a voice in research design, implementation, and dissemination. In a CBPR-oriented research project, community partners participate in defining research questions of interest to the community, assist with designing context-appropriate participant recruitment strategies, and may participate in data collection and data analysis activities. In addition, results of CBPR-oriented projects are presented not just in academic venues, but using alternate dissemination approaches such as community meetings and social media. Partners representing community groups might also be involved in presenting research results, both in academic and non-academic venues [31].

Researchers and communities can use CBPR to gain a more complete picture of patient work from diverse perspectives. Approaching patient work research from a CBPR perspective can assist researchers with gaining access to contexts and groups that might be inaccessible when viewing research strictly from an academic perspective.

CBPR is not a specific methodology, but rather a theoretical orientation to research that emphasizes meaningful partnerships between researchers and communities. CBPR in consumer health IT applies methods from user-centered design and participatory design fields² and also approaches from the social sciences and ethnography such as observation and interviews [32]. Unertl and colleagues examined projects that integrated CBPR and informatics design, recommending eight principles: 1) Viewing community as a unit of identity, 2) Understanding the existing strengths and resources within the community, 3) Building collaborative partnerships in all research phases, 4) Integrating research results for mutual benefit, 5) Viewing research and partnership building as a cyclical and iterative process, 6) Empowering both academic and community partners through co-learning opportunities, with awareness of social inequalities, 7) Incorporating positive and ecological perspectives into research, 8) Disseminating knowledge to all partners [33].

Although application of CBPR is still in the early stages to health informatics research, studies have shown significant promise for improving the fit between technology interventions and patient needs [34, 35].

4. Participatory Design Methodologies: Listening to the Voices of the Intended Users

Once researchers have access to patient work contexts, identifying appropriate methodologies to gather relevant data is critical. Design Science Research focuses on a systematic view of technology in use, rather than separated from use [36]. Design Science methodologies such as participatory design workshops seek to increase the amount and quality of input from intended end users into technology design processes [37]. A group of methodologies especially well-suited to understanding patient contextual factors are participatory design methods. Using participatory design methods,

² See also: A. Kushniruk et al., Participatory design and health IT evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

researchers seek input on the emerging design of an artifact, process, service or other entity [38-40]. Participatory design shares common elements with user-centered design, but places more emphasis on co-designing products, technology, or services with intended end users. Participatory design methods go beyond merely asking patients what they want technology to look like or needed functions. Rather, participatory design methods involve intended end users in hands-on activities to help intended end users think about their needs and how technology products might assist with meeting those needs. The emphasis on collaboration in participatory design methods can result in products tailored to an individual's or group's interests and needs. These approaches seem especially well suited to patient contexts, where researcher- or technology-focussed perspectives inherent in technology can result in a poor fit to patient needs [14].

A commonly used participatory design approach is a design workshop. Design workshops differ from focus groups in the degree of active participation and the collaborative nature of the work [37]. One approach to design workshops incorporates three activities: priming, designing, and debriefing. The priming activity helps to set the stage for the design activity by engaging participants in the topic and encouraging participants to begin thinking about their experiences in new ways. For example, participants could be asked to fill out a worksheet about their experience managing a chronic disease or to take photos of places, people, and resources in their community that either are barriers to or facilitators for healthy behavior choices.

The design activity can take many different formats, such as creating a paper prototype for a human-computer interface or developing ideas for processes that could be used to support health-related activities. The focus of the design activity is on looking beyond narrow constraints and engaging in creative development of solutions to problems or questions. Specific types of design activities serve different purposes, so thoughtful selection of an appropriate design activity for the research questions should be considered. Probing activities [41] can be used to explore current experiences and expectations, while generative activities [42] can be used for the co-design of technology or processes. Generative toolkits are used in design science to provide a set of materials that participants can build on while thinking about how they would design a process or a product [43]. Design activities are typically very hands-on and action-oriented, allow participants to actively engage in design based on their expertise.

Finally, the debriefing activity focuses on understanding participant perspectives about artifacts or concepts created in the design activity. Debriefing activities could take various formats, such as individual or small group interviews or having individual participants reporting back to the full group about the artifacts they have created.

The methods used in design science and participatory design often produce extensive amounts of multimedia data, requiring careful consideration of efficient and effective approaches to data analysis. Although workshop participants are experts on the topic being explored (e.g. self-management of a chronic disease; use of technology in managing health), eliciting theoretical concepts from the dataset and acting on the design concepts developed by participants are the responsibility of researchers. Participatory design research is also frequently iterative, with future design workshop cycles building on experiences from participants in prior workshops.

There are multiple examples of applying design science and participatory design approaches to health information technology design and development. Byrne and Gregory reported on a large-scale participatory design project in rural South Africa in which “shared ground” was achieved through design shops among national government officials, local community members and researchers in order to build a system of

indicators that reflected child health in the region [44]. Other projects have used PD in both small-scale [45] and large-scale projects [46].

5. How can eHealth IT stakeholders benefit from participatory and patient-engaged methods?

Stakeholders in eHealth include patients, their families and caregivers, health care providers (including hospitals, clinics, pharmacy, etc.), health care purchasers, technology developers and device manufacturers, and policy makers. All of these can benefit from patient-engaged approaches to research and design of eHealth tools.

Patients, families and caregivers – rigorous and detailed description of barriers and facilitators to the everyday routines of health maintenance and chronic illness management creates opportunities for technology designers to develop tools that can be tailored to the needs of these actors. Better understandings of the information infrastructures involved in these activities can enable time-saving, safety-enhancing links between them, e.g. health care records, transportation schedules, school systems, pharmacies, grocery stores, family members, social media, and medical such as glucometers, pill boxes, and inhalers.

Health care purchasers – governments and other purchasers gain benefits from patient engagement technologies through improved tailoring of and adherence to therapeutic goals and resulting reduced cost to organizations and society, and also through improved efficiencies gained by automation of previously manual tasks such as filing insurance claims.

Technology developers and device manufacturers – these stakeholders gain through increased acceptance and use of tools that are co-designed to meet users' needs.

Policy makers – these stakeholders will be presented with new challenges from eHealth. While the potential financial benefits of a healthier population are real, policy makers will be required to engage in challenging analysis of issues related to the ongoing digitization of health and related information. Capitalists in data science will push for increased access to the details of life and health of individuals. Simultaneously, privacy and security will be increasingly at risk of malicious attack.

6. Challenges in implementation

Methodologies

As with any instance of field research, the above approaches require not only valid and reliable methods but also an appropriate implementation [47]. An analysis of implementation challenges from two studies of Patient Work in home and community setting [48] found four categories of challenges, related to:

- Researcher-participant partnership (e.g., mutual trust, common ground)
- Participant characteristics (e.g., patients' cognitive limitations, lack of participant transportation for research visits)
- Research logistics and procedures (e.g., travel distances, problems recruiting); and
- Scientific quality and interpretation (e.g., combining data from multiple sources).

In CBPR and participatory design, issues of partnership and balancing the researcher-participant relationship are especially important, given the responsibilities held by community stakeholders [49]. In our experience, the presence of a third party such as a healthcare delivery entity or local nonprofit makes for an even bigger challenge, introducing additional regulations, priorities, and institutional history with the community.

Translation into design

While examples of effective use of the methods exist, skills for translating results of participatory design and Patient Work studies into actionable eHealth design and implementation specifications are not widespread. Findings from these studies can potentially impact interaction design, information infrastructure negotiation and access, security and privacy design elements, connectivity with traditional EHR data, and dissemination strategies. Such skills can be developed and disseminated through professional and educational networks.

7. Conclusion

Developing a base of evidence to support design and implementation decisions in eHealth is essential for ensuring the safe, effective, and efficient deployment of these technologies. Stakeholder engagement through participatory methods and user studies that focus on Patient Work activities produce information that can be disseminated and used as guidance for design. Funders that support these activities may include corporate and government entities. We recommend that funders and professional societies support the further development of frameworks, methods, and training and dissemination infrastructure to enable widespread adoption of these approaches, which will in turn provide the best representation of a range of patient needs to the designers and implementers of tools.

Recommended further readings

1. J. Simonsen, T. Robertson, *Routledge International Handbook of Participatory Design*, Routledge, New York, 2013.
2. G. Eysenbach, H. Jimison, R. Kukafka, D. Lewis, P.Z. Stavri, (Eds.), *Consumer Health Informatics*, Springer, New York, 2005. pp. 34-60.
3. B.A. Israel, E. Eng, A.J. Schulz, E.A. Parker, *Methods for Community-Based Participatory Research for Health*, Jossey-Bass, San Francisco, 2nd edition, 2012.
4. J. Corbin, A.L. Strauss, Managing Chronic Illness at Home: Three Lines of Work, *Qualitative Sociology* **8(3)** (1985), 224-47.
5. P. Carayon (Ed.), *Handbook of human factors and ergonomics in health care and patient safety*, CRC Press, Boca Raton, 2011.

Food for thought

1. How is a patient work approach to consumer health informatics design different from biomedical and behavioral change approaches?
2. You are developing a smartphone app to help teens with asthma manage their medications. What skills do you need on your development team?
3. Consider a health problem that you or a friend lives with. What are the Patient Work tasks involved in managing the health issue? Where does the information come from to make decisions? What are the cues to action?
4. Who are the stakeholders of eHealth technology, and what are their priorities?

References

- [1] B. Kaplan, P. F. Brennan, Consumer informatics supporting patients as co-producers of quality, *J Am Med Inform Assoc* **8** (2001), 309-16.
- [2] K. D. Mandl, I. S. Kohane, Tectonic shifts in the health information economy, *N Engl J Med* **358** (2008), 1732-7.
- [3] WHO, *WHO – E-Health*, <http://www.who.int/trade/glossary/story021/en>, last access 11 February 2016.
- [4] L. Beatty, S. Lambert, A systematic review of internet-based self-help therapeutic interventions to improve distress and disease-control among adults with chronic health conditions, *Clin Psychol Rev* **33** (2013), 609-22.
- [5] G. Eysenbach, The law of attrition, *J Med Internet Res* **7** (2005), e11.
- [6] C. Free, G. Phillips, L. Galli, L. Watson, L. Felix, P. Edwards, et al., The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review, *PLoS Med* **10** (2013), p. e1001362.
- [7] T. D. Giardina, S. Menon, D. E. Parrish, D. F. Sittig, H. Singh, Patient access to medical records and healthcare outcomes: a systematic review, *J Am Med Inform Assoc* **21**(4) (2014), 737-41.
- [8] P. Wicks, D. L. Keininger, M. P. Massagli, C. de la Loge, C. Brownstein, J. Isojärvi, et al., Perceived benefits of sharing health data between people with epilepsy on an online platform, *Epilepsy Behav* **23** (2012), 16-23.
- [9] C. H. Yu, R. Bahlaiwal, A. Laupacis, E. Leung, M. S. Orr, S. E. Straus, Systematic review and evaluation of web-accessible tools for management of diabetes and related cardiovascular risk factors by patients and healthcare providers, *J Am Med Inform Assoc* **19** (2012), 514-22.
- [10] H. Jimison, P. Gorman, S. Woods, P. Nygren, M. Walker, S. Norris, et al., Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved, *Evid Rep Technol Assess (Full Rep)* (2008), 1-1422.
- [11] J. L. Marquard, T. Zayas-Cabán, Commercial off-the-shelf consumer health informatics interventions: recommendations for their design, evaluation and redesign, *J Am Med Inform Assoc* **19** (2012), 137-42.
- [12] T. Zayas-Cabán, B. E. Dixon, Considerations for the design of safe and effective consumer health IT applications in the home, *Qual Saf Health Care* **19** (Suppl 3) (2010), i61-7.
- [13] S. A. Mulvaney, R. L. Rothman, M. S. Dietrich, K. A. Wallston, E. Grove, T. A. Elasy, et al., Using mobile phones to measure adolescent diabetes adherence, *Health Psychology* **31**(1) (2012), 43.
- [14] N. Ramanathan, D. Swendeman, W. S. Comulada, D. Estrin, M. J. Rotheram-Borus, Identifying preferences for mobile health applications for self-monitoring and self-management: Focus group findings from HIV-positive persons and young mothers, *International Journal of Medical Informatics* **82** (2013) e38-e46.
- [15] R. S. Valdez, R. J. Holden, L. L. Novak, T. C. Veinot, Transforming consumer health informatics through a patient work framework: connecting patients to context, *Journal of the American Medical Informatics Association* **22**(1) (2015), 2-10.
- [16] R. J. Holden, C. C. Schubert, R. S. Mickelson, The patient work system: An analysis of self-care performance barriers among elderly heart failure patients and their informal caregivers, *Applied Ergonomics* **47** (2015), 133-150.
- [17] A. L. Strauss, S. Fagerhaug, B. Suczek, C. Wiener, The work of hospitalized patients, *Social Sience and Medicine* **16**(9) (1982), 977-986.

- [18] R. J. Holden, P. Carayon, A. P. Gurses, P. Hoonakker, A. S. Hundt, A. A. Ozok, et al., SEIPS 2.0: a human factors framework for studying and improving the work of healthcare professionals and patients, *Ergonomics* **56** (2013), 1669-1686.
- [19] A. L. Strauss, *Continual permutations of action*, Aldine de Gruyter, New York, 1993.
- [20] L. G. Militello, R. J. B. Hutton, Applied Cognitive Task Analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands, *Ergonomics* **41** (1998), 1618-1641.
- [21] D. R. Millen, *Rapid Ethnography: Time Deepening Strategies for HCI Field Research*, presented at the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques, New York, NY, 2000.
- [22] S. Shiffman, A. A. Stone, M. R. Hufford, Ecological momentary assessment, *Annu Rev Clin Psychol* **4** (2008), 1-32.
- [23] J. Corbin, A. L. Strauss, Managing Chronic Illness at Home: Three Lines of Work, *Qualitative Sociology* **8** (1985), 224-247.
- [24] C. Mattingly, L. C. Garro, *Narrative and the cultural construction of illness and healing*, Univ of California Press, 2000.
- [25] R. S. Mickelson, M. Willis, R. J. Holden, Medication-related cognitive artifacts used by older adults with heart failure, *Health Policy & Technology* **4**(4) (2015), 387-398.
- [26] C. R. May, D. T. Eton, K. Boehmer, K. Gallacher, K. Hunt, S. MacDonald, et al., Rethinking the patient: using Burden of Treatment Theory to understand the changing dynamics of illness, *BMC health services research* **14** (2014), 281.
- [27] M. Rigby, E. Ronchi. *OECD-NSF Workshop: Building a Smarter Health and Wellness Future - Summary of Key Messages*. 2011. <http://www.oecd.org/sti/economy/oecd-nsfworkshopbuildingasmarterhealthandwellnessfuture.htm>, last access 11 February 2016.
- [28] M. Rigby, S. Koch, D. Keeling, P. Hill, A. Alonso, E. Maeckelbergh, *Developing a New Understanding of Enabling Health and Wellbeing in Europe: Harmonising Health and Social Care Delivery and Informatics Support to Ensure Holistic Care*, European Science Foundation: Standing Committee of the Social Sciences, 2013.
- [29] C. Showell, P. Turner, The PLU problem: are we designing personal ehealth for people like us? *Stud Health Technol Inform* **183** (2013), 276-80.
- [30] B. Israel, E. Eng, A. Schulz, E. Parker, Introduction to methods in CBPR for health, in *Methods for community-based participatory research for health*, 2nd edition, Jossey-Bass, 2012, pp. 3-26.
- [31] P. G. Chen, N. Diaz, G. Lucas, M. S. Rosenthal, Dissemination of results in community-based participatory research, *American Journal of Preventive Medicine* **39** (2010), 372-378.
- [32] K. M. Unertl, C. L. Schaefbauer, T. Campbell, C. Senteio, K. A. Siek, S. Bakken, et al., Integrating community-based participatory research and informatics approaches to improve the engagement and health of underserved populations, *J Am Med Inform Assoc* (2015), 1-15. <http://dx.doi.org/10.1093/jamia/ocv094>.
- [33] K. M. Unertl, C. L. Schaefbauer, T. Campbell, C. Senteio, K. A. Siek, S. Bakken, et al., Integrating community-based participatory research and informatics approaches to improve the engagement and health of underserved populations, *J Am Med Inform Assoc* (2015), 1-15. <http://dx.doi.org/10.1093/jamia/ocv094>.
- [34] S. Grieb, Development of a tailored mHealth platform to improve HIV testing and linkage to care in a new immigrant receiving community, in *142nd APHA Annual Meeting and Exposition* (November 15-November 19, 2014), 2014.
- [35] R. de la Harpe, H. Lotriet, D. Pottas, M. Korpela, Socio-technical approach to community health: designing and developing a mobile care data application for home-based healthcare, in South Africa, *The Journal of Community Informatics* **9**(2) (2012).
- [36] P. Johannesson, E. Perjons, A Method Framework for Design Science Research, in: *An Introduction to Design Science*, Springer, 2014, pp. 75-89.
- [37] J. Simonson, T. Robertson, *Routledge international handbook of participatory design*, Routledge, 2012.
- [38] J. Clemensen, S. B. Larsen, M. Kyng, M. Kirkevold, Participatory design in health sciences: using cooperative experimental methods in developing health services and computer technology, *Qualitative health research* **17** (2007), 122-130.
- [39] D. Schuler, A. Namioka, *Participatory Design: Principles and Practice*, LEA, Hillsdale, NJ, 1993.
- [40] J. Gregory, Scandinavian approaches to participatory design, *International Journal of Engineering Education* **19** (2003), 62-74.
- [41] E. B.-N. Sanders, P. J. Stappers, Probes, toolkits and prototypes: three approaches to making in codesigning, *CoDesign* **10** (2014), 5-14.
- [42] G. N. Vavoula, M. Sharples, Future technology workshop: A collaborative method for the design of new learning technologies and activities, *International Journal of Computer-Supported Collaborative Learning* **2** (2007), 393-419.

- [43] E. B. N. Sanders, P. J. Stappers, Co-creation and the new landscapes of design, *Co-design* **4** (2008), 5-18.
- [44] E. Byrne, J. Gregory, Co-constructing local meanings for child health indicators in community-based information systems: The UThukela District Child Survival Project in KwaZulu-Natal, *International Journal of Medical Informatics* **76S** (2007), S78-S88.
- [45] J. H. Gennari, M. Reddy, Participatory design and an eligibility screening tool, in: *Proceedings of the AMIA symposium*, 2000. p. 290.
- [46] S. Pilemalm, T. Timpka, Third generation participatory design in health informatics - making user participation applicable to large-scale information system projects, *Journal of Biomedical Informatics* **41** (2008), 327-339.
- [47] R. J. Holden, C. K. Or, S. J. Alper, A. Joy Rivera, B. T. Karsh, A change management framework for macroergonomic field research, *Appl Ergon* **39** (2008), 459-74.
- [48] R. J. Holden, A. M. McDougald Scott, P. L. T. Hoonakker, A. S. Hundt, P. Carayon, Data collection challenges in community settings: Insights from two field studies of patients with chronic disease, *Quality of Life Research* **24** (2015), 1043-1055.
- [49] B. A. Israel, A. J. Schulz, E. A. B. Becker, Review of community based research: Assessing partnership approaches to improve public health, *Annual Review of Public Health* **19** (1998), 173-202.

Ensuring Evidence-Based Safe and Effective mHealth Applications

Bárbara VALLESPIN^{a,1}, Joan CORNET^a and Anna KOTZEVA^b

^a*Mobile World Capital, Barcelona, Spain*

^b*Agency for Health. Quality and Assessment of Catalonia*

Abstract. The Internet and the digitalization of information have brought big changes in healthcare, but the arrival of smartphones and tablets represent a true revolution and a new paradigm is opened which completely changes our lives. In order to validate the impact of these new technologies in health care, it is essential to have enough clinical studies that validate their impact in wellbeing and healthcare of the patient. Traditional regulatory organisations are still looking for their role in this area. If they follow the classical path of medical devices, we get to a technical, administration and economic collapse. This contribution first presents the main indicators showing the potential of mHealth adoption. It then proposes a classification of mobile health care apps, and presents frameworks for mHealth evaluation. Regulation of mHealth as part of the evaluation process is discussed. Finally, the necessary steps and challenges that have to be taken into account by the industry to prepare the entrance of these technologies into the EU market is analysed.

Keywords. mHealth, regulation, clinical evidence, framework, evaluation, smartphones.

1. mHealth as a transformative factor of care delivery in healthcare systems

Healthcare is based in a wide sense both on data and information. Up until recently it is hospitals and healthcare providers who have obtained and held this information, which has not been accessible to the people to whom it relates. Information to the citizen and patient on how to lead a healthy life has come from professionals, or in general advice texts. However, a transformation is now occurring whereby citizens can create and interpret large volumes of data to enable them to ensure a healthy lifestyle, as well as to interact with healthcare providers.

With the advent of Internet, called by Manuel Castell “*the information society*” [1], a new paradigm is opened which changes the way we live, work, communicate and enjoy our free time. The Internet and the digitalization of information have brought big changes, but the arrival of smartphones and tablets represents a true revolution. Everyone can be connected no matter where he/she is located. Information follows the person. Nowadays, most of the citizens in the world have a tool which has more capacity of data processing than those computers from decades ago that took up a whole floor of a building and cost dozens of millions of dollars. And above all, most of the

¹ Corresponding author: Bárbara Vallespín, mHealth Manager, Mobile World Capital, MediaTic – Roc de Boronat 117, 08018 Barcelona, Spain, bvallespin@mobileworldcapital.com.

population has an emotional connection to their smartphones. We only have to see how much we stress out when we forget our phone at home or when the battery goes dead and we have no chance to plug it in order to recharge it.

The number of mobile connections and subscribers to phone services is growing exponentially. By 2020, there will be 6.100 million people using smartphones, while nowadays (2015) there are 2.600 million [2][3]. The number of smartphone users is growing and also the number of healthcare applications.

In order to look at the evidence for safe mHealth applications, it is necessary to define what we understand by “Mobile Health” (mHealth). Below there are three different definitions of how mHealth is perceived from different perspectives:

- *“mHealth seeks to improve individuals’ health and wellbeing by continuously monitoring their status, rapidly diagnosing medical conditions, recognizing behaviours, and delivering just in time interventions, all in the user’s natural environment”* [4].
- *“Medical and public health practices supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants and other wireless devices”* [5].
- *“mHealth also includes lifestyle & wellbeing applications, personal guidance systems, health information and medication reminders and telemedicine provided wirelessly”* [21].

mHealth covers different aspects of health, wellness, prevention, education, diagnostic, monitoring (with the use of wearables), follow up (treatment adherence) and contributes a new dimension to the collection of large amounts of data. Therefore we are not talking about using new technologies in healthcare, but how mobile technologies can help in the process of healthcare delivery transformation, covering various illnesses such as diabetes, heart failure, COPD, hypertension and mental health. But above all mHealth can open the door to personalized medicine, empowering the patient/citizen by providing bigger responsibility in the management of their health or condition.

One of the weak points of mHealth is the lack of sufficient number of empirical studies that validate their impact on wellbeing and health of the patient. Too frequently, studies on mHealth solutions have been based on “How can these technologies be introduced in the healthcare system?” instead of “How the healthcare systems can be more sustainable, more secure and efficient with the help of mHealth technologies?”

Lack of wide-spread agreement among experts on common research methods for mHealth assessment hinders the generation of reliable and comparable knowledge regarding the impact of mobile innovations. Also, many evaluations performed are based on specific disease groups, which limits their generalisability. Some evaluation studies provide neutral or negative results on the impact of eHealth; however, often, evaluations were conducted on pilots without having implemented the necessary organisational changes. Other evaluation studies point to positive impact of mHealth [5].

mHealth applications have grown exponentially. There are above 100.000 healthcare applications in the market including wearables, monitoring devices and others, getting to a “tsunami” of technologies that day by day invade the market. On the other hand, traditional regulatory organisations or medical evaluation institutions are still looking for their role in this area. If they follow the classical regulatory path of medical devices for mHealth, we may get to a technical, administration and economic

collapse, as the duration of the clinical trials enforced by regulation and the high costs associated make classical procedures unfeasible. Even more, mHealth solutions need regular updates which should also follow long and costly regulatory procedures. The FDA (Food and Drug Association, U.S.) has shown a cautious position by stating that there are “no binding recommendations” for mHealth regulation [13]. The European Commission has been also cautious and has initiated the mHealth Green Paper 2014, which consists of an open consultation without any recommendation as a result [21].

The lack of clear rules or guidelines for mHealth regulation is producing uncertainty in the industry and also lack of confidence of healthcare professionals. We need therefore to look for creative and innovative ways to create mHealth evaluation. At Mobile World Capital, we are working on this direction, searching consensus on a common framework for mHealth assessment among the different countries and European regions, through Medical Evaluation Agencies or similar organizations around Europe.

2. Socioeconomic impact and market readiness of mHealth

It is claimed in a report by PWC from June 2013 about the socio-economic impact of mHealth [14], “mHealth could save 99 billion EUR in healthcare costs in the European Union (EU) and add 93 billion EUR to the EU GDP in 2017 if its adoption is encouraged” (Figure 1). We might agree or not in this figure, but when thinking about these numbers, there is no doubt that it is worth giving mHealth a fair opportunity to realise and validate its potential.



Figure 1. Economic impact of mHealth Fehler! Verweisquelle konnte nicht gefunden werden.

Aging population and chronic diseases are a major problem throughout not only Europe, but also around the globe [16]. As people become more aware of their condition and become more informed on their diseases through the technologies that provide availability to medical information, they can and should also start taking a much more active role in the management of their disease. Also, more and more people start focussing on wellness and prevention, which leads to good healthy habits, helps to avoid certain practices that are well known as triggers of certain future diseases, and provides citizens with tools, services and products that can help them to take an active role in the healthcare ecosystem. Many mHealth applications have this as their objective, but raise significant challenges in assessing their impact. This is also

congruent with the challenge of recognising that healthcare resources are limited both in terms of healthcare professionals and budget, and it is important that new models are brought into the system to face this new era we are entering into.

If we take a look at the benchmarking analysis carried out by Research2guidance on the mHealth App Market Ranking [17], “*Denmark, Finland, The Netherlands, Sweden and the UK offer the best market conditions for mHealth companies in the EU*”. The main indicators that show a higher potential for mHealth adoption are:

- Regulatory frameworks for mHealth are in place, and guidelines on standards and interoperability following the European paths are adopted by trusted governmental bodies.
- High adoption of mHealth by healthcare professionals and patients/citizens, which lead to new service delivery processes and new communication channels to allow a much more active role of the patient.
- High level of digitalisation, integration and sharing of healthcare information by tools such as the electronic healthcare record, personal health record, ePrescription and many others.
- Strategic roadmap on mHealth is supported by the national government and policy makers that will facilitate the integration within the system at large scale.

In order to start with the deployment of mHealth within the EU, the industry should carefully choose those countries in which they will start the integration and penetration of this technology. As shown in the Figure 2 below, five countries offer the best market conditions for mHealth industry to establish new businesses [17].

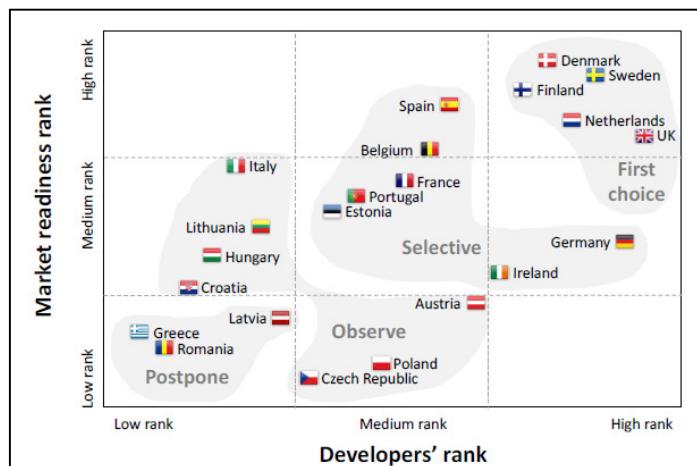


Figure 2. Country rank in market readiness for mHealth [17].

As indicated in the report on benchmarking analysis [17], “50% of the mHealth practitioners say that a good country ranking depends on how open doctors are for applying and integrating mHealth solutions into their patient treatments and communication. As there is no general reimbursement of mHealth services in all EU countries, this high rating of the doctor channel in the top country builds more on their

general openness to use new technologies rather than the existence of business models for doctors and mHealth services that work already today”.

One important set of players for acceptance and deployment of mHealth services is the governments and policy makers. Once they become aware of the benefits and potential of mHealth, they have to play also a very active and important role in spreading those benefits among the citizens and making possible the integration within the healthcare system. The establishment of concrete roadmaps and action plans aligned with the political agendas will help to position the country as a leading player for mHealth deployment in the EU scene.

3. Classification of mobile ‘apps’ and related solutions in healthcare

Although the number of mobile health apps is large and growing, most have only simple functionalities built into them. An analysis [18] of the apps available to consumers through the iTunes app store resulted in categorization of apps based on whether they could:

- **Inform:** Provide information in a variety of formats (text, photo, video)
- **Instruct:** Provide instructions to the user
- **Record:** Capture user entered data
- **Display:** Graphically display user entered data/output user entered data
- **Guide:** Provide guidance based on user entered information, and may further offer a diagnosis, or recommend a consultation with a physician/a course of treatment
- **Remind/Alert:** Provide reminders to the user
- **Communicate:** Provide communication with HCP/patients and/or provide links to social networks

There is a small subset of apps with complex functionality (e.g. electrocardiogram (ECG) readers, blood pressure monitors, blood glucose monitors), however it is recognized that most of the mHealth apps available today are only simple in design and do little more than provide information.

An alternative approach to classify mobile applications is to place them according to their use as part of the care continuum, sometimes called the “patient journey”: overall wellness and healthy living, diagnosis/self-diagnosis, healthcare professional visit, follow up and further information, prescription filling and medication compliance [18].

The working group on mHealth assessment of the Agency for Health care Quality and Assessment of Catalonia (AQuAS) and the mHealth Competence Centre of Mobile World Capital has developed and suggests a new taxonomy for classification of mHealth applications and services which combines the three aspects: 1) functionality and intended use, 2) type of clinical condition, and 3) potential risk (Figure 3).

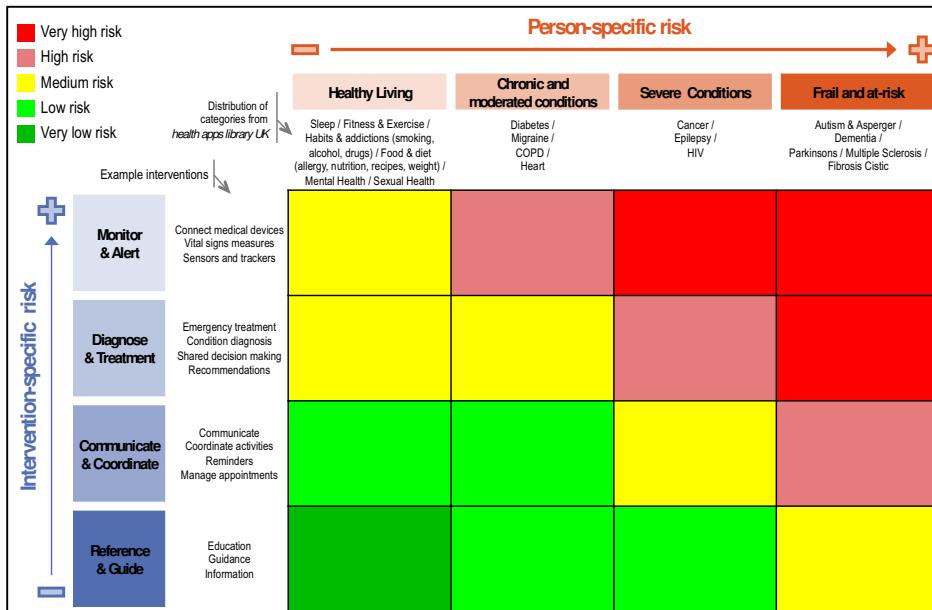


Figure 3. Proposed taxonomy of mHealth and related risks on the use of the solution for patient safety by AQuAS and MWC[®].

4. Frameworks for mHealth evaluation

In order to inform policy and user decisions, generation of reliable scientific evidence of mHealth benefits through systematic evaluation is crucial. There is need to assess the impact and empirically demonstrate benefit and best use of mHealth solutions as part of care delivery and health/disease management. However, health care innovations based on mHealth solutions have several features that make scientific evaluation challenging, such as fast technologic turnover and strong influence of design and organizational context. To answer the particularities of mHealth applications and services, the evaluation approach should be iterative and involve the views of all relevant stakeholders early in the process. Some initiatives exist for rating, validation and certification of marketed mobile apps, but these offer partial assessment generally focused on usability and data privacy and protection (such as myhealthapps.net, AppSaludable, HealthApp Library NHS, iMedical apps) [30].

On the other hand, there are numerous frameworks and models developed for evaluation of eHealth services (including health IT and telemedicine) and these are extensively based on commonly accepted evaluation methodologies, such as Health Technology Assessment. After a review and analysis of these, they could be easily adapted to meet mHealth evaluation needs through an overarching assessment approach (Table 1).

Table 1. Key evaluation frameworks for telemedicine and health IT that could be adapted to mHealth evaluation.

Reference	Telemedicine	Health IT systems	
	Kidholm K et al. 2012 [27] A model for assessment of telemedicine applications: MAST.	Catwell & Sheikh, 2009 [28] Evaluating eHealth interventions: the need for continuous systemic evaluation	Yusof et al. 2008 [29] An evaluation framework for health information systems: human, organization and technology-fit factors (HOT-fit)
Target group	Health professionals, patients, managers, policy makers.	Developers/design teams	Researchers, practitioners s (clinicians/GPs)
Goal/ Approach	Its aim is to inform the formulation of safe, effective, health policies that are patient focused and seek to achieve best value. It is a multidisciplinary process that summarizes information about the medical, social, economic and ethical issues related	Comprehensive overall evaluation approach, multifaceted, multidisciplinary approach and facilitates continuous systematic evaluations throughout the lifecycle of an eHealth intervention.	Provides evaluation dimensions for addressing the fit between human, organization, and technology factors. HOT-fit should be applied in a flexible way, taking into account different contexts and visions, stakeholders' point of views, phases in the system development life cycle, and evaluation methods.
Foundation	Literature review, EUNetHTA, IOM	Literature: 1) cognitive and usability engineering methods for the evaluation of clinical information systems, 2) socio-technical and contextual considerations	1) The IS Success Model of DeLone & McLean, 2) The IT-Organization Fit Model, 3) Literature review; critical appraisal of health information systems studies; 4) Pilot testing developed framework (case study clinical setting)
Dimensions	<p>Preceding considerations Purpose of the telemedicine application? Relevant alternatives? International, national, regional or local level of assessment? Maturity of the application?</p> <p>Multidisciplinary assessment 1) Health problem and characteristics of the application 2) Safety 3) Clinical effectiveness 4) Patient perspectives 5) Economic aspects 6) Organizational aspects 7) Socio-cultural, ethical and legal aspects Transferability assessment: 1) Cross-border, 2) Scalability, 3) Generalizability</p>	Documenting the complex relationships between: (1) political, (2) social, (3) organizational, and (4) technical worlds. Continuous systematic evaluations (eHealth intervention lifecycle: (1) inception (e.g. vision, goals & needs) (2) requirements & analyses (3) design, develop & test (4) implement & deploy	1) Human factors: system use, user satisfaction 2) Technology factors: system, information, and service quality 3) Organizational factors: structure, environment, communication, 4) Net benefits: impact on users, performance; efficiency, effectiveness, etc.; organizational impact (e.g. costs); clinical impact (quality of life, care, communication/information access).
Evaluation methods	Measure efficacy, effectiveness, safety, usual methods are: RCT, cohort studies, quasi-experimental design Economic evaluation; cost-	1) Formative iterative evaluations using simple prototypes of the eHealth intervention may be used for requirements	Qualitative, quantitative or a combination of both approaches: 1) Formative evaluation to identify system problems

	<p>effectiveness analysis Qualitative approach: interview, focus group, questionnaires, etc.</p>	<p>elicitation and analyses 2) Once a working model of the system is available, empirical evaluations can be completed, which could include the collection of quantitative and/or qualitative data, depending on the goals and scope of the study and the stage of development</p>	<p>as they emerged and to improve the system as it was developed. 2) Qualitative methods to generate a fuller description of the healthcare setting and its cultural issues and to understand why the system functioned well or poorly in a particular setting. Face-to-face interviews (including users, clinicians and IT staff) about their system us</p>
--	--	--	--

5. Regulation of mHealth as part of the evaluation process

Despite the fact that mHealth applications are numerous and getting more popular due to all the potential benefits described above, they are still under regulated and may pose risks to the health and safety of consumers, as well as to the privacy and security of consumer health information.

Health and safety refers to the physical health and wellbeing of a user of the application. Characteristics of the application that have influence on patient safety are related to its functionalities - in particular to the appropriateness, accuracy and reliability of used information. For example, a mHealth application may provide inaccurate information or recommendations on how to treat a condition causing negative impact on a patient's overall health.

On the other hand, patient privacy and security refers to safeguarding protected health information (PHI) [18]. Privacy is an individual's right to control access to his/her PHI. Security is the device's or user's ability to protect PHI from unauthorized disclosure either when stored on the device or transmitted to another device. Security requires technical safeguards, such as encryption, workstation security, and access controls, while privacy focuses more on an organization's policy and procedure for protecting PHI [20].

Safety and transparency of information were identified as one of the main issues for mHealth uptake in the public consultation on the Green Paper on mobile health of the European Commission [21]. In the public consultation, a majority of respondents thought that safety and performance requirements of lifestyle and wellbeing apps are not adequately covered by the current EU legal framework while calling for a strengthened enforcement of data protection and medical devices rules.

For conventional medical devices, all these issues are addressed by the manufacturer who must fulfil the requirements establish by the relevant regulatory authorities. Patient safety is thus guaranteed by controlling that only safe and effective devices reach the market.

Currently, the debate around the regulation of mHealth applications and services is getting momentum and classification algorithms are proposed by FDA and the European institutions to support decisions as to whether a certain mHealth application is a medical device or not. A governing principle in both regulations is the concept of "intended use" of the application and this determines the applicability of medical device regulation. Updates and amendments of existing regulation are currently

underway in both the US and the EU in order to better respond the needs for evaluation and marketing of mobile applications for health use [22].

In February 2015 the US Department of Health and Human services at the FDA issued Guidance for Industry and FDA staff on Mobile Medical Applications to explain the position of FDA on this topic [22]. In order not to stifle innovation and waste resources, FDA decided to limit its regulatory reach by identifying clearly the specific group of mobile applications which are subject of regulation. Thus, three categories of mobile apps are defined:

- a) Regulated mobile medical apps (those complying with the definition of medical device)
- b) Mobile apps subject to enforcement discretion (may meet the definition of medical device, but pose a lower risk to the public)
- c) Unregulated mobile apps (do not meet the definition of medical device)

The majority of available mobile apps on the market currently fall in the unregulated groups b) and c). There are six subcategories within the category b) enforcement discretion and each of them has a policy basis for existing [23]:

1. Patient self-management
2. Patient trackers
3. Access to contextually relevant information
4. Patient communication and telemedicine
5. Simple, professional calculators
6. Connectors to Electronic Health Records

In this guidance, the FDA lists a number of examples of mobile apps, to assist manufacturers in determining if a product is a mobile medical app and to follow the associated controls established by the regulation.

In the EU there is no integrated health regulation framework with a single regulatory body, such as the FDA in the US. The EU regulates mHealth in a number of ways: by means of medical devices regulation, regulation of personal health data, reimbursement of healthcare rules, and product liability. To be legally introduced in the EU market, a medical device should bear the Conformité Européenne (CE) mark. The CE mark states that the device has been assessed before being placed on the market and meets EU safety, health and environmental protection requirements. mHealth as a service is not regulated, but the software to provide the service is under the e-Commerce directive (SaaS: Software as a Service) [34].

The European medical device directive (MEDDEV2007/47) contemplates the software (stand-alone) in the definition of the medical device. Particularly, according to the EU directive, a medical device is defined as: "Any instrument, apparatus, appliance, software, material or other article, whether used alone or in combination, including the software intended by its manufacturer to be used specifically for diagnostic and/or therapeutic purposes and necessary for its proper application, intended by the manufacturer to be used for human beings for the purpose of: diagnosis, prevention, monitoring, treatment or alleviation of disease, diagnosis, monitoring, treatment, alleviation of or compensation for an injury or handicap, investigation, replacement or modification of the anatomy or of a physiological process, control of conception, and which does not achieve its principal intended action in or on the human body by pharmacological, immunological or metabolic means, but which may be assisted by

such means". Thus, since a mobile app is a software, it is a medical device if its intended use falls within the above definition.

An effort is needed to implement the medical device regulations to the mobile app market, since the market of the medical devices, with which the competent authorities for medical devices of the member states are familiar, is far different from that of the mobile apps [24].

We can conclude that existing regulations (FDA, CE mark) are relevant to address certain risks, but cover only mHealth technologies classified as medical devices. A broader evaluation of the impact of mHealth services should be coherent with existing regulation for medical devices but goes beyond it. In this sense, regulatory requirements can be considered as part of the initial stages of an evaluation approach. In order to inform policy and practice decisions, further generation of reliable scientific evidence through systematic evaluation is crucial to assess the impact and empirically demonstrate benefit and best use of mHealth solutions as part of care delivery and health and disease management.

6. Challenges for implementation and adoption of mHealth solutions

Multiple barriers, such as regulatory, economic, structural and technological, are limiting the adoption of mHealth. Also, the non-existence of clear business and exploitation models behind the implementation of mHealth services makes it difficult to expand and deploy these new technologies for the benefit of patients and professionals. The industry still is a little reluctant to invest efforts and budget in certain pilots and initiatives that seem to be far from the market.

The main two changes foreseen for 2020 in the field of mHealth are around *data integration* and *interoperability* of services and platforms [25]. Both are necessary to support sharing of information between patients and professionals, healthcare centres and easy implementation of new solutions. Both will help avoiding isolated silos that decentralise information and make difficult the taking of decisions based on an aggregated pull of data available. Also, in the next coming years, it is foreseen that more and more medical apps will be developed and introduced in the market as a regulatory and legal framework is being agreed and adopted in the EU.

When looking into the biggest barriers for deployment, privacy issues and clear regulation frameworks are the most relevant issues that have been identified. This is not a surprise, since data management and sharing in the field of healthcare is one of the most important topics and goes directly linked to the use of technologies.

In order to encourage the adoption of these new solutions, it is important to take into account some actions that can facilitate the process, as for example the approval of an mHealth strategic plan within the regional or national strategies of the healthcare and social government departments. Also, adding up to the regulatory framework, the creation of innovative business models to provide sustainability of both health services and information/IT services which at the same time fulfill the objectives of all stakeholders implied in the process is important. And last, it is important to raise awareness among citizens, patients and professionals, through training programs and communication campaigns that show the benefits and added value coming from mHealth.

Another concern raised by professionals in the field is the lack of clinical evidence linked to the impact of mHealth solutions. As seen in Figure 4, very little scientific

evidence yet exists. This, together with the weak regulatory situation makes it difficult for professionals to act as “prescribers of apps” for the patients.

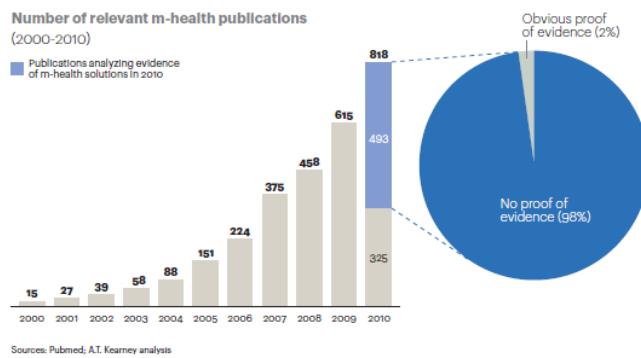


Figure 4. Number and level of evidence of mHealth publications 2000 - 2010 [26].

To overcome the barriers mentioned above (regulatory, economic, structural and technological), strong support from policy makers is needed which will enable and accelerate adoption by the system and final users. They need to “formulate policies that can drive adoption of mHealth solutions. The national and regional payers need to support these policies by creating facilitative reimbursement mechanisms that ease the adoption of mHealth solutions across patients and healthcare providers”.

As a summary, for each barrier identified that prevents from deploying mHealth, some actions to be executed are proposed below:

Regulatory framework

Actions needed:

- Regulations should effectively address issues as certification, standardization and interoperability to help increase the confidence and trust of both healthcare professionals and patients.
- Concrete roadmap and timeline on when relevant policies and regulations may be introduced and what might be addressed by such measures should be developed.
- Regulations should be pro-innovation and aimed at introducing measures that enable affordable and ubiquitous healthcare.

Standardization and Interoperability

Actions needed:

- Regulators have to work together to ensure interoperability and standardization guidelines for various mobile health ecosystem participants (device vendors, content creators and healthcare providers).
- Ensuring standardization and interoperability among solutions will help
 - plug-and-play solutions development
 - easy adoption for end-users
 - facilitate scaling

Certification of Applications

Actions needed:

- Regulators should facilitate speedy approvals for vendors and software developers.
- The intent of governments and regulators should be to enable the rapid creation of a healthy mobile health ecosystem that benefits both patients and market players.
- It is important for regulators to follow a harmonized approach to ensure greater applicability of certified devices and applications across regions to encourage greater participation of device vendors and solution developers.

Recommended further readings

1. GSMA and PA Consulting, *Policy and Regulation for Innovation in Mobile Health*. Report 2011. <http://www.gsma.com/mobilefordevelopment/wp-content/uploads/2012/04/policyandregulationforinnovationinmobilehealth.pdf>, last access 11 February 2016.
2. Cusack CM et al., *Health information technology evaluation toolkit: 2009 Update*, AHRQ publication No09-0083-EF, Agency for Healthcare Research and Quality, Rockville, MD.
3. Nykänen P et al., Guideline for good evaluation practice in health informatics (GEP-HI), *Int J Med Inform* **80** (2011), 815–27.
4. Kumar S et al., Mobile health technology evaluation: the mHealth evidence workshop, *Am J Prev Med* **45**(2) (2013), 228-36.
5. Rigby M et al., Evidence Based Health Informatics: 10 Years of Efforts to Promote the Principle, in: Sérousi B, Jaulent M-C, Lehmann CU (eds.), *Evidence-based Health Informatics – IMIA Yearbook of Medical Informatics*, Schattauer, Stuttgart, 2013. pp 34-46.

Food for thought

1. Which clinical areas can benefit the most from mHealth solutions?
2. Which kind of role can have patients/citizens/consumers on the regulatory and legal process, and is this appropriate?
3. Why is the US mHealth market growing faster than the EU market?
4. What do you think about the taxonomy proposed by AQuAS and MWC (compare Figure 3)?

References

- [1] M. Castells, *The Rise of the Network Society*, in: The Information Age: Economy, Society and Culture. 2nd ed., Wiley Blackwell, Oxford, 2010.
- [2] GSMA intelligence, <https://gsmaintelligence.com>, last access 11 February 2016.
- [3] Ericsson Mobility report, <http://www.ericsson.com/ericsson-mobility-report>, last access 11 February 2016.
- [4] S. Kumar, W. Nilsen, M. Pavel, M. Srivastava, Mobile health: Revolutionizing healthcare through transdisciplinary research, *Computer* **46**(2013), 28-35.

- [5] WHO, *mHealth – New horizons for health through mobile technologies*, Global Observatory for eHealth, vol 3, http://www.who.int/goe/publications/goe_mhealth_web.pdf, last access 11 February 2016.
- [6] R.S. Valdez, R.J. Holden, L.L. Novak, T.C. Veinot, Transforming consumer health informatics through a patient work framework: connecting patients to context, *Journal of the American Medical Informatics Association* **22**(2015), 2-10.
- [7] M. Fiordelli, N. Diviani, P.J. Schulz, Mapping mHealth research: A decade of evolution, *Journal of Medical Internet Research* **15**(2013): e95.
- [8] C. Free, G. Phillips, L. Galli, L. Watson, L. Felix, P. Edwards, V. Patel et al, The Effectiveness of Mobile-Health Technology-Based Health Behaviour Change or Disease Management Interventions for Health Care Consumers: A Systematic Review, *PLoS Medicine* **10**(2013).
- [9] N.G. Cortez, I.G. Cohen, A.S. Kesselheim, FDA Regulation of Mobile Health Technologies, *New English Journal of Medicine* **371** (2014), 372-379.
- [10] S.R. Steinhubl, E.D. Muse, E.J. Topol, Can mobile health technologies transform health care? *JAMA* **310**(2013), 2395-6.
- [11] P. Klasnja, W. Pratt, Managing health with mobile technology, *Interactions* **21** (2014), 66-69.
- [12] P.R. Sama, Z.J. Eapen, K.P. Weinfurt, B.R. Shah, K. Schulman, An Evaluation of Mobile Health Application Tools, *JMIR mHealth and uHealth* **2**(2014), e19.
- [13] R.S. Weinstein, A.M. Lopez, B.A. Joseph, K.A. Erps, M. Holcomb, G.P. Barker, E.A. Krupinski, Telemedicine, telehealth, and mobile health applications that work: opportunities and barriers, *American Journal of Medicine* **127**(2014):183-7.
- [14] US Food and Drug Administration, *Guidance for Industry and Food and Drug Administration Staff: Mobile Medical Applications*, Document issued on 9 of February 2015, <http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM263366.pdf>, last access 11 February 2016.
- [15] PricewaterhouseCoopers, *Socio-economic impact of mHealth: An assessment report for the European Union*, June 2013, http://www.gsma.com/connectedliving/wp-content/uploads/2013/06/Socio-economic_impact-of-mHealth_EU_14062013V2.pdf, last access 11 February 2016.
- [16] OECD, *Health at a glance: Europe 2012*, 2012, http://ec.europa.eu/health/reports/european/health_glance_2012_en.htm, last access 11 February 2016.
- [17] Research2Guidance, *EU Countries' mHealth App Market Ranking*, 11th may 2015.
- [18] IMS Institute, *Patient Apps for Improved Healthcare. From novelty to mainstream*, October 2013
- [19] HIPAA, 'Protected Health Information': What Does PHI Include? <http://www.hipaact.org/protected-health-information-what-does-phi-include>, last access 11 February 2016.
- [20] D.F. Schulke, The regulatory arms race: mobile-health applications and agency posturing, *Boston University Law Review* **93** (2014), 1699-1752.
- [21] European Commission, *Green paper on mobile Health (mHealth)*, Brussels, April 10, 2014, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=5147, last access 11 February 2016.
- [22] FDA, *Mobile medical applications: Guidance for Industry and food and drug Administration Staff*. February 9, 2015, <http://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf>, last access 11 February 2016.
- [23] B.M. Thompson. *FDA Regulation of Mobile Health* (2nd ed). Chester Street Publishing, USA, 2013.
- [24] F. Censi, E. Mattei, M. Triventi, G. Calcagnini, Regulatory frameworks for mobile medical applications, *Expert Review Medical Devices* **12**(3) (2015) 1-6.
- [25] Research2Guidance, *mHealth App Developer Economics report*, May 2014.
- [26] ATKearney, *Improving the evidence for mobile health*, 2011.
- [27] K. Kidholm, A.G. Ekeland, L.K. Jensen, J. Rasmussen J, C.D. Pedersen, A. Bowes, S.A. Flottorp, M. Bech, A model for assessment of telemedicine applications: MAST, *International Journal of Technology Assessment in Health Care* **28**(2012), 44-51.
- [28] L. Catwell, A. Sheikh, Evaluating eHealth interventions: the need for continuous systemic evaluation, *PLoS Medecine* **6**(2009), e1000126.
- [29] M.M Yusof, J. Kujlis, A. Papazafeiropoulou, L.K. Stergioulas, An evaluation framework for Health Information Systems: human, organization and technology-fit factors (HOT-fit), *International Journal of Medical Informatics* **77**(2008), 386-98.
- [30] *My health apps*, <http://myhealthapps.net>, last access 11 February 2016.
- [31] *Estrategia de calidad y seguridad en aplicaciones móviles de salud*, <http://www.calidadappsalud.com>, last access 11 February 2016.
- [32] *Health apps library*, <http://apps.nhs.uk>, last access 11 February 2016.
- [33] *iMedicalApps*, <http://www.imedicalapps.com>, last access 11 February 2016.

- [34] European Commission, *Guidelines on the qualification and classification of stand-alone software used in healthcare within the regulatory framework of medical devices*, http://ec.europa.eu/health/medical-devices/files/meddev/2_1_6_ol_en.pdf, last access 11 February 2016.

Systematic Reviews and Meta-Analysis of Health IT

Christine URQUHART^{a,1} and Rosemary CURRELL^b

^a*Aberystwyth University, UK*

^b*Bury St Edmunds, Suffolk, UK*

Abstract. This contribution examines how systematic reviews contribute to the evaluation of health IT planning and implementation. It defines and explains the systematic review process and how higher level overviews of health IT can be conducted. A reprise of some of the Cochrane reviews relating to health IT, particularly those conducted for the Cochrane Effective Practice and Organization of Care Group (EPOC), provides examples of the type of question that can be answered (at least in part) by a Cochrane-type systematic review. The contribution also discusses the benefits and limitations of the systematic review process using examples of reviews on telemedicine, nursing records, and home uterine monitoring in pregnancy.

Keywords. Evidence-based medicine, review, meta-analysis, telemedicine, health information systems, nursing records, evidence-based practice.

1. Introduction

Decisions about treatments by individual healthcare professionals, and decisions concerning health service delivery options and public health interventions should be made on the basis of the best evidence available at the time. Evidence-based medicine stresses the importance of using evidence from randomized controlled trials (RCT's) as these are likely to provide much more reliable information than other sources of evidence. Hence any summaries, reviews and critiques should give most weight to evidence derived from studies using the most rigorous research designs. From early beginnings in the 1990s, the growth of "systematic reviews" of the healthcare evidence has been dramatic, and now covers far more than treatments alone.

The first sections of the contribution describe the systematic review process, with emphasis on the procedures applied to reviews for the Cochrane Collaboration and the guidelines for the Effective Practice and Organization of Care (EPOC) Cochrane group. We describe the type of searching required for reviews that deal with health information technology, discuss the types of research design that qualify for inclusion, and consider the status of such reviews within EPOC. Policymakers may require overviews, meta-reviews of some aspect of e-health, and the contribution examines the theoretical approaches that may underpin such overviews.

The case study of the work on an EPOC Cochrane review of telemedicine then discusses how the original questions that concern the technology will change as the

¹ Corresponding author: Dr. Christina Urquhart, Senior Lecturer, Aberystwyth University, Department of Information Studies, UK, cju@aber.ac.uk.

technology is integrated into clinical practice. This may mean that the question of interest to the health informatics specialist – does telemedicine benefit patient care – is not any easier to answer. The contribution also reflects on the experience of updating a systematic review of the effect of nursing records on nursing practice. The conclusions set out some pointers for those doing systematic reviews of health information technology. We also provide guidance for those designing health IT evaluations, to ensure that the research designs are sufficiently rigorous for inclusion in Cochrane reviews, thus providing high quality evidence for policymakers, professionals and the public.

2. The systematic review process

Systematic reviews of the research literature are a rigorous approach to collecting and appraising the research evidence on a topic of significance to clinicians and policymakers. The main steps of any systematic review are:

- 1) define the questions and comparisons to be investigated;
- 2) retrieve research evidence (using appropriate databases and explicit, reproducible search strategies);
- 3) sift the retrieved items to find studies that are in scope (on the topic), and which meet the required quality criteria to be included;
- 4) assess the validity of the study findings, and assess the risk of bias;
- 5) prepare a synthesis of the studies included in the review.

The Cochrane Collaboration is the main international body that prepares, maintains and publishes systematic reviews that aim to help healthcare decision making by clinicians, managers and policymakers.

A systematic review must be conducted in the same disciplined way as any other research study, and must begin with a protocol that sets out the research question to be answered and the scope of the review. Systematic reviews carried out within the Cochrane collaboration are conducted under the guidance of the relevant subject area review group and begin with the scrutiny and approval of the review protocol. Eligibility criteria for the studies to be included in the review are specified in advance. The search strategies are intended to find as many relevant studies as possible, and the search strategies are included in the published review, to enable others to check how the studies were collected and collated. The processes for appraisal are also systematic, so that the process for inclusion and exclusion of studies is fair and explicitly justified in the published review. The transparency of the review process is very important, as individual clinicians and policymakers do not have time to check through the evidence for themselves. They need to be confident that a systematic review has minimized any bias, by collecting and appraising the evidence in a systematic way.

According to the Cochrane Handbook [1], “the key characteristics of a systematic review are:

- a clearly stated set of objectives with pre-defined eligibility criteria for studies;
- an explicit, reproducible methodology;
- a systematic search that attempts to identify all studies that would meet the eligibility criteria;

- an assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias; and
- a systematic presentation, and synthesis, of the characteristics and findings of the included studies.”

Many systematic reviews contain meta-analyses. Meta-analysis is the use of statistical methods to merge and summarize the results of independent studies [2]. Systematic reviews should be able to assess whether there is sufficient evidence to say whether an intervention (treatment, or mode of healthcare delivery) is effective or not. Meta-analysis, by combining data from more than one compatible and relevant studies, can provide more precise estimates of the effects of an intervention, usually depicted as a forest plot. This helps to clarify whether an intervention is truly effective, on the basis of all the evidence when some studies indicate there is an effect, and others do not. During the process of meta-analysis, it should be possible to explore the consistency of the evidence, and why some studies point to an effect in one direction, and others show a different type of effect. Meta-analysis is only possible where the studies being brought together are sufficiently similar in their conduct, inclusion and exclusion criteria and outcomes. This can be relatively straightforward in studies of different treatment interventions, but is much more difficult in studies of health care delivery methods and health informatics interventions.

Many health IT interventions fit within the scope of the Cochrane EPOC (Effective Practice and Organization of Care) group. For most of the other Cochrane groups, the studies included for the review will be limited to randomized control trials. In such trials, participants are assigned to the intervention or not, randomly (i.e., they have an equal probability of being assigned to any group). Procedures are put in place to ensure all participants in all study groups are treated the same for the condition of interest, except for the type of intervention they receive. There may be some other undiscovered variables that affect what happens, but random assignment (and a suitably large sample) should ensure that these are similarly distributed among the treatment and control groups. In other words, if the intervention makes a difference in outcomes between the treatment and control group, we can be reasonably secure in concluding that the intervention is responsible and that the difference is not due to some other variable.²

Randomized control trials are suited to evaluating whether drugs are effective, but for trials of interventions that involve modes of healthcare delivery, it may not be appropriate, or possible, to organize a randomized control trial. The Cochrane EPOC group allows for inclusion of controlled before and after research designs, in which two groups of participants are studied before an intervention, and after an intervention. Only one group receives the intervention, the other group acts as the control. Procedures should be in place to ensure that each group is as similar to each other as possible, to minimize the possibility of falsely ascribing an effect to the intervention when something else made the difference. It is important that both groups are studied after the intervention, as well as before, and that as far as possible this is contemporaneous.

² See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

Another type of research design that can be included in a Cochrane EPOC review is the interrupted time series design, in which a series of measurements are made before and after the intervention.³ For implementation of an IT system, the interrupted time series design is feasible, and sensible in many circumstances, as it is often not possible to have a control group, but one difficulty of including studies for a Cochrane review is the need for at least three time point measurements prior to the intervention (installation of the system). Researchers are much more likely to provide data for many time points after implementation, but often there is only one time point measurement prior to the implementation.

Searching for studies for a health IT review is sometimes more complicated than searching the literature for a clinical topic. It is important to remember who might be doing the independent research or evaluation for an IT-related topic. Agencies that deal with research on health services delivery may fund evaluations of IT systems, or health IT projects, but evaluations are often small scale and internal (unlike drug trials externally funded by pharmaceutical companies). Evaluations may be conducted by postgraduate students, who may be working as members of staff in the organization under evaluation or in association with a staff team. Publications on health services research may be found in a variety of databases, beyond the purely clinical databases. It is also important to remember that research that is more to do with pure computer science is quite likely to appear in conference proceedings, rather than a journal, as this is quite normal for computer science research. Therefore it may be necessary to look at publications from ACM (Association for Computing Machinery), Dissertation Abstracts (and university repository collections such as e-theses online – EthoS), HMIC (Health Management Information Consortium), as well as general databases such as SCOPUS, Web of Science, and the clinical databases such as MEDLINE. Google Scholar is increasingly useful in locating reports, and non-journal literature for the social sciences.⁴

3. Cochrane Collaboration and health IT reviews

Given the money spent on health IT, it is surprising that there are not more current reviews of health IT in the Cochrane Library for the EPOC group. A check in July 2015 of the EPOC website showed that, under the review category of “delivery of health care” there were seven items in the information and communication technology category. Of these, two were at the early stage, with published protocol but no published review: delivery arrangements for health systems in low-income countries [3]; and patients’ record systems in dental practice [4]. Two reviews (telemedicine versus face to face patient care [5]; nursing record systems [6]) are discussed in more detail in this contribution. Of the remaining three published reviews, one review on computerized advice on drug dosage to improve prescribing practice [7]) concludes that the quality of the studies was low so these results must be interpreted with caution. Another review on a topical subject (smart home technologies for health and social care

³ See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁴ See also: A. Georgiou, Finding, appraising and interpreting the evidence, in: *ibid.*

support [8]) found that no studies met the inclusion criteria. This is not a particularly useful conclusion for policymakers.

Perhaps the first stage of systematic reviewing on topical IT subjects is an overview of the various applications. A systematic review of healthcare applications for smartphones [9] mainly categorizes the applications and outlines how they work. There is less emphasis on the quality of the research design, but this type of systematic review is a useful “state of the art” review for policymakers. The final review of the Cochrane group (patient reminder and recall systems to improve immunization rates [10]) found that telephone reminders worked, but the evidence was limited to studies from developed countries. That is a pity, as there are opportunities for developing countries to implement neat, cheap solutions to healthcare problems using technology, just as mobile phone based money transfer services have revolutionized microfinance in parts of Africa.

There are, of course, other Cochrane Groups that deal with health IT topics. The Cochrane Consumers and Communication group has produced several reviews that deal with email communication (between professionals, and between professionals and patients, for co-ordination of healthcare appointments and reminders) and mobile phone messaging. This group has also published a review on interactive health communication applications for patients with a chronic disease (usually web-based information package with at least one other support mechanism (social support, behaviour change or decision support) [11]. This review concluded that there was evidence to support the use of such interventions, but noted the need for larger studies and larger sample sizes.

Other Cochrane clinical groups have examined the use of information and communication technology, as the case study example presented below discusses for telemedicine within clinical practice. Researchers in the social construction of technology (see, for example [12]) would expect to see a shift from the perspective of the technology as an innovation (and its acceptability to practice) to a later emphasis that has accepted the technology as part of practice and aims to find out exactly how the technology can benefit the delivery of care.

Technology implementations take time – training, awareness sessions, technical trials, adjustments, teething problems, changes to practice, growing acceptance (hopefully) and further tweaking of the system to make it work better for a clinical team. The timescale required from initial planning, through design, implementation and realization of benefits (or withdrawal of the system) means that an interrupted time series research design may be a more feasible, and more valuable way of assessing the effectiveness of an e-health implementation such as a records system rather than a randomized control trial, or controlled before and after design that did not cover an appropriate time period before and after implementation.

For those funding or conducting a systematic review one question is the best time to do a review. One of the possible problems with the smart home review [8] was the timing – was this done too early, when many reports came from early pilot uncontrolled trials done by enthusiastic IT champions, rather than larger, properly controlled trials? On the other hand, from the national policymaker perspective, some good quality, indicative evidence might be useful to decide whether to fund larger trials. There has been some growth what are generally termed “rapid reviews” to provide some quick answers to urgent questions for national and regional policymakers and managers, as traditional Cochrane reviews can, and usually do, take years to complete. One way of speeding up the review process is a review of reviews, essentially an

overview (and critique) of existing systematic reviews on a topic area. This is the approach taken for the overview of delivery systems for health care in low income countries [3]. An overview can be a useful way of bringing together all the so-called “systematic” reviews on a topic, although Cochrane overviews may only include existing Cochrane reviews.

Cochrane reviews can answer some questions about the contribution of information and communication technology to clinical care, but the evidence is, as noted, scattered and the emphasis may vary according to whichever Cochrane group is dealing with the topic. Some of the questions need answers supplied through qualitative evidence, as we will later discuss. The next section considers some theoretical approaches to modelling the acceptance (and use) of technology, and the normalization process theory that can be used to examine acceptance and adoption of technology within a complex system. The notion that implementation can be viewed as a neat, linear, and rational process has long been challenged [13]. We may be able to identify, in hindsight, what some of the barriers to acceptance were, but removing the barriers is no guarantee of success. Prediction is difficult, particularly when changes to behavior and practice are concerned.

4. Theoretical approaches to modelling IT implementation

This section considers some theoretical models on IT acceptance and use, and how such models illuminate the process of IT evaluation, and help to add value to the conventional Cochrane systematic review.

The Technology Acceptance Model (TAM) is a well-known model that attempts to explain IT acceptance and use. A review [14] of the application of TAM to health care lists and defines the variables that have been examined in TAM and related studies that may use TAM2, Theory of Planned Behavior, Theory of Reasoned Action or the Unified Theory of Acceptance and Use of Technology. The review specifically included studies that examined the quantitative relationships between the variables, and found that the health care set often started with the TAM conceptual framework but may have added or removed variables from the TAM set. It was sometimes unclear what definitions the study authors had used, and the operationalization of the constructs varied among studies. However, the review concluded that there were generally significant relationships between perceived usefulness of a system and clinicians’ intentions to accept and use an IT system. Perceived ease of use seems to correlate with perceived usefulness.

The lessons for Cochrane reviews dealing with health IT implementations are that some TAM variables are worth measuring, but care is needed about definitions and measurement of the variables, particularly for the added variables, that often try to assess the fit between the technology and workplace practice. For health care settings, the review concludes that more studies that elicit beliefs held by clinicians about health IT use will help to refine theories about acceptance and use of IT by clinicians.

An overview, a meta-review, of systematic reviews and meta-syntheses concerned with e-health implementation [15] used normalization process theory as a conceptual framework for the overview. This theory considers: 1) coherence (how people make sense of the e-health implementation, how they can ascribe a value to it, and appreciate their role); 2) cognitive participation (involvement and initiation); 3) collective action

(workability, and organizational support); 4) reflexive monitoring (how people appraise and perhaps reconfigure the system).

This type of meta-review helps to understand “why” and “how” questions, rather than the “does it work or not” question that is associated with a Cochrane intervention review. The type of evidence that is required, and which may be very dependent on the situation of e-health implementation, is the effect on roles and responsibilities, risk management, effective ways to engage with professionals, and transparent “working through” – ensuring that the improvements to obtain clinical benefits are clear to everyone. Normalization process theory could be argued to take a systematic approach to the bricolage and improvisations described as a normal part of information systems development [13].

Earlier work we undertook on the Cochrane review of telemedicine had identified a discrete group of studies, mostly conducted in the late 1980’s and early 1990’s on the use of telemedicine for home uterine monitoring in pregnancy. There had been considerable debate about the effectiveness of such home uterine monitoring for preventing preterm birth. When we undertook the review and meta-analysis of these studies for the Cochrane Pregnancy and Childbirth Group, we also wanted to explore why the researchers took the positions they did, and what some of the assumptions were. We required a theoretical framework that went beyond the acceptance and use of the technology itself, a framework that allowed for the differences in clinical perspective so evident in the letters to the editors of many of the journals publishing home uterine monitoring studies. We used a socio-technical interaction network (STIN) approach – STIN uses some ideas from the social construction of technology and actor-network theory but develops these into a usable framework for a method of enquiry.

This framework [16] has eight stages: 1) Identify a relevant population of system interactors; 2) Identify core interactor groups; 3) Identify incentives; 4) Identify excluded actors and undesired interactions; 5) Identify existing communication forums; 6) Identify resource flows; 7) Identify system architectural choice points; 8) Map architectural choice points to socio-technical characteristics. Using this framework [17], it was obvious that patient educators were (mostly) excluded actors, the motivations and incentives in the trials varied considerably, and the type of desirable interaction between monitoring center staff (and/or midwives) was unclear. Studies varied in the type and frequency of contact established between the pregnant women and the clinical centers receiving the monitoring data. Monitoring might be intended as an additional check, as part of an intensive care scheme with regular contacts with healthcare staff, or it could simply complement an educational program with an emphasis on self-care, prompted by monitor use. The system architecture choice points appeared to be: education supplement (monitoring to supplement a woman’s own interpretation of self-palpation, providing objective data for interpretation by trained staff) or education empowerment (interpretation of the monitoring data discussed by staff and the pregnant woman together).

An alternative set of architecture choice points was: maintaining contacts (reinforcing a regimen of regular contacts) or providing additional data for face-to-face contacts with health professionals. There is some overlap, as assumptions are mixed up with values about clinical care, and a system that aims for empowerment may look very similar to a system set up to enhance the regular face-to-face contacts. The use of the framework also highlighted the lack of data on resource costings, and the lack of data on the acceptability of the monitoring to the women themselves.

Doing the STIN analysis helped when doing the Cochrane review itself [18] as it was easier to identify some of the hidden assumptions about the possible clinical benefits of the system. In fact, the conclusions of the review suggested that the uterine monitoring that was possible at the time most of the studies were conducted was not clinically beneficial. This did seem to be a technical solution trying to find a suitable clinical problem to solve, rather than the other way round.

5. Challenges in the art of systematic reviewing – the case of telemedicine

We conducted a review of telemedicine at the end the 1990s and although various telemedicine applications had been under trial around the world for 30 years or more, they had largely been small pilot studies, with few studies and publications meeting the standards required for a Cochrane review. At the time of publication in 2000, only seven trials met the inclusion criteria [5]. At the last attempt to update the review in 2009, the world of health informatics and telemedicine had changed dramatically.

Telemedicine is now a major sector of health informatics and its implementation is a concern for health policy makers and commissioners around the globe. However, there is no agreed definition of 'telemedicine' and thus there is a problem for any systematic review of the topic, from the very outset. The first principle of all systematic reviews is a precise definition of the intervention to be studied. Telemedicine has evolved over the last fifty years and technology and applications have changed considerably over that time; change in technology has been dramatic in the last ten years or so. Technology has been applied to a whole range of illnesses and health care conditions with varying degrees of success, and any one telemedicine application can now make use of a number of different technologies. For example, remote monitoring at home from a specialist cardiac unit can include the transmission of real time ECG signals and consultation using videoconferencing, while diabetes monitoring can be by electronic monitoring of blood glucose levels and text messaging responses [19-20], or any combination of such technologies.

Telemedicine may include telehealth care, telemonitoring, telecare and other terms, and indexing terms need careful scrutiny as the search strategy for a review in this area is developed. Technologies that might come under the umbrella of 'telemedicine' may now include forms of video-conferencing including Skype, use of telephone modems to relay results of clinical monitoring in real time or as 'store and forward' applications, use of internet delivered interventions for mental health or health promotion, web based bulletin boards, mobile phone text messaging etc, etc, and any of these in combination, e.g. [21-26].

The literature shows that both studies and reviews of telemedicine may be initiated by clinical specialists, with the expectation that a particular clinical need might be met using some form of telemedicine or telecare. Reviews of heart failure monitoring or diabetes monitoring are common examples. The definition of interventions to be compared may be very different in studies and reviews initiated this way compared with studies and reviews initiated by service commissioners who may have an interest in the usefulness of a particular technology application to meet a whole range of chronic conditions [27].

Not only will definitions of the applications be different, but the outcomes chosen for the studies and the reviews will also reflect the different interests of the authors. Clinically initiated reviews may have very precise patient outcomes such as

biochemical measures, whereas organizationally initiated studies and reviews of care given to patients with similar conditions may focus on more process focused measures such as hospital bed days and readmission rates [24].

Systematic reviews are usually expected to consider all published studies meeting their criteria, and reviews of health informatics will always have to take account of the health systems in which the interventions are embedded. Interventions in private, public, not for profit, primary, secondary, tertiary, university care settings will all come together in a single review. The care providers in telemedicine studies vary, even within apparently equivalent applications. The health care professional receiving information or giving feedback to the patient may be a nurse, a doctor, or an allied health professional, and may be a generalist or a specialist, different studies of the same topic making use of health professionals in quite different ways [23].

Telemedicine may be initiated as a possible solution to social as well as clinical problems, and a number of telemedicine projects have been designed to reach populations who would otherwise be underserved by health services, because of their remoteness from a major centre, or because they are underprivileged economically; for example community hypertension surveillance in some areas of the USA [28] and video conferenced consultations for prisoners [29].

Without doubt, use of telemedicine is always going to be a complex intervention and this makes the task of producing rigorous and credible systematic reviews of the topic difficult. Not only must the reviewers match each study to the inclusion and exclusion criteria agreed in the review protocol and record the quality of the research design and conduct of each study to be included, they must also exercise common sense about which studies can sensibly be considered together. This is not just a matter of whether study design, subjects and outcomes are sufficiently similar to allow meta-analysis, but whether the aspects of the applications described above are sufficiently similar in practical terms for the clinicians, commissioners or policy makers with an interest in the topic to perceive the review as credible and useful for their purposes. This has proved problematic in updating an old review of telemedicine, where the technology, its availability and applications have outstripped the original criteria for review, and these issues are common to other fields of health informatics. Questions of 'What to review?', 'When to review?' and 'Who should review?' need to be asked at the very beginning, together with that most important question 'Why review?' and it is always possible that there is not a good enough answer to that fundamental question.

Telemedicine therefore provides a good illustration of the problems associated with systematic reviews of complex interventions and of health informatics reviews in particular. What the telemedicine review has also shown is that it is possible to design robust research studies of health informatics applications if sufficient thought is given at the outset of the implementation. Sadly many studies were identified in the course of the review, that appeared to be 'after thoughts', or had not had good research expertise input, and were not able therefore, to make the useful contribution to the health informatics literature that they might otherwise have done.

6. Reflections on updating a nursing records review

Updating the Cochrane Nursing Records review has also been problematical. Some of the issues have been similar to those encountered in the telemedicine review, but nursing records have also revealed other problems, which may be found in other areas

of health informatics research. The original review was published in 2003, and the most recent published update was in 2009 [6].

Some of the conclusions of the first systematic review in 2003 indicated that some important questions could only be answered through the use of qualitative evidence, and further work was done to synthesize the qualitative evidence [30]. Qualitative research on nursing records systems, together with other evidence from surveys was collated to answer questions on the format of the record (structured versus free text, for what type of practice), occasions when information exchange about nursing care may not and should not be recorded formally, and the effective organization of the nursing record. The synthesis of the qualitative research emphasized themes around compliance, confidentiality and recording to avoid future litigation problems. There is the formal record, but this may not truly reflect the quality of nursing care, particularly as nursing care is continuous in many situations. The purported time savings from using a computerized system rather than manual recording seemed elusive, and rarely reflect changes in quality of care.

The update to the 2009 Cochrane review is a larger task than the previous update, which is a good sign in many ways. There are more studies to be included. Although a large number of evaluations have to be listed in the excluded studies section of the review as they are uncontrolled before and after studies, there are some larger trials that are well-designed, including controlled before and after studies, and interrupted time series studies with sufficient time points to be included in the Cochrane review. It is likely that the analysis of the data across the studies will be limited to descriptive analysis as the exact definition of patient or practice outcomes depends on situation and even within one trial, across two main regions, there seem to be differences in practice on the frequency and format of documentation.

Although one possible benefit of computerized documentation might be more comprehensive documentation, done more efficiently, it is not possible to claim that time saved on documentation is necessarily allotted to direct patient care activities and therefore signals a better quality of care. The situation in which care is provided is likely to be as, if not more important. For example, the presence or absence of personal advice to nursing staff on how to improve the quality of care may be more influential than an information system that is designed to promote better quality documentation, and prompt care that accords to guidelines. From the patient perspective, it would be useful to be assured that the implementation of an electronic records system was associated with a reduced risk of falls, or reduced risk of pressure ulcers, but the risk depends on the situation of care.

Meta-analysis of some patient outcomes may be possible for the current update of the nursing records review (due to be published in early 2016), but it is likely to be limited. Even for the home uterine monitoring review that did have a strong clinical focus, the meta-analysis for that was complicated by the time lag between the time most of the studies were conducted and the time of the review. New clinical knowledge meant that many of the clinical outcomes that were the focus of the studies over ten years ago were not the patient outcomes of current clinical interest. The quality of the reporting in the studies and some problems with their research designs also meant that the quality of evidence was moderate [18], as the forest plots demonstrate. Clinical outcomes of interest change, and so do management and policy priorities.

What does not change is the need for high quality research design of trials, as advocated by Archie Cochrane [31], in his well-known reflections for improvement of health services. Cochrane was concerned about effective and efficient clinical care, and

wished to avoid unnecessary interventions that may do more harm than good to the patient. Perhaps IT enthusiasts need to remember that bit about unnecessary interventions and avoidance of harm. IT systems may allow us to do more, but is that actually helping clinical practice and patient care? The value of reviews such as the Cochrane nursing records review is that they do illuminate how that question may be answered.

7. Conclusion

We have discussed how systematic reviews of health information technology can be done, and what they can contribute to the evidence about the contribution of technology to improving the quality of patient care, and professional practice. Those doing or commissioning the review need to be clear about what to review, and when. An early review may simply scope the status of some possible applications, a later review may indicate where more work is required to realize benefits.

Cochrane reviews are largely concerned with quantitative evidence, but for many health information technology evaluations, qualitative evidence may be useful. There are theoretical frameworks that can help reviewers, as well as those designing the evaluation.

Cochrane reviews largely deal with randomized control designs and these are not necessarily ideal, or feasible, for those designing an evaluation of a health information technology implementation. Some Cochrane review groups such as EPOC allow for the inclusion of controlled before and after, and interrupted time series designs. These are often far more feasible and appropriate for health information technology evaluation, but there are no guidelines, for example, on the spacing of time points before and after an implementation. This type of guidance should come from health informatics experts.

It is disappointing that there is so little evidence on health information technology interventions from EPOC itself, but some of the evidence is scattered among reviews from other Cochrane groups. Those looking for evidence of the effectiveness of telemedicine now need to look for evidence from different clinical specialties. This is a sign of progress, and an indicator that overviews of reviews may be necessary. The clinical evidence of health informatics may have to wait until the technology is well integrated into clinical practice, but this does not necessarily help those funding, conducting, and evaluating early large scale implementations.

The value of a Cochrane review on health information technology may often be the learning associated with the process, and the questions that are raised, rather than the actual evidence that demonstrates increased (or decreased) efficiency and effectiveness of practice. Reviewers and users of reviews need to be aware of socio-technical frameworks, and the contribution of qualitative research to evaluations of health informatics interventions.

Recommended further readings

1. F. Lau, C. Kuziemsky, M. Price, J. Gardner, A review on systematic reviews of health information system studies, *Journal of the American Medical Informatics Association* **17** (2010), 637-645.

2. C.P Friedman, J. Wyatt, *Evaluation Methods in Biomedical Informatics*. Springer, New York, 2010 (2nd edition).
3. D. Gough, S. Oliver, J. Thomas, *An Introduction to Systematic Reviews*, Sage, London, 2012.

Food for thought

1. Why do so few evaluations of new IT systems use a controlled trial design or use an interrupted time series design? How can we improve this situation, to provide better evidence?
2. What type of qualitative evidence could complement a controlled trial of a health information technology, and provide evidence that is useful to clinical practice, managers and policymakers?
3. Is it worthwhile doing Cochrane reviews of health information technology? How can we ensure that health information technology development and utilisation is based on equally robust evidence to that required in other health domain areas?

References

- [1] J.P.T Higgins, S. Green, *Cochrane Handbook for Systematic Reviews of Intervention*, Version 5.1.0. Cochrane Collaboration, 2011, <http://www.cochrane-handbook.org>, last access 11 February 2016.
- [2] G.V. Glass, Primary, secondary, and meta-analysis of research, *Educational Researcher* **5** (1976), 3-8.
- [3] A. Ciapponi, S. Lewin, G. Bastias, L. Dudley, S. Flottorp, M. Gagnon, S. Garcia Marti, C. Glenton, C.A. Herrera, C.I. Okwundu, N. Opiyo, A.D. Oxman, T. Pantoja, E. Paulsen, B. Peñaloza, G. Rada, F. Suleman, C. Wiysonge, Delivery arrangements for health systems in low-income countries: an overview of systematic reviews, *Cochrane Database of Systematic Reviews* **5** (2014), CD011083, DOI: 10.1002/14651858.CD011083.
- [4] M. Nasser, Z. Fedorowicz, T. Newton, C. Van Weel, J.J. van Binsbergen, F.A. Van de Laar FA, Patients record systems: effects on dental practice and patient oral health outcomes, *Cochrane Database of Systematic Reviews* **7** (2010), CD008606, DOI: 10.1002/14651858.CD008606.
- [5] R. Currell, C. Urquhart, P. Wainwright, R. Lewis, Telemedicine versus face to face patient care: effects on professional practice and health care outcomes, *Cochrane Database of Systematic Reviews* **2** (2000), CD002098. DOI: 10.1002/14651858.CD002098.
- [6] C. Urquhart, R. Currell, M.J. Grant, N.R. Hardiker, Nursing record systems: effects on nursing practice and healthcare outcomes, *Cochrane Database of Systematic Reviews* **1** (2009), CD002099, DOI: 10.1002/14651858.CD002099.pub2.
- [7] F. Gillaizeau, E. Chan, L. Trinquart, I. Colombet, R. Walton, M. Rège-Walther, B. Burnand, P. Durieux, Computerized advice on drug dosage to improve prescribing practice, *Cochrane Database of Systematic Reviews* **11** (2013), CD002894, DOI: 10.1002/14651858.CD002894.pub3.
- [8] S. Martin, G. Kelly, W.G. Kernohan, B. McCreight, C. Nugent, Smart home technologies for health and social care support, *Cochrane Database of Systematic Reviews* **4** (2008), CD006412, DOI: 10.1002/14651858.CD006412.pub2.
- [9] A.S.M. Mosa, I. Yoo, L. Sheets, A systematic review of healthcare applications for smartphones, *BMC Medical Informatics and Decision Making* **12** (2012), 67.
- [10] J.C. Jacobson Vann, P. Szilagyi, Patient reminder and recall systems to improve immunization rates, *Cochrane Database of Systematic Reviews* **3** (2005). CD003941. DOI: 10.1002/14651858.CD003941.pub2.
- [11] E. Murray, J. Burns, S. See Tai, R. Lai, I. Nazareth I, Interactive Health Communication Applications for people with chronic disease, *Cochrane Database of Systematic Reviews* **4** (2005), CD004274, DOI: 10.1002/14651858.CD004274.pub4.
- [12] W.E. Bijker, T.P. Hughes, T. Pinch, D.G. Douglas, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, Cambridge, MA, MIT Press, 2012.
- [13] C. Ciborra, *The Labyrinths of Information: Challenging the Wisdom of Systems*, Oxford: OUP, 2002.

- [14] R.J. Holden, B-T. Karsh, The Technology Acceptance Model: its past and its future in health care, *Journal of Biomedical Informatics* **43** (2010), 159-172.
- [15] F.S. Mair, C. May, C. O'Donnell, T. Finch, F. Sullivan, E. Murray, Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review, *Bulletin of the World Health Organization* **90** (2012), 357-364.
- [16] R. Kling, G. McKim, A. King, A bit more to IT: scholarly communication forums as socio-technical interaction networks, *Journal of the American Society for Information Science and Technology* **54** (2003), 46-67.
- [17] C. Urquhart, R. Currell, Home uterine monitoring: a case of telemedicine failure? *Health Informatics Journal* **16** (2010), 165-175.
- [18] C. Urquhart, R. Currell, F. Harlow, L. Callow, Home uterine monitoring for detecting preterm labour, *Cochrane Database of Systematic Reviews* **1** (2015), CD006172, DOI: 10.1002/14651858.CD006172.pub3.
- [19] A. K Woodend, H. Sherrard, M. Fraser, L. Stuewe, T. Cheung, C. Struthers, Telehome monitoring in patients with cardiac disease who are at high risk of readmission, *Heart Lung* **37** (2008), 36-45.
- [20] R.M. Bergenstal, R.L. Anderson, D.M. Bina, M.L. Johnson, J.L. Davidson, B. Solarez-Johnson, D.M. Kendall, Impact of modem-transferred blood glucose data on clinician work efficiency and patient glycemic control, *Diabetes Technology and Therapeutics* **7** (2005), 241-247.
- [21] M.A. Loane, S.E. Bloomer, R. Corbett, D.J. Eedy, N. Hicks, H.E. Lotery, C. Mathews, J. Paisley, K. Steele, R. Wootton, A comparison of real-time and store-and-forward teledermatology: a cost-benefit study, *British Journal of Dermatology* **143** (2000), 1241-1247.
- [22] P.B. Jacklin, J.A. Roberts, P. Wallace, A. Haines, R. Harrison, J.A. Barber, S.G. Thompson, L. Lewis, R. Currell, S. Parker, P. Wainwright, Virtual Outreach Project Group, Virtual outreach: economic evaluation of joint teleconsultations for patients referred by their general practitioner for a specialist opinion, *BMJ* **327** (2003), 84.
- [23] J.C. Fortney, J.M. Pyne, M.J. Edlund, D.E. Robinson, D. Mittal, K.L. Henderson, Design and implementation of the Telemedicine-Enhanced Antidepressant Management Study, *General Hospital Psychiatry* **28** (2006), 18-26.
- [24] K.H. Dansky, J. Vasey, K. Bowles, Impact of telehealth on clinical outcomes in patients with heart failure, *Clinical Nursing Research* **17** (2008), 182-199.
- [25] J.E. Gray, C. Safran, R.B. Davis, G. Pompilio-Weitzner, J.E. Stewart, L. Zaccagnini, D. Pursey, Baby CareLink: using the internet and telemedicine to improve care for high-risk infants, *Pediatrics* **106** (2000), 1318-1324.
- [26] S. Shea, R.S. Weinstock, J. Starren, J. Teresi, W. Palmas, L. Field, P. Morin, R. Goland, R. Izquierdo, I.T. Wolf, M. Ashraf, C. Hillman, S. Silver, S. Meyer, D. Holmes, E. Petkova, L. Capps, R. Langue; for the IDEATEL Project, A randomized trial comparing telemedicine case management with usual care in older, ethnically diverse, medically underserved patients with diabetes mellitus, *Journal of the American Medical Informatics Association* **13** (2006), 40-51.
- [27] C. Sanders, A. Rogers, R. Bowen, P. Bower, S. Hirani, M. Cartwright, R. Fitzpatrick, M. Knapp, J. Barlow, J. Hendy, T. Chrysanthaki, M. Bardsley, S.P. Newman, Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: a qualitative study, *BMC Health Services Research* **26** (2012), 220.
- [28] N. Artinian, J. Flack, C. Nordstrom, E. Hockman, O. Washington, K. Jen, M. Fathy, Effects of Nurse-managed telemonitoring on blood pressure at 12 month follow-up among urban African Americans, *Nursing Research* **56** (2007), 312-322.
- [29] G. Hastings, *Primary nurse practitioners and telemedicine in prison care: an evaluation*, in S. Zoog, S. Yarnall, eds., The Changing Health Care Team, MCSA, Seattle, 1976. pp. 54-59.
- [30] C. Urquhart, R. Currell, Reviewing the evidence on nursing record systems, *Health Informatics Journal* **11** (2005), 33-44.
- [31] A.L. Cochrane, Effectiveness and Efficiency: Random Reflections on Health Services, The Rock Carling Fellowship Lecture. Nuffield Provincial Hospitals Trust, 1972, reprinted with BMJ, 1989.

Going Beyond Systematic Reviews: Realist and Meta-Narrative Reviews

Terese OTTE-TROJEL^{a,b,1} and Geoff WONG^c

^a *Institute of Health Policy & Management, Erasmus University Rotterdam*

^b *NNIT, Public & Healthcare Advisory, Copenhagen, Denmark*

^c *Nuffield Department of Primary Care Health Sciences, University of Oxford*

Abstract. Health information technologies are complex interventions whose effects differ across contexts. To improve our understanding of the effects of health information technologies, approaches are needed that utilize evidence beyond experimental results in order to provide explanatory answers to how and why a given technology works. The relatively new realist and meta-narrative review approaches are introduced as important methods in synthesising and analysing evidence in the field of health informatics. A common purpose of these two review approaches is to help create a sense of evidence about complex interventions that enables an understanding of how and why they work. A detailed description of the principles and objectives of the two types of reviews is presented. Key steps required to conduct each of the reviews are summarized, and examples of how the review approaches have been applied to topics related to health informatics are provided. Limitations of the two review approaches are discussed.

Keywords. Evaluation, evaluation methodology, health information technology, medical informatics, narrative review.

1. The value of realist and meta-narrative reviews for gathering evidence in the field of health informatics

The effect of health information technology often differs across settings. Also, there are numerous accounts of technologies that have worked in one context, but failed in another. In this contribution, we examine approaches to advancing the health informatics evidence base in ways that allow for explaining these varied effects under differing contexts.

An important notion in any such exploration of health information technologies is that these typically represent complex interventions whose effects are influenced by the interplay of several interconnected parts [1], acting in non-linear and emergent ways [2]. First, there are a large number of interacting components that affect the implementation of a given technology. Implementation of an Electronic Health Record (EHR) system in a large hospital, for example, typically entails involvement of the executive board, managerial and clinical leadership, front line physicians, nurses and other staff as well as technical, financial, customer service, and legal departments. Second, there are a large number of complex behaviours required by those delivering

¹ Corresponding author: Dr. Terese Otte-Trojel, NNIT, Public & Healthcare Advisory, Oestmarken 3A, Soeborg, Denmark, tetr@nnit.com.

services supported by an EHR as well as those receiving health care services in such a system. EHR system implementation relies on actions by people to carry out complicated and skill-demanding tasks in a coordinated fashion. Third, since the EHR is used in the work processes of clinical and technical staff throughout the hospital, sometimes in interplay with patients, a large number of groups or organizational levels must be served by the system.² Fourth, it is likely that there will be variability in outcomes of EHR use, as the EHR system may be used differently in different parts of the hospital. In some cases, system implementation may even have negative implications, for instance due to disrupted workflow. Fifth, a high degree of flexibility or tailoring of the EHR system is required. Typically, even with standard EHR products provided by large vendors, considerable customization is required to fit the EHR into the organization, and several rounds of modification can be expected as the system matures. These and further characteristics of complexity (see [2] for further methodological discussion on complexity) are important to keep in mind when studying implementation of health information technology. The multi-faceted, dynamic, and social properties of the context in which the technologies are implemented make it unlikely that a given technology will work similarly in different contexts.³

Systematic Cochrane reviews⁴ that draw on experimental studies of the effects of interventions have been conducted for several decades, proving to be indispensable for gathering evidence on effects of 'simple interventions' such as a new medication [3]. However, while Cochrane-type reviews are useful for such simpler interventions, their ability to incorporate heterogeneity across primary studies with respect to research design, characteristics of the study population, the context in which the intervention is implemented, types of interventions, and outcome indicators is limited. In fact, Cochrane reviews expressly seek to filter out all variance. Accordingly, Cochrane reviews have primarily focused on estimating the effect size of an intervention, asking questions such as 'does this intervention work and how well?' and seeking deterministic answers such as ' $a + b = c$ ' [4]. However, for most health information technology interventions such results are not meaningful. As reasoned above, complex health information technologies are embedded in open, social systems; rely on human action and interaction; and are continually affected by the organizational and socio-political context. Such technologies do not lend themselves to 'recipes'; a recipe for one context at one space in time cannot be assumed transferable to another context at another space in time. Thus, complete reliance on the Cochrane review with a relatively narrow focus on effectiveness limits our ability to build a useful evidence base in health informatics. In fact, relying solely on evidence generated from systematic Cochrane-like reviews that expressly filter out contextual influence and human factors may give decision- and policy makers only partial, or even misleading, information on which to base decisions.

Hence, there is a need for review approaches that synthesize data in a way that allows for incorporating the influence of context and dealing with heterogeneity. Further, approaches are required that can utilize evidence beyond experimental results

² See also: S. Marceglia, Domains of Health IT and tailoring of evaluation: Practicing Proces Modeling for Multi-Stakeholder Benefits, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

³ See also: K. Cresswell, Evaluation of Implementation of Health IT, in: *ibid.*

⁴ See also: C. Urquhart et al., Systematic Reviews and Meta-Analyses of Health IT, in: *ibid.*

in order to provide explanatory answers to how and why a particular intervention works. In this contribution, we look closer at such approaches for reviewing contextually relevant and real world evidence. Importantly, these review approaches do not replace Cochrane-type reviews. However, depending on the research question, they may be better suited to understanding an intervention's effects than the Cochrane review or may be provide insights that complement the findings generated by a Cochrane review.

2. Introduction to the realist and meta-narrative review approaches

The realist and the meta-narrative reviews take centre stage in this contribution, although many other types of review approaches exist. Examples include Meta-Ethnography, Grounded Theory, Thematic Synthesis, Textual Narrative Synthesis, Meta-study, Critical Interpretive Synthesis, Ecological Triangulation and Framework Synthesis (for more information about these approaches see [5]). A common purpose of the realist and meta-narrative review approaches is to "help make sense of heterogeneous evidence about complex interventions applied in diverse contexts in a way that informs policy" [6], thus allowing systematic exploration of and explanations for how and why complex interventions work. A realist review does this by building and testing theories about how a given intervention will work. We focus on realist reviews in this contribution as, unlike many other theory driven interpretive review approaches, they have a coherent analytical process that enables sense-making of the relationship between context and outcomes. Meta-narrative reviews make sense of complex interventions by elucidating and exploring the implications of different conceptualizations and applications of a given construct – an approach that is missing from many other theory driven review approaches. In the following, we give a more detailed account of the origin, philosophical principles, and objectives of the two types of reviews.

2.1 Realist review

The realist review, in common with realist evaluation [7], is based on a realist philosophy of science. Its goal is to systematically examine how contextual factors influence outcomes, through mechanisms [8]. This core aim of the realist evaluation has often been summarized in the question of 'what works, how, for whom, in what circumstances and to what extent?'. To answer these questions, realist evaluation "seeks to unpack the mechanism of how complex programmes work (or why they fail) in particular contexts and settings" [9]. Its philosophical lens is realism, which assumes that there is an external reality, but that this reality is modified through human actions and perceptions. The implication of this philosophical lens is an understanding of complex, social interventions as constantly perceived, generated and altered [10]. In the realist conceptualization of the world, this understanding is integral to explaining why some interventions work and others do not, and is therefore part of any realist review of a complex intervention. As we noted earlier, this understanding is pertinent to effectively implementing health information technology in real world systems and organizations.

A realist review is theory-driven. This implies that the review starts with articulation of key theories about how an intervention is assumed to work, which are

then explored, tested and refined in the review [11]. Such theories are called programme theories; a programme theory outlines the assumptions about how an intervention is expected to achieve desired outcomes, for whom, in what circumstances and why [8]. To develop programme theories, a realist review seeks to tease out and describe the relationships between context and mechanisms that create outcome pattern.[9] In the realist conceptualization of how the world ‘works’, mechanisms are the ‘agents of change’ that affect whether an interventions brings about any effects [10]. There are many definitions of mechanism, but they can usefully be conceptualized as hidden entities, processes, or social structures that operate in particular contexts to generate certain outcomes [12]. As such, mechanisms are seen as causal processes that tend to, but not always, occur under a particular set of conditions - activation of a mechanism is thus contingent on the context in which an intervention is implemented [4]. Context may be conceptualized as “those features of the conditions in which programmes are introduced that are relevant to the operation the programme mechanisms” [10]. These may be social, economic or political characteristics of the geographical area in which the intervention is implemented. However, it can also be more local features pertaining to the particular setting or even population receiving the intervention. Outcomes patterns are “the intended and unintended consequences of programmes, resulting from the activation of different mechanisms in different contexts” [10]. An important point to note is that context, mechanism and outcomes are linked. An outcome (O) occurs because it has been caused by a mechanism (M) that has been ‘triggered’ under specific context(s) (C) – often summarized in the heuristic $C+M=O$.

During analysis in a realist review, a feature in an intervention is only conceptualized as a context because data indicates that it has caused a specific outcome to occur through a certain mechanism. In other words realist analysis does not produce ‘free-floating’ lists of context, mechanisms and outcomes but configurations of context-mechanisms-outcomes – often referred to as CMO configurations. These form the basic explanatory building blocks of a programme theory. Within a realist programme theory there may be several CMO configurations and a complete programme theory contains an explanation of both the CMO configurations and the relationships between these [13].

Realist reviews have an explicit and coherent explanation for why it is that programme theories from one context may be relevant to another. Analysis in realist reviews focuses on the causal mechanisms found within programmes – specifically the behaviour of a mechanism in different contexts and the outcome(s) caused. Any justification for learning from or extrapolating the explanation for how and why an outcome occurs in one setting as well as in another is based on the assumption that the same mechanism(s) are found in both contexts. Any such assumption must then be tested against data.[6] As an example, a programme theory may suggest that under certain contexts a health information technology can trigger the mechanism of patient engagement to produce a certain outcome. The reviewer may be able to learn more about the behaviour of this mechanism by gathering data from other fields of research where it has been studied and not just in health informatics. The idea of focusing on mechanisms within programmes, rather than types of interventions, as the unit of analysis is especially useful in emerging areas of research where evidence on effectiveness is still limited [3].

The ultimate goal of a realist review is to provide explanatory propositions that make visible the contingencies that are likely to affect whether an intervention will

generate intended outcomes.[10] Hence, it does not produce deterministic theories that can always predict outcomes across contexts. Recommendations possible through a realist review are thus likely to take the format: 'In circumstances such as A, try B, or when implementing C, watch out for D' [9]. The analytic focus on the causal mechanisms within programmes that generate given outcomes in a given context may provide guidance to policy makers or practitioners about ways to tweak organizational structures or processes to most likely activate relevant mechanisms [4].

2.2 *Meta-narrative review*

Greenhalgh et al. developed the meta-narrative review approach in 2004. It is particularly suited to topics where there are different perspectives about the nature of the topic [14]. Thus, it is intended for use when a topic has been differently conceptualized, theorized and empirically studied by different groups of scientists. The underlying assumption in the meta-narrative approach is that key constructs mean and are valued differently to groups of scientists who (implicitly or explicitly) belong to different research traditions and/or paradigms [15]. (Please note that in this contribution we will for the sake of brevity use the term research tradition to include both research tradition and/or paradigm.) Specifically, Greenhalgh and colleagues developed the review method to help explain the apparently disparate data encountered in a review of research from a wide range of research traditions, namely diffusion of innovation in healthcare organizations. The authors found that constructs such as 'diffusion', 'innovation', 'adoption' and 'reutilization' had been conceptualized and studied very differently by researchers from a wide range of traditions including psychology, sociology, economics, management and philosophy [16].

Meta-narrative review uses a constructivist philosophical lens, which proposes that science progresses in paradigms; that is, knowledge is produced within particular research traditions, which have their own assumptions about theory, the legitimacy of study objects, research questions and knowledge [16]. As Greenhalgh et al. pointed out "an empirical discovery made using one set of concepts, theories, methods and instruments cannot be satisfactorily explained through a different paradigmatic lens" [17]. The meta-narrative review thus makes sense of complex, heterogeneous, and conflicting bodies of literature by identifying and analysing the belief systems that exist within a research tradition or paradigm. As such, a research tradition becomes the unit of analysis in the meta-narrative review [17]. Through combining and comparing findings generated within different research traditions or paradigms, an overarching narrative can be illuminated that provides a richer picture of the topic area than would be possible to obtain by including only one perspective. Key questions that a meta-narrative review will seek to answer are (1) Which research traditions or paradigms have considered this broad topic area? (2) How has each tradition conceptualized the topic? (3) What theoretical approaches and methods did they use? (4) What are the main empirical findings? and (5) What insights can be drawn by combining and comparing findings from different traditions? [14]

The meta-narrative review and the realist review share several properties [3]. However, in comparison to a realist review, a meta-narrative review deliberately focuses its analysis on the implicit and explicit assumptions, value systems, world views and so on of the researchers, and not just the theories that explain the behaviour of interventions in different contexts [6]. Like the realist review, the meta-narrative

review offers a strategy to assist decision and policy makers to use a conflicting body of research to guide decisions.

3. Description of review approaches and steps

A set of steps has been proposed to guide efforts to use the two review approaches.

3.1 Realist review

In 2005, Pawson et al. proposed guidelines for conducting a realist review, consisting of five steps of an iterative and non-linear nature [9]. These guidelines have been further expanded and detailed in Pawson [8] as well as through the RAMESES project (which will be introduced later in the contribution). In the following, we summarize the five review steps, while pointing the reader to the aforementioned resources for more thorough explanations of the steps.

Step 1: Clarify scope. The first activity in this step is to identify the review questions, which may be sharpened as new information and insights emerge. It is advised that this step is conducted in close collaboration with the commissioner(s) of the review to ensure that the findings match their needs and expectations. Subsequently, the reviewers should map the programme theory(ies) that explain(s) how the given intervention works. This mapping exercise should result in the articulation of the key programme theories that the review will explore. This may entail doing exploratory searches to come up with a list of possibly relevant programme theories that are then grouped, categorized or synthesized. This product of this step – an initial programme theory (or theories) that focuses on the needs and expectations of the commissioner(s) of the review - should then be further refined with data from documents.

Step 2: Search for relevant evidence. The search for material should be purposive and iterative, that is, geared to continually capture emerging primary research data to refine program theories. As opposed to a Cochrane review, the realist review includes all types of study designs, reasoning that information about a programme theory and other intervention processes are captured in a variety of sources including peer-reviewed quantitative and qualitative studies as well as grey literature such policy documents, business plans, and websites. In other words, a broad range of document types may be able to contribute to programme theory refinement in a realist review. The search process may comprise four iterative search strategies. The first strategy is an exploratory background search to scope the literature. As programme theories start to emerge, the reviewers may then be able to refine inclusion criteria, thus further focusing the search. Upon agreement on a shortlist of programme theories, purposive searching may be applied to explore and test the corresponding hypotheses, making extensive use of snowballing. Searching may continue to be needed even when the review may be close to completion, as additional data may continually be needed to refine programme theory.

Step 3: Select and appraise documents and extract data. From searching the literature, reviewers will hopefully have identified documents that may possibly contain data that might be useful for programme theory refinement. Reviewers still need to decide if a document does in fact contain the data needed data. In practice, this

process of determining if a document contains the data needed often takes place when reading the full text of the document. Document selection in a realist review, is thus based on relevance. Relevance refers to whether any document retrieved during searches can contribute data to build or test a certain programme theory or aspects of it. For any relevant data, a judgment has to be made about the rigour of the method(s) (if any) that has been used to produce the data. In other words, rigour concerns the credibility of the method(s) used to produce the pieces of data. In extracting data from the included material, the realist logic of analysis plays an important structuring role: data should be extracted about programme theories, context, mechanisms and outcomes configurations that will help in programme theory refinement.

Step 4: Synthesize evidence and draw conclusions. The goal of the realist review is to use data to build one or more programme theories that explain what has caused the outcome patterns observed under different contexts; that is, why an intervention generates particular outcomes in particular contexts through one or more mechanisms. Theory building entails making inferences about 'CMO' configurations and the place of these configurations within a programme theory (or theories). Reviewers should thus make clear how they derived to such inferences and what data have been used to develop and support them. In addition clarity is needed on how what they have conceptualized as context, mechanism and outcomes with relevant data. The value of such clarity is that readers and users of the review's findings are provided with transparency – they can see for themselves the links from data to programme theory. Since the findings from realist analysis are dependent on contextual factors, conclusions should not lead to deterministic formulae. Instead they must indicate the contingencies upon which it is expected that a mechanism will be triggered in a context to process certain outcomes such as 'If A, then B' or 'In the case of C, D is unlikely to work'.

Step 5: Disseminate, implement and evaluate. With the audience in mind, reviewers should explain the relevance of one or more key programme theories that emerged from the analysis and highlight the strength of evidence for the main conclusions. In doing so, the reviewers may want to test out recommendations and conclusions with key stakeholders in order to place focus on actions that are feasible in the given policy context. This will entail dialogue with practitioners and policy-makers to apply recommendations in particular contexts.

3.2 Meta-narrative review

Based on experiences deriving from conducting the first meta-narrative review, Greenhalgh et al. [17] suggested a set of phases that should be followed to conduct a meta-narrative review. In common with realist review, iterative refinements are often needed - it is normal and appropriate for the review objectives, question and scope to transform as the review progresses as the key research traditions are uncovered and become better understood. This review approach is guided by six principles. These are: pragmatism (reviewers should include what makes most sense for the intended audience); pluralism (the topic should be illuminated from multiple perspectives); historicity (deepest understanding of a topic comes from studying its evolution over time); contestation (conflicting data from different research traditions should be examined to generate higher-order insights); reflexivity (reviewers should continually reflect on the emerging findings); and peer-review (emerging findings should be presented and discussed with an external audience).

We summarize here the phases from Greenhalgh's approach (more details maybe found in [17]).

Step 1: Planning. In this initial phase, a preliminary scoping of the literature can be done to discern the research traditions dominant in the field of review. Then, a multidisciplinary research team should be formed. The guiding principle for the team composition is that the researchers' scientific backgrounds cover all relevant research traditions identified through the scoping of the literature. The team should formulate some initial, broad research questions, and should agree with the review's commissioners about the desired outputs. As a last step in the planning phase, the team should plan regular meetings, which may also include representatives of the review's intended audience.

Step 2: Search. The search process should start with initial and intuitive browsing of literature and consultation with experts and stakeholders in each research tradition, with the aim of identifying the array of perspectives and approaches within each relevant research tradition. Reviewers may want to consider the parameters of each tradition such as its scope, historical roots, key constructs and assumptions, commonly asked research questions and methods used, main empirical findings, and strengths and limitations. Upon reaching agreement about key narratives emerging through this initial search, the reviewers should conduct a systematic search for empirical papers within each research tradition. The reviewers can make use of electronic and paper-based databases and sources as well as tracking references of seminal conceptual references to accumulate a relevant subset of primary research.

Step 3: Mapping. This phase, which often takes place in parallel with the searching phase, entails mapping the key conceptual, theoretical, methodological, and instrumental components of each research tradition. Furthermore, reviewers should outline key actors and events in the unfolding of the tradition over time, including main findings and how they came to be discovered, using the prevailing narrative styles used by scientists in the selected traditions.

Step 4: Appraisal. Using the quality criteria determined by experts within a study's particular research tradition, the review team should appraise the quality of each primary study. The purpose of the quality appraisal is to determine which studies within a tradition are considered to be examples of high quality. These will then be analysed to ascertain what data they can contribute to building a narrative of that research tradition. Then, data elements should be extracted that can contribute to constructing a narrative of how research on a topic evolved in a particular tradition. The review team may want to develop a data extraction form to summarize key data items such as the papers' research question, theoretical basis, study design, validity and robustness of methods, sample size and power, nature and strength of findings, and validity of conclusions for each empirical study.

Step 5: Synthesis. In the synthesis phase, the review team should compare and contrast all the key dimensions of the topic that have been researched within each research tradition. The aim is to generate higher-order data by comparing conflicting findings and explaining them in terms of contestation between the different traditions from which the findings were generated. Greenhalgh et al. propose that this synthesis is guided by a set of questions, such as (1) What is the range of research questions that different groups of scientists have asked about each of the dimensions of the research paradigm? Can these questions be grouped meaningfully and classified across traditions? (2) What are the commonalities of research findings across paradigms, and where the empirical findings from different paradigms are conflicting, to what extent

can discrepancies be explained? (3) Given the ‘rich picture’ of the topic area achieved from these multiple perspectives, what are the overall key findings and implications for practice and policy? and (4) What are the main gaps in the evidence on this topic and where should further primary research be directed?

Step 6: Recommendations. Through reflection, multidisciplinary dialogue and consultation with the intended users of the review, the review should distil and discuss recommendations for practice, policy and further research.

To ensure that realist reviews and meta-narrative reviews are being executed consistently and rigorously, checklists and publication standards have been proposed for both approaches. Specifically, in 2013 the Realist And MEta-narrative Evidence Syntheses: Evolving Standards (RAMESES) for reporting realist and meta-narrative reviews were published.[11,14] Further, quality criteria have been set forth, which should be used to determine the rigour with which a review has been done. For a complete overview of this methodological work dedicated to advancing the review approaches, the interested reader can benefit from visiting the RAMESES project website [18] and reviewing the project outputs. Researchers interested in realist and meta-narrative approaches may also wish to consider joining the RAMESES email listserv (www.jiscmail.ac.uk/RAMESES)

4. Application of realist and meta-narrative reviews in health informatics

The realist and meta-narrative review approaches have been applied on topics as diverse as the effects of joint health and safety committees, school feeding programmes, diabetes education programmes, innovations in health service delivery and organization, and the role of urban municipal governments in reducing health inequities. However, thus far, the application of the realist and meta-narrative reviews in the field of health information has been limited. In the following we highlight an example of a realist review (case A) and a meta-narrative review (case B) of literature in health informatics to demonstrate the nature of the evidence they can generate.

Case A: Internet-based medical education: a realist review of what works, for whom and in what circumstances.

Wong et al. [19] conducted a realist review of literature on Internet-based medical education, which is increasingly offered by means of online courses. The objective of the review was to contribute to a limited evidence base on what actually works when offering medical education via the Internet. Specifically, it aimed to generate recommendations that could inform the development and assessment by (potential) users of Internet-based medical courses. The realist review included 249 studies that were believed to be relevant to testing two main theories to explain variation in course-takers’ satisfaction and outcomes: Davis’s Technology Acceptance Model and Laurillard’s model of interactive dialogue. Studies were included with various designs and outcomes as long as they addressed interventions that used the Internet to support learning, involved doctors or medical students, and reported a formal evaluation. The included material was used to test and refine the two main theories. The review established that course-takers are more likely to follow a course if they perceive the course to have advantage relative to non-Internet alternatives, high ease of use and compatibility with their values and norms. Further, the review found that ‘interactivity’ can result in effective learning, but only if course-takers receive formative feedback,

for example through dialogue with a tutor, fellow students or virtual tutorials. Hence, in designing and/or choosing an Internet-based medical course, it is important to consider how the interaction between technology and course-taker can be made most meaningful. This will require efforts to improving the fit between the technical attributes of the course and the needs and priorities of the course-takers. As a way of offering recommendations, the review provides a set of questions that can spur helpful lines of thought. Due to the varying contexts in which internet-based medical courses are implemented, these questions are not considered to offer deterministic guidance that will always generate desirable outcomes.

Case B: Tensions and Paradoxes in Electronic Patient Record Research: A Systematic Literature Review Using the Meta-narrative Method.

Greenhalgh et al. [20] conducted a meta-narrative review on Electronic Patient Records (EPRs) with the goal of explaining how the Electronic Patient Record and its implementation had been conceptualized and studied in different research traditions. The review team included scholars representing key research traditions that had addressed the topic including health information systems, change management in health services research, information systems, computer-supported cooperative work, and more. The review included twenty-four systematic review and ninety-four primary studies that helped develop an understanding of the key constructs related to EPRs within each of these research traditions. In the synthesis, key tensions between the constructs in the different research traditions were identified and illuminated. The identified tensions centred on the conceptualization of seven constructs: (1) The EPR (whether it was thought as a 'container' or 'itinerary'); (2) The EPR user (as either an 'information-processer' or 'member of socio-technical network'); (3) The organizational context (as 'the setting within which the EPR is implemented' or 'the EPR-in-use'); (4) Clinical work and knowledge (as 'decision making' or 'situated practice'); (5) The process of change (as 'the logic of determinism' or 'the logic of opposition'); (6) The impact of change and definition of success (as 'objectively defined' or 'socially negotiated'), and lastly, (7) Definition of complexity and scale ('the bigger the better' or 'small is beautiful'). The findings raise questions about the positivistic assumptions typically underlying EPR implementations by bringing forth insights from a variety of perspectives. For example, the findings suggest that EPR use will always require human input to re-contextualize knowledge. Further, even though administrative tasks may be made more efficient by the EPR, primary clinical work may become less efficient, since paper-based recording of information may provide more flexibility on the work floor. Lastly, contrary to a widely held belief, smaller EPR systems may sometimes be more efficient and effective compared to larger ones. Hence, these findings from outside and inside the health informatics research tradition offer food for thought for EPR researchers and policy-makers that can be considered in their future EPR work.

5. Limitation of review approaches

The preceding text has focused on the potentials of the realist and the meta-narrative reviews and has demonstrated their application in the field of health informatics. This has included two specific examples that show the potential usefulness of the approaches to help make sense of heterogeneous evidence about complex interventions

in diverse contexts. Yet, at the same time, there are some limitations to the review approaches that may restrict their application and relevance for busy decision- and policymakers. In the following we highlight some main limitations as they have been discussed in the literature.

An important limitation is the extensive resources and expertise that may be needed to conduct these reviews. Both approaches may entail exploration and appraisal of literature from a potentially vast number of disciplines. Further, refining theories and narratives can become an infinite task as new information can be expected to emerge continually. It is thus not always clear when saturation has been reached. In addition to the time-consuming nature, both approaches (and especially the meta-narrative) require cross-disciplinary expertise, which can be difficult to muster in practice. Moreover, although quality standards and training materials have been proposed for conducting the reviews, there is no fixed formula that can be used slavishly to generate findings. The implications are that due consideration needs to be given by any researchers wishing to use these approaches on what the scope and focus of their review will be. As we have shown, the answers that can be generated through these review approaches tend to be complex and contingent upon several factors. Thus, answers achievable from these types of reviews may be relatively difficult for decision- and policymakers to utilize in practice [10]. Thus, a challenge for anyone using these approaches will be to make sure that their answers are useful to their intended audience. Hence, considerable time is likely to be needed to think through how to present the answers to the questions asked. In addition, as with all reviews, time and effort will need to be spent to plan and put together a review team with the relevant expertise.

Up until 2013, when the RAMESES protocols were published, there was little methodological guidance for reviewers and appraisers of realist or meta-narrative reviews. While the advent of the RAMESES protocol is likely to have improved clarity and consistency, such protocols cannot remove reviewers' interpretive judgments, which are integral to the synthesis process [17]. Due to the subjective nature of the approaches, it is questionable whether another review team would arrive at the same results. Thus, to ensure validity, much effort must be put into providing transparency into the processes used to develop and refine theories, thereby allowing the reader to understand clearly how the review was carried out and what data were used. While not necessarily permitting reproducibility, this transparency is important to allow readers to pinpoint and debate exactly where their disagreement lies with the findings of an approach.

Lastly, while it is considered a strength because it allows for inclusion of more studies with explanatory power, the inclusion of weaker study designs (such as case study reports) can arguably limit the inferences that might be drawn from the data. Moreover, many studies do not report detailed information about the interventions and the processes surrounding the interventions, making it hard to build, test and refine theories. As for all other reviews realist and meta-narrative reviews are only as good as the primary data on which they build.

6. Conclusion

The realist and meta-narrative review approaches hold great potential to complement traditional Cochrane-type reviews. The realist review focuses on unpacking the relationships between contexts and mechanisms that cause an intervention's outcomes.

The meta-narrative review seeks to create clarity on the conceptualization of complex topics where there is a lack of clarity or disagreement about key constructs. Both review approaches aim to inform decision and policymaking in complex policy areas. The review approaches have been applied in various research fields and disciplines, and their use may be expected to increase with the publication of the RAMESES quality standards and training materials, in that these likely make it easier to conduct and publish realist and meta-narrative reviews. Yet, so far their application in the field of health informatics has been sparse. This may be related to the fact that both approaches are relatively new, and to their potential limitations such as their time-consuming nature and the challenges of producing 'simple' recommendations.

Acknowledgements

The authors are thankful to Professor Thomas G. Rundall for his helpful comments on an earlier draft of this contribution.

Recommended further readings

1. R. Pawson, *The Science of Evaluation: A realist manifesto*, Sage, London. 2013.
2. T. Greenhalgh, G. Robert, F. Macfarlane, P. Bate, O. Kariakidou, R. Peacock, Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review, *Social Science & Medicine* **61** (2005), 417-430.
3. *Outputs (including publications standards, methodological guidance, and training materials) from the RAMESES projects*, http://www.ramesesproject.org/index.php?pr=Project_outputs, last access 11 February 2016.
4. R. Anderson, New MRC guidance on evaluating complex interventions, *BMJ* **337** (2008), a1937.
5. R. Pawson, N. Tilley, *Realistic Evaluation*, Sage, London. 1997.

Food for thought

1. In your day-to-day work can you think of any circumstances when the answers produced through a realist review might be more useful than those from a Cochrane systematic review?
2. Can you think of one or more (contested) topics within health informatics that could benefit from a meta-narrative review?
3. What are the most important skills required to conduct realist or meta-narrative reviews? Do you possess these skills?
4. If you wanted to conduct a realist or meta-narrative review, how would you start?
5. How would you tailor the findings of your realist or meta-narrative review so that they are understandable and useful for your specific audiences (e.g. policy and decision makers, other researchers and so on)?

References

- [1] P. Craig, P. Dieppe, S. Macintyre, S. Michie, I. Nazaret, M. Petticrew, Developing and evaluating complex interventions: the new Medical Research Council guidance, *BMJ* **337** (2008), a1655.
- [2] R. Anderson, New MRC guidance on evaluating complex interventions, *BMJ* **337** (2008), a1937.
- [3] B. Snistveit, S. Oliver, M. Vojtikova, Narrative approaches to systematic review and synthesis of evidence for international development policy and practice, *Journal of Development Effectiveness* **4**(3) (2012), 409-429.
- [4] J. Rycroft-Malone, B. McCormack, A.M. Hutchinson, K. DeCorby, T.K. Bucknall, B. Kent, A. Schultz, E. Snelgrove-Clarke et al., Realist synthesis: illustrating the method for implementation research, *Implementation Science* **7** (2012), 33.
- [5] E. Barnett-Page, J. Thomas, Debate. Methods for the synthesis of qualitative research: a critical review, *BMC Medical Research Methodology* **9** (2009), 59.
- [6] T. Greenhalgh, G. Wong, G. Westthorp, R. Pawson, Protocol - realist and meta-narrative evidence synthesis: Evolving Standards (RAMESES), *BMC Medical Research Methodology* **11** (2011), 115.
- [7] R. Pawson, N. Tilley, *Realistic Evaluation*, Sage, London, 1997.
- [8] R. Pawson, *Evidence-Based Policy: A Realist Perspective*, Sage, London, 2006.
- [9] R. Pawson, T. Greenhalgh, G. Harvey, K. Walshe, Realist review – a new method of systematic review designed for complex policy interventions, *Journal of Health Services Research & Policy* **10**(Suppl 1) (2005), 21-34.
- [10] R. Pawson, N. Tilley, Realist Evaluation, 2004. http://www.communitymatters.com.au/RE_chapter.pdf, last access 11 February 2016.
- [11] G. Wong, T. Greenhalgh, G. Westthorp, J. Buckingham, R. Pawson, RAMESES publication standards: realist syntheses, *BMC Medicine* **11** (2013), 21.
- [12] B. Astbury B, F.L. Leeuw, Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation, *American Journal of Evaluation* **31**(3) (2010), 363-381.
- [13] G. Wong, T. Greenhalgh, G. Westthorp, R. Pawson, Development of methodological guidance, publication standards and training materials for realist and meta-narrative reviews: the RAMESES (Realist And Meta-narrative Evidence Syntheses - Evolving Standards) project, *Health Services and Delivery Research* **2**(30) (2014).
- [14] G. Wong, T. Greenhalgh, G. Westthorp, J. Buckingham, R. Pawson, RAMESES publication standards: meta-narrative reviews, *BMC Medicine* **11** (2013), 20.
- [15] T. Greenhalgh, *Meta-narrative mapping: a new approach to the systematic review of complex evidence*, in: B. Hurwitz, T. Greenhalgh, V. Skultans (eds.), *Narrative Research in Health and Illness*, Wiley. 2008.
- [16] T. Greenhalgh, G. Robert, F. Macfarlane, P. Bate, O. Kariakidou, Diffusion of innovations in service organizations: systematic review and recommendations, *The Milbank Quarterly* **82**(4) (2004), 581–629.
- [17] T. Greenhalgh, G. Robert, F. Macfarlane, P. Bate, O. Kariakidou, R. Peacock, Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review, *Social Science & Medicine* **61**(2) (2005), 417-430.
- [18] RAMESES project website, <http://www.ramesesproject.org>, last access 11 February 2016.
- [19] G. Wong, T. Greenhalgh, R. Pawson, Internet-based medical education: a realist review of what works, for whom and in what Circumstances, *BMC Medical Education* **10** (2010), 12.
- [20] T. Greenhalgh, H. Potts, G. Wong, P. Bark, D. Swinglehurst, Tensions and Paradoxes in Electronic Patient Record Research: A Systematic Literature Review Using the Meta-narrative Method, *The Milbank Quarterly* **87**(4) (2009), 729–788.

This page intentionally left blank

Part III: Ensuring the Relevance and Application of Evidence

This page intentionally left blank

Quality of Health IT Evaluations

Pirkko NYKÄNEN^{a,1} and Johanna KAIPIO^b

^a*University of Tampere, School of Information Sciences, Tampere, Finland*

^b*Aalto University, School of Science, Espoo, Finland*

Abstract. Health IT evaluation studies have often been found to be of limited quality. To address this problem, several guidelines and frameworks have been developed as tools to support improvement of the quality of evaluation studies. In this contribution, we review available guidelines and then present the Good Evaluation Practice Guideline in Health Informatics (GEP-HI) in more detail. GEP-HI is a comprehensive guideline which supports especially planning and execution of a health IT evaluation study. The GEP-HI guideline helps to overcome the quality problems related to weak study planning and methodological study design. We also discuss application of GEP-HI on an evaluation project and discuss the need to publish systematically following the recognised publication guidelines. Finally we discuss the future trend on multi-method evaluation approaches.

Keywords. Evaluation, quality, scope, guideline, framework, health IT.

1. Introduction

Evaluation is the means to assess the quality, value, effects and impacts of health IT in the health care environment. Evaluation is defined as the “act of measuring or exploring properties of a health information system, in planning, development, implementation, or operation, the result of which informs a decision to be made concerning that system in a specific context” [1, p. 480]. Evaluation offers methods and tools to collect evidence about the benefits, quality, effects and impacts of health IT.

In health care practice health IT applications offer challenging opportunities to improve the health care system’s functioning, effectiveness and outcomes as well as health care services quality and delivery, but there are also problems and unanticipated effects related to the use of IT [2,3].² It is of utmost importance in the health care environment that health IT provides the optimal and safe results and therefore health IT applications need to be evaluated with robust methodologies and evaluation results are to be reported following structured reporting standards [4,5].

Evaluation is difficult, it deals with values and norms and various organizational contexts and stakeholders’ interests, and it has to fight for funding and support [6]. Additionally, many potentially applicable methodologies and methods exist and evaluation results are to be analysed and interpreted in the study context [4]. A challenge to improve the quality of evaluation studies is to apply a systematic

¹ Corresponding author: Prof. Dr. Pirkko Nykänen, University of Tampere, School of Information Sciences, PO Box 607, FIN-33014 University of Tampere, Finland, Pirkko.Nykänen@uta.fi.

² See also: F. Magrabi et al., Health IT for patient safety and improving the safety of health IT, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

approach, plan the evaluation study carefully and execute the study following systematic evaluation guidelines [7].

In this contribution we review the focus and scope of the published evaluation studies, discuss the quality problems related to these studies, and present health IT evaluation guidelines and frameworks, namely the GEP-HI guideline.

2. The scope and quality of evaluation studies

The quality of an evaluation study is dependent on many factors, e.g. on the objectivity of the study and on the independence of evaluators, referring to their independence on economic interests, on intellectual interests and on the various stakeholders' interests. An evaluation study must also be scientifically well-established on robust theories and methodologies³, and the study should be performed following the principles of scientific research.

As there are many interesting focuses for evaluation, e.g. economics, efficiency, usability of health IT, safety, privacy and security, compliance with the clinical process, functionality of health IT, effects and impacts on health care outcomes, there is also need to use many different methods suitable for measuring the evaluation criteria of interest. These potential methods cover e.g. qualitative and quantitative methods, statistics, heuristics, ethnography, human-system interaction observations, data mining and quality analysis, cost-effectiveness and cost-benefit analyses.⁴ There are also many potential perspectives for evaluation, representing various stakeholders' viewpoints, e.g. managerial, clinical, technical, and these viewpoints may be studied at various levels of health care system – local, regional and national, or even at EU or international levels.⁵ It is of utmost importance, when planning an evaluation study, to elaborate and define how these issues are related to the study: Coverage of scientific robustness, relevance to the current purpose of the study, best fit of the important characteristics to the specific current need and relevance of the methods. The quality also comprises the quality in publication of the study.

Evaluation studies have been performed since 1960's, however, most with a rather narrow scope, often focusing on how health IT systems are related to professionals' roles, change management and user involvement [8]. Later the studies have covered also the success aspects and lessons learned in implementation and development of health IT systems. In many cases evaluation has been led by research interests to develop methodologies or to study the health care processes. From 1980's onwards also management issues, user acceptance and adoption of health IT systems in health care organizations have been studied. Kaplan and Shaw made a review of how aspects related to people, organizational and social issues have been considered in health IT evaluations [9]. They emphasized the need to pay more attention to these issues during

³ See also: J. Brender, Theoretical basis of health IT evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁴ See also: D. Luzzi et al., Economic evaluation of health IT, in: *ibid.*

⁵ See also: L. Lee et al., Understanding stakeholder interests and perspectives in evaluations of health IT, in: *ibid.*

health IT system design, implementation and use and emphasize the need to integrate multi-method evaluation to the whole life cycle of the health IT system.⁶

Van der Loo analysed evaluation studies published between 1967 and 1995 with regard to type of the system, study design, data collection, economic evaluation and type of effect measure [10]. He found performance of users to be a criterion in many studies, as well as time savings and costs of patient care. By contrast, user satisfaction was an evaluation criterion only in 11% of studies. These results reflect well the situation in 1990's, usability and effects and impacts of health IT on health care outcomes and quality were not highly valued issues that time.

Ammenwerth and de Keizer [11] found that explanatory research and quantitative methods have dominated evaluation studies during the last 20 years. However, the studies on outcome quality and costs of patient care, patient satisfaction and patient behaviour have received more attention recently as well as studies on quality of processes. At the same time, the number of laboratory studies and technical evaluation studies has declined while the number of studies focusing on the influence of health IT on quality of care processes or outcome of patient care has increased. It was typical for early studies, e.g. evaluation of decision support systems and expert systems, that the focus was on technical issues, on hardware and software quality and on system performance issues. Rigby [12] also noted that the focus of evaluation of a health IT system changes during its life cycle: During the implementation phase evaluation addresses often technical aspects, but with a completed system focuses on impacts on patient care. Ammenwerth and de Keizer found also that the number of inter-organisational studies has increased reflecting the trend towards cooperative and shared care. In many recent studies, user satisfaction and efficiency of patient care are the most frequently addressed evaluation criteria, and there is a slowly growing trend towards evaluation studies covering more than one evaluation criterion [11]. Also the health IT developments discussed in [12,13] showed the importance of evaluation, the challenges and the problems in integrated methodologies and in endorsing the importance of robust studies.

These reviews show that explanatory research and quantitative methods have dominated health IT evaluation research for a long time, but studies focusing on process quality or outcome quality of patient care have increased lately. Also human, social and organizational aspects have been included in evaluation studies to some extent, and promisingly research in this aspect is growing [9].⁷ Qualitative methods have been used in studies focused on user acceptance, usability and usefulness or on organizational and social impacts of health IT. Many evaluations have been focused on one user's individual context: how a health professional uses the system in the specific context, how he/she accepts the technology and how technology fits into his/her work processes, leaving the more wide usage contexts untouched.

The reviews also reveal quality problems in evaluation studies: there are problems related e.g. to weak study planning without a systematic, scientific methodology, false implicit assumptions made in the study, experimental errors in the research setting, under- or over-interpretation of the results or false conclusions, inclusion of non-

⁶ See also: P.J. Scott et al., Mixed methods: a paradigm for holistic evaluation of health information system, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁷ See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: *ibid.*

neutral evaluators, intra- or inter-organizational variability in the evaluation object or novelty of technology. In some cases there seem also to be problems in selecting such evaluation methods that are capable to measure those variables and aspects that describe the phenomenon under study. These quality problems show that there is real need to improve the quality of evaluation studies.

Since the early 2000's researchers in the health informatics field have pointed out the need to include also user perspective and usability, in the evaluation studies [14, 15]. This need is motivated by recognition of the importance of good usability as the health IT systems are planned for use in health care clinical practice.⁸ We analyse briefly also the published usability evaluation studies, on their methodological approach and quality problems.

A review in [16] showed that in general, empirical usability studies are heavily affected by the traditional approaches to evaluation of human-computer interaction. Usability studies have mainly applied traditional usability evaluation methods, particularly usability testing and inspections. A considerable number of usability evaluation studies have concentrated on the later phases of health IT development, and evaluation has been focused on systems that are already in use. Recent usability studies are characterized by a narrow focus on user and usability issues, emphasis on summative evaluation rather than on design or development, isolated system development, and emphasis on information systems and data management [16]. Also in many usability studies there is a lack of understanding of the contextual aspects of usability, and the characteristics of clinical work contexts, though the widely known definitions for usability [17,18] emphasize the need to understand usability as a contextual property.

There are quality problems also in the usability evaluation studies, e.g. the studies often focus on a single end-user group perspective, user interface components, or use of the system in a specific context, but do not provide a comprehensive picture of usability of a large-scale healthcare information system. Further, the evaluation studies rarely discuss the relationship between single-system development and the existing technology setting in healthcare, or the characteristics of various use contexts in which the evaluated system is used. Typically the studies discuss summative usability results on working systems; this leads to a focus on the problems with adopting current systems in a given healthcare environment and diverts concentration from the design or development of new, better systems. In general, the quality of usability evaluation studies could be enhanced by applying appropriate usability evaluation methods other than traditional ones, e.g. contextual inquiry [19], which takes into consideration the contextual aspects of healthcare environment. The planning of evaluation studies should aim at providing evaluation results but also understanding of the underlying problems and support for the development work.

Usability evaluation studies have in most cases been done separately from other evaluation studies, e.g. from effectiveness, effects and impacts evaluation studies. However, it would be beneficial to integrate usability and other evaluation criteria in a study. Usability is an important quality aspect of health IT and usability is heavily dependent on good health IT system design, implementation and adoption [20]. Some of the recent usability studies have already emphasized a more holistic view for

⁸ See also: R. Marcilly et al., From usability engineering to evidence-based usability in health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

usability evaluation and called for multi-method approaches and focus more on health professionals' clinical collaboration. Among others, Kushniruk [21, 22] has expressed a concern that when the technology applications become more complex, evaluation methodologies will need to be continually refined in order to keep pace. Multi-method approaches are called for as opposed to using one single method [23]. Kaipio emphasizes [16] that the user-centred approach will play an essential role in future health IT design and development, because knowledge and understanding of the needs of various user groups, as well as clinical use contexts, are important and should be part of all health IT design and development phases.

These findings are supported in [24] by emphasizing the need for predictive evaluation methods to accurately identify usability issues that arise from the interaction, sharing and communication requirements of clinical work. Furthermore, health IT systems have many user groups and the users use the systems for various purposes, such as clinical, nursing, medical, administrative, managerial, statistical and economical purposes and today, also patients and citizens use health IT systems for their personal health and wellness purposes. Therefore, it is important to consider the variety of uses and user contexts, and also the variety of evaluation criteria, since the nature of clinical work, as well as the physical and organizational environments in the workplace, may differ significantly between healthcare units.

This state of affairs of evaluation studies and quality challenges emphasizes the need for systematic approaches and guidelines to design and to carry out different kinds of evaluation studies to provide evidence about the impacts and actual efficiency, quality, usability and safety of health IT [25].

3. Guidelines and frameworks for health IT evaluations

3.1. Overview on available guidelines and frameworks

Many frameworks and guidelines developed for evaluation exist, aiming at supporting and improving studies so that health IT evaluation is conducted to the highest methodological and scientific standards. These frameworks and guidelines differ in terms of generality, specificity and timing related to system development phases and theoretical underpinning. In this section, we will introduce some selected frameworks and guidelines for health IT evaluation.

Kaplan has suggested an evaluation framework of 4C - Communication, Control, Care and Context [26]. This 4C model calls for multi-method longitudinal design of formative and summative evaluations.⁹ Shaw has introduced the CHEATS-framework [27] which identifies six aspects to be important in evaluation: Clinical, Human and Organizational, Educational, Administrative, Technical and Social.

A Model for ASsessment of Telemedicine applications (MAST) [28] lists aspects of evaluation within seven domains of outcomes: health problem and characteristics of the application; safety; clinical effectiveness; patient perspectives; economic aspects; organizational aspects; and socio-cultural, ethical and legal aspects. MAST is planned to be a toolkit, a checklist of issues that need to be considered in evaluation. MAST is

⁹ See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

based on a health technology assessment (HTA) approach and focuses on telemedicine systems, not on health IT generally. MAST does not consider the execution and management of an evaluation project.¹⁰

Cusack et al. developed the AHRQ toolkit [29] to provide step-by-step guidance for developing evaluation plans for health IT projects. AHRQ assists evaluators in defining the goals for evaluation, in identifying what is important to stakeholders, what needs to be measured to satisfy stakeholders, what is realistic and feasible to measure, and how to measure these items. The AHRQ toolkit is very useful from the methodological point of view. It can be applied within other more generic guidelines. The toolkit does not, however, give guidance on the evaluation project itself, how to manage it, how to carry out the project, or how to complete and report the study.

A life-cycle framework for evaluation by Clarke et al. [30] is focused on how to evaluate health IT interventions while the system is being designed, developed and deployed. The model is formative and relates evaluation to the phases of the system development. The life-cycle evaluation framework is a valuable tool to monitor the development process and the deployment of a new system.

The HOT-fit evaluation framework [31] considers Human, Organization and Technology factors and recognizes interrelated dimensions of health IT success and determines both benefits and satisfaction. 'Fit' in the framework concerns the ability of health IT system, stakeholders and clinical practices to align with each other.

ISO standards of human-centred design include guidelines for usability evaluation studies [18,32] and state that the following information is needed: Description of the intended goals, a description of the components of the context of use including users, tasks, equipment and environments, target or actual values of effectiveness, efficiency, and satisfaction for the intended contexts. Evaluation should involve the following tasks: Allocating of resources, planning of the evaluation, carrying out sufficient testing, analysing the results and prioritizing issues and proposing solutions, and communicating the solutions appropriately [32].

Many analyses (e.g. [33-34]) have emphasized that we need to understand the changes that health IT systems bring into a complex health care system. The socio-technical assessment tool, STAT-HI [35], focuses on the socio-technical aspects of systems and their implementation in the health care organizational environment.

A recent contribution for health IT evaluation is made by WHO in the context of health IT systems for developing countries [36] (on this topic, see also *Chapter 26: Evaluation of Health IT in low-income countries*). This comprises of nine high-level principles, the Bellagio principles, which cover e.g. evidence-based culture, high quality evaluative data collection and stakeholder engagement (see also *Chapter 1: The need for evidence in health informatics*). The principles emphasize the overall responsibility to ensure that health IT solutions and policies are subjected to, and informed by, rigorous evaluations and that evaluation findings should be used and contribute to evidence generation, synthesis and documentation, including peer-reviewed articles.

¹⁰ For details on MAST, see: P. Doupi, Evolving Health IT Systems Evaluation: The Convergence of Health Informatics and HTA, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

3.2. GEP-HI guideline to design and execute an evaluation study project

The discussed frameworks and guidelines for evaluation aim at improving the quality of health IT evaluation studies. They provide guidance and support on how to plan and perform evaluation studies, but their application scopes and contexts vary as well as theoretical foundations and methodological approaches. Thus a comprehensive evaluation guideline was developed to support especially planning and execution of an evaluation study to overcome the quality problems related to weak study planning and methodological choices. The guideline was named GEP-HI, Good Evaluation Practice Guideline for Health Informatics [7].

The starting point for the GEP-HI guideline development was the existing knowledge, experience and literature on evaluation studies, methodologies, guidelines development, codes of ethics and good implementation practices. In particular the following review materials and textbooks provided the foundation for preparation of the guideline [1, 9, 37-39].

The GEP-HI guideline has been developed through an informal consensus-seeking process, without balloting, in the community of health IT evaluation experts, and it has been regularly in open discussion through the HISEVAL website (<http://iig.umit.at/efmi>) and many conference workshops. The primary authors of GEP-HI were all participants of the ESF HIS-EVAL Workshop [4, 5] and active in the EFMI (European Federation for Medical Informatics) and IMIA (International Association of Medical Informatics) working groups dealing with evaluation of health IT systems (for the details on these working groups, see the book appendix).

The objective of the GEP-HI is to give advice on how to design and carry out evaluation studies in various health IT contexts [7]. The guideline lists issues to consider, and gives recommendations on how to design evaluation studies, how to make methodological choices, how to conduct studies and how to define evaluation criteria at specific phases of the health IT life cycle.

When applied, the GEP-HI guideline has potential to raise the quality of evaluation studies through careful planning, and thus contribute to the accumulation of the scientific evidence base. GEP-HI is complemented by the STARE-HI – Statement on Reporting of Evaluation Studies in Health Informatics [38], which provides guidance on how to report an evaluation study (for details on STARE-HI, see *Chapter 24: Publishing health IT evaluation studies*).

The GEP-HI guideline [7] is divided into parts corresponding to the phases of an evaluation study (Figure 1). The theoretical background for the study phases is compliant with the information system development models. Implementation is an iterative spiral; the topics are in general repeated in depth or breadth to achieve progress during all phases, and feedback loops urge to revisit earlier phases when new aspects, additional information, or changes in context appear.

The phases of GEP-HI guideline are:

- *Preliminary outline* presenting the purpose of the study and the first ideas on why, for whom, and how the evaluation should take place,
- *Study design* clarifying the design issues for the evaluation study.
- *Operationalization of methods* making the methodological approach and methods concrete and compliant with the system type, the organization and the information need.
- *Project planning* developing plans and procedures for the evaluation project.

- *Execution of the evaluation study* accomplishing the designed evaluation study.
- *Completion of the evaluation study* reporting, accounting, archiving of evaluation study results, finalization of outstanding issues and formal closure of the evaluation study.

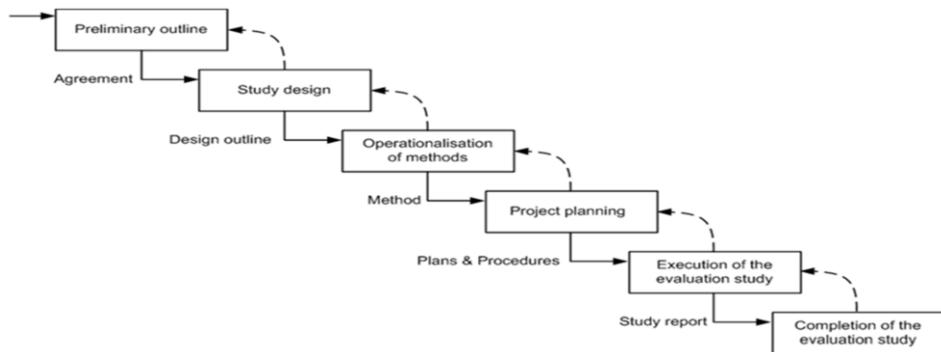


Figure 1: Phases of a health IT evaluation according to the GEP-HI guideline [7].

To progress from one phase to the next phase a formal acceptance is required from the relevant stakeholders of the planned evaluation study. For each phase a list of items is presented (Table 1) and these should be carefully considered during the study. All phases together contain some 60 detailed items, which are presented in relation to the evaluation study phases. When designers and executors of evaluation studies address these items, the plan, structure, objectives and results of the studies will become more robust and consequently the studies contribute an important step towards evidence-based health informatics.

Table 1. Items to be considered at each phase of a health IT evaluation according to GEP-HI.

Phase no.	Phase	Items of the phase
1	Preliminary outline	<ul style="list-style-type: none"> - Purpose of the study - Primary audience - Identification of the study funding party(ies) - First identification of stakeholders - Identification of required expertise - The organizational and user context of the evaluation study - Object of evaluation, type of health IT - First exploration of evaluation methods to be used - Ethical and legal issues - Budget - Preliminary permissions for publication - Result of preliminary outline - Formal acceptance to proceed to the next phase

2	Study design	<ul style="list-style-type: none"> - Detailed rationale and objectives for the study - Key evaluation issues, questions, indicators - Stakeholder analysis/Social Network analysis - Study methods - Organizational context, the study setting - Technical setting, the type of health IT - Participants from the organization - Project timeline - Material and practical resources - Establishment of the study team - Risk analysis and quality management - Budget - Ethical and legal issues - Strategy for reporting and disseminating the results - Result of study design - Formal acceptance to proceed to the next phase
3	Operationalization of methods	<ul style="list-style-type: none"> - Study type - Approach - Assumptions and feasibility assessment - Frame of reference - Timing - Justification of the methodological approach - Expertise - Outcome measures - Avoiding Bias - Quality control on data (measures) - Participants - Ethical and legal issues - Result of operationalization of methods - Approval of operationalization of methods
4	Project planning	<ul style="list-style-type: none"> - Project management - Study flow - Evaluation activity mapping - Quality management - Risk management - Recruitment of necessary staff - Inform all relevant stakeholders - Result of project planning - Approval of project planning
5	Execution of the evaluation study	<ul style="list-style-type: none"> - Undertake the study, collect data and interpret observations - Quality control of findings and observation of changes - Continuous project management, quality and risk management - Regular reports - Final result of execution of the evaluation study
6	Completion of the evaluation study	<ul style="list-style-type: none"> - Accounting - Archiving - Reports and publications

The strength of the GEP-HI guideline is in forcing the user to go through a checklist of relevant issues that might otherwise only act informally as tacit knowledge, or even be overlooked. This systematic approach will increase the likelihood of an evaluation outcome with the desired level of accuracy and precision and hence an increased effectiveness, and additionally encourage the adoption of a scientifically

valid approach in an evaluation study. The performed evaluation study should be reported following the STARE-HI reporting statement [38].¹¹

4. Discussion and conclusions

The scope of evaluation studies is wide. There are small snapshot-type studies just to get an insight on the ongoing development, or longitudinal wide studies following the health IT system for a long time with various users and use contexts, or impact studies to assess the changes that have been implemented by the health IT system and possibly also measure the effects of the system on efficiency and outcomes of the health care organization. All these require different planning, management and methods, different resources and expertise.

There are quality problems in reported evaluation studies, e.g. weak planning, missing systematic, scientific methodology for evaluation, false assumptions and conclusions, experimental errors and weak attention to intra- or inter-organizational variability. Quality problems in usability evaluations are related e.g. to the narrow focus of the study, or evaluation of the use in a very specific context, or leaving out the relationship between single-system development and the existing technology setting in healthcare or the characteristics of various use contexts.

In health care clinical context, the environment consists of many health IT applications, of which several are used simultaneously. Evaluation of these health IT environments should address the relevant evaluation criteria from a broad viewpoint. Health IT systems need to be seen as integrated parts of a wider technology environment and the objectives of evaluation should be framed with respect to the clinical situation and use contexts. A challenge is to cover the wide variety of users of the health IT systems and the numerous purposes these systems serve, and the diversity of clinical surroundings in healthcare organizations where the systems are implemented and used.

The GEP-HI guideline has been developed to overcome the identified quality problems in evaluation study design and execution. GEP-HI can be applied to different kinds of health IT evaluation studies, irrespective of whether the object of study is an IT application or a method like nursing classification or data security practice. In small evaluation studies not all phases of the GEP-HI guideline may be needed. The guideline is applicable at various phases of a health IT project, starting from design and development, over application or system implementation and installation, and ending with the study of effects and impacts in routine use.

The GEP-HI guideline and other discussed health IT evaluation guidelines and frameworks support detailed planning of the evaluation study and thus help to consider carefully the study objectives and operationalisation of methods for the specific evaluation study in planning; most of them are applicable to different types of studies, such as feasibility, effectiveness, efficiency and impact evaluation, and to studies with various scopes. GEP-HI supports application of multi-method approach in evaluation and integration of usability evaluation with the other evaluation criteria e.g. interoperability, security, effects and impacts evaluation.

¹¹ See also: P. Doupi, Evolving Health IT Systems Evaluation: The Convergence of Health Informatics and HTA, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

Rigby et al. [40] listed the ten core principles that are essential for the effectiveness in all evaluations: Preliminary planning, stakeholder analysis, health issue and eHealth application, safety, clinical effectiveness, user experience, economic aspects, organisational aspects, ethical and legal issues and reporting of the studies. These are well covered by the GEP-HI guideline. For the future, it is important to define the quality aspects of interest in an evaluation study and to plan the study carefully and apply approaches that enable integration of various evaluation criteria in the specific evaluation study. This is needed to get a complete picture of the health IT system under study, from the users', developers' and from health care organisational perspectives. Systematic guidelines help to pay attention to all relevant issues and to plan carefully the evaluation study and execute the study following scientific principles of research.

Recommended further readings

1. E. Ammenwerth, J. Brender, P. Nykänen, H.U. Prokosch, M. Rigby, J. Talmon, Visions and strategies to improve evaluation of health information system. Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, *Int J Med Inform* **73** (2004), 479-91.
2. E. Ammenwerth and N. de Keizer, An inventory of evaluation studies of information technology in health care. Trends in evaluation research 1982 – 2002, *Meth Inf Med* **1** (2005), 44-56.
3. C.F. Friedman, J. Wyatt, *Evaluation Methods in Medical Informatics*, 2nd edition, Springer, New York, 2006.
4. M. Rigby, E. Ammenwerth, M-C. Beuscart-Zephir, J. Brender, H. Hyppönen, S. Melia, P. Nykänen, J. Talmon, N. de Keizer, Evidence based health informatics: 10 years of efforts to promote the principle, in: *IMIA Yearbook of Medical Informatics 2013*, Schattauer, Stuttgart, 2013. pp. 1-13.

Food for thought

1. What are the challenges and possibilities to integrate various evaluation frameworks or guidelines in one evaluation study? Think e.g. integration of GEP-HI and MAST or integration of GEP-HI and HOT-fit?
2. How to integrate usability evaluation with standard evaluation framework, e.g. within the GEP-HI guideline?
3. Analyse the differences between a small-scale and a wide-scale evaluation study planning and execution. Differences may be related to e.g. objectives of the study, scope of the study, evaluation criteria, methods applied, reporting of the study. As examples of various study scopes you may consider e.g. evaluation of the user acceptance of one health IT system in one hospital department (small-scale) and evaluation of the impacts of an electronic health record system (EHR) at many levels of the health care system (primary care, specialized care, tertiary care).

References

- [1] J. Brender, *Handbook of evaluation methods for health informatics*, Elsevier Academic Press, USA, 2006.
- [2] E. Ammenwerth, N. Shaw, Bad health informatics can kill – Is evaluation the answer, *Editorial. Methods Inf Med* **44** (2005), 1-3.
- [3] A.G. Ekeland, A. Bowes, S. Flottorp, Effectiveness of telemedicine: A systematic review of reviews, *Int J Med Inform* **79** (2010), 736-771.
- [4] E. Ammenwerth, J. Brender, P. Nykänen, H.U. Prokosch, M. Rigby, J. Talmon, Visions and strategies to improve evaluation of health information systems. Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, *Int J Med Inform* **73** (2004), 479-91.
- [5] *Declaration of Innsbruck – Results from the European Science Foundation sponsored Workshop on Systematic Evaluation of Health Information Systems (HIS-EVAL)*, April 4-6th, 2003. Published in: Ammenwerth E, Brender J, Nykanen P, Prokosch HU, Rigby M, Talmon J, et al., Visions and strategies to improve evaluation of health information systems; Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, *Int J Med Inf* **73**(6) (2004), 479-91.
- [6] M. Rigby, *Evaluation: 16 Powerful Reasons Why Not to Do It - And 6 Over-Riding Imperatives*, in: V. Patel, R. Rogers, R. Haux (eds.), *Medinfo 2001: Proceedings of the 10th. World Congress on Medical Informatics*, IOS Press, Amsterdam, 2001, pp. 1198-1202.
- [7] P. Nykänen, J. Brender, J. Talmon, N. de Keizer, M. Rigby, M-C. Beuscart-Zephir, E. Ammenwerth, Guideline for good evaluation practice in health informatics, *Int J Med Inform* **80** (2011), 815-827.
- [8] E.M.S.J. van Gennip, J.T. Talmon (eds.), *Assessment and evaluation of information technologies*, IOS Press, Amsterdam, 1995.
- [9] B. Kaplan, N. Shaw, Future directions in evaluations research: People, organisational and social issues, *Methods Inf Med* **43** (2004), 215-231.
- [10] R. van der Loo, *Overview of Published Assessment and Evaluation Studies*, in: E.M.S.J. van Gennip, J.T. Talmon (eds.), *Assessment and evaluation of information technologies*, IOS Press, Amsterdam, 1995, pp. 261-282.
- [11] E. Ammenwerth, N. de Keizer, An inventory of evaluation studies of information technology in health care, Trends in evaluation research 1982 – 2002, *Meth Inf Med* **1** (2005), 44-56.
- [12] M. Rigby, *Evaluation – the Cinderella science of ICT in health*, in: C. Kulikowski, R. Haux (eds.), *IMIA Yearbook of Medical Informatics 2006*, Schattauer, Stuttgart, 2006, pp. 114-120.
- [13] J. Talmon, *Evaluation and implementation: A call for action*, in: C. Kulikowski, R. Haux (eds.), *IMIA Yearbook of Medical Informatics 2006*, Schattauer, Stuttgart, 2006, pp. 11-15.
- [14] J. Zhang, Human-Centred Computing in Health Information Systems, Part 1: Analysis and Design, *Journal of Biomedical Informatics* **38** (2005), 1-3.
- [15] B. Chaudhry, J. Wang, S.Wu, M. Maglione, W. Mojica, E. Roth, S.C. Morton, P.G. Shekelle, Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care, *Annals of Internal Medicine* **144** (2006), 742-752.
- [16] J. Kaipio, *Usability in healthcare: Overcoming the mismatch between information systems and clinical work*, Aalto University publication series, Doctoral dissertations 105, Unigrafia Oy, Helsinki, 2011, <http://lib.tkk.fi/Diss/2011/isbn9789526043340/isbn9789526043340.pdf>, last access 11 February 2016.
- [17] J. Nielsen, *Usability Engineering*, Academic Press, San Diego, USA, 1993.
- [18] ISO 9241-11, *Ergonomic Requirements for Office Work with Visual Display Terminals, part 11: Guidance on Usability*, International Organization for Standardization, Geneve, 1998.
- [19] H. Beyer, K. Holtzblatt, *Contextual Design: Defining Customer-Centered Systems*, Academic Press, San Diego, USA, 1998.
- [20] M-C. Beuscart-Zephir, P. Elkin, S. Pelayo, R. Beuscart, *The human factors engineering approach to biomedical informatics projects: State of the art, results, benefits and challenges*, in: A Geissbuhler, R Haux, C Kulikowski (eds.), *IMIA Yearbook of Medical Informatics 2007*, Schattauer, Stuttgart, 2007, pp. 109-127.
- [21] A. W. Kushniruk, Evaluation in the Design of Health Information Systems: Application of Approaches Emerging from Usability Engineering, *Computers in Biology and Medicine* **32** (2001), 141-149.
- [22] A.W. Kushniruk, V.L. Patel, Cognitive and usability engineering methods for the evaluation of clinical information systems, *J Biomed Inform* **37**(2004), 56-76.
- [23] M.W.M. Jaspers, A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence, *Int J Med Inform* **78** (2009), 340-353.
- [24] P.J. Edwards, K.P. Moloney, J.A. Jacko, F. Sainfort, Evaluating usability of a commercial electronic health record: a case study, *Int J Hum Comput Stud* **66** (2008), 718-728.
- [25] E. Ammenwerth, S. Gräber, G. Herrmann, T. Bürkle, J. König, Evaluation of Health Information Systems – Problems and Challenges, *Int J Med Inform* **71** (2003), 125 –135.

- [26] B. Kaplan, Addressing organizational issues into the evaluation of medical systems, *J Am Med Inform Assoc* **4** (1997), 94-101.
- [27] N.T. Shaw, CHEATS: A generic evaluation framework for information communication technologies, *Comp Biol Med* **32** (2002), 209-220.
- [28] K. Kidholm, A. Bowes, S. Dyrehauge, A. Granström Ekeland, S.A. Flottorp, L. Kvistgaard et al., *The MAST Manual - MAST-Model for assessment of telemedicine*, <http://www.epractice.eu/en/library/315610>, last access 11 February 2016.
- [29] C.M. Cusack, C. Byrne, J.M. Hook, J. McGowan, E.G. Poon, A. Zafar, *Health information technology evaluation toolkit: 2009 Update*, AHRQ Publication No. 09-0083-EF, Agency for Healthcare Research and Quality, <http://www.ahrq.gov>, last access 11 February 2016.
- [30] K. Clarke, R. O'Moore, R. Smeets, J. Talmon, J. Brender, P. McNair et al., A methodology for evaluation of knowledge-based systems in medicine, *Artificial Intelligence in Medicine* **6** (1994), 107-121.
- [31] M.M. Yusof, J. Kuljis, A. Papazafeiropoulou, K. Stergioulas, An evaluation framework for health information systems: Human, organization and technology-fit factors (HOT-fit), *Int J Med Inform* **77** (2008), 386-398.
- [32] ISO 9241-210 (2010) *International Standard: Ergonomics of Human-system Interaction – Part 210: Human-centred Design for Interactive Systems*, First edition 2010-03-15, Reference number ISO 9241-210:2010(E).
- [33] M.M. Yusof, A. Papazafeiropoulou, R.J. Paul, K. Stergioulas, Investigating evaluation frameworks for health information systems, *Int J Med Inform* **77** (2008), 377-385.
- [34] M. Berg, Medical work and the computer-based patient record: a sociological perspective, *Meth Inf Med* **7** (1998), 294-301.
- [35] P.J. Scott, J. Briggs, STAT-HI: A socio-technical assessment tool for health informatics implementations, *Open Med Inform J* **4** (2010), 214-20.
- [36] WHO, *Call to Action on Global eHealth Evaluation - Consensus Statement of the WHO Global eHealth Evaluation Meeting*, Bellagio, http://www.ghdonline.org/uploads/The_Bellagio_eHealth_Evaluation_Call_to_Action-Release.docx, last access 11 February 2016.
- [37] C.F. Friedman, J. Wyatt, *Evaluation Methods in Medical Informatics*, 2nd edition, Springer, New York, 2006.
- [38] J. Talmon, E. Ammenwerth, J. Brender, N. de Keizer, P. Nykänen, M. Rigby, STARE-HI – Statement on Reporting of Evaluation Studies in Health Informatics, *Int J Med Inform* **78** (2009), 1- 9.
- [39] P. Nykänen, J. Enning, J. Talmon, D. Hoyer, F. Sanz, C. Thayer, R. Roine, M. Vissers and F. Eurlings, Inventory of validation approaches in selected health telematics projects, *Int J Med Inform* **56** (1999), 87-96.
- [40] M. Rigby, E. Ammenwerth, M.-C. Beuscart-Zephir, J. Brender, H. Hyppönen, S. Melia, P. Nykänen, J. Talmon, N. de Keizer, *Evidence based health informatics: 10 years of efforts to promote the principle*, in: IMIA Yearbook of Medical Informatics 2013, Schattauer, Stuttgart, 2013. pp. 1-13.

Publishing Health IT Evaluation Studies

Elske AMMENWERTH^{a,1} and Nicolet F. de KEIZER^b

^aUMIT – University for Health Sciences, Medical Informatics, Hall in Tirol, Austria

^bAcademic Medical Center, Amsterdam, Netherlands

Abstract. Progress in science is based on evidence from well-designed studies. However, publication quality of health IT evaluation studies is often low, making exploitation of published evidence within systematic reviews and meta-analysis a challenging task. Consequently, reporting guidelines have been published and recommended to be used. After a short overview of publication guidelines relevant for health IT evaluation studies (such as CONSORT and PRISMA), the STARE-HI guidelines for publishing health IT evaluation studies are presented. Health IT evaluation publications should take into account published guidelines, to improve the quality of published evidence. Publication guidelines, in line with addressing publication bias and low study quality, help strengthening the evidence available in the public domain to enable effective evidence-based health informatics.

Keywords. Medical informatics, publishing, evaluation studies, guideline .

1. Introduction

Progress in science is based on evidence from well-designed studies, normally in individual peer-reviewed publications, and also sometimes in repositories of studies. This evidence is often collected and aggregated in the form of systematic literature reviews and meta-analyses. A systematic literature review typically involves a detailed and comprehensive plan and search strategy derived *a priori*. The goal is to add strength and reduce selection bias by identifying, appraising, and synthesizing all relevant studies on a particular topic. A meta-analysis, in addition, comprises statistical method to synthesize the data from several studies into a single quantitative estimate or summary effect size [1].

However, while preparing systematic reviews and meta-analyses on health IT evaluation studies, reviewers have been confronted with three major challenges leading to possible bias and low quality of published evidence: Publication bias, summarizing the problem that studies with unfavourable outcome may not be published due to stakeholder pressure or related political reasons [2];² low quality of the conducted evaluation study;³ and poor reporting quality of the published evaluation study, where often important information needed to understand, interpret, reproduce or generalize

¹ Corresponding author: Prof. Dr. Elske Ammenwerth, Institute for Biomedical Informatics, Eduard Wallnöfer Zentrum 1, 6060 Hall in Tirol, Austria, elske.ammenwerth@umit.at.

² See also: A. Georgiou, Finding, appraising and interpreting the evidence, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

³ See also: P. Nykänen et al., Quality of health IT evaluations, in: *ibid.*

the findings of a study is missing in a study paper. This contribution will address this last challenge: low reporting quality.

2. Reporting quality of health IT evaluation studies

The interpretation of studies of health IT is highly context-specific [3]. Thus, especially information on the study context is important to allow the reader to judge the generalizability or relevance to their setting of the published study. Missing context information endangers the evidence-base of health informatics [4]. As Shekelle writes: “The generalizability of evidence will remain low unless more systematic, comprehensive, and relevant descriptions and measurements are made regarding how the technology is utilized, the individuals using it, and the environment it is used in” [3].

Nevertheless, health IT evaluation studies often show insufficient reporting quality. For example, while reviewing 23 randomized health IT trials, Jamal et al. found an insufficient description of the health IT intervention, of allocation or randomization procedures, or of data collection procedures [5]. Likewise, while reviewing 257 health IT studies, Chaudhry et al. found insufficient description of the health IT intervention, the implementation process and the organizational context while reviewing health IT evaluation studies [6]. Eisenstein et al. analysed 134 economic health IT evaluations and found that many studies did not report on key information such as invested financial and personal resources or cost elements. Talmon et al. analyzed the reporting quality of 47 health IT trials and found that title and abstract often missed important information such as the type of evaluated health IT [7]. Shekelle et al. analysed 258 health IT evaluation studies and found that only very few studies reported sufficient information on the organizational and technical context, including health IT usage and users [3].

In a study specifically analysing publication quality, de Keizer et al. reviewed the quality of 120 randomly chosen health IT evaluation studies [8]. They found varying degrees of reporting quality. Often, the evaluated health IT intervention (including functionality, usage, and workflow), the involved study population, and methods or instruments for data collection or data analysis were not described in sufficient detail. Also, no improvement in reporting quality was visible between 1980 and 2005.

Consequently, several reviewers expressed the strong need to improve reporting quality and to develop reporting standards for publication of health IT evaluation studies [5,6,8]. In medical science, guidelines to improve publication quality such as CONSORT [9] or PRISMA [10] have existed for many years. While these guidelines may be helpful on a general basis, they do not cover specific aspects of health IT evaluation studies. Therefore, in 2009, STARE-HI was proposed as a specific guideline for health IT evaluation papers [11]. In 2011, in addition, CONSORT-eHealth [12] for specific types of health IT evaluation studies was published.

In this contribution, we will first present and discuss the applicability of the guidelines from medical sciences such as CONSORT. We will then present the motivation and details of the STARE-HI guideline in more detail.

3. General publication guidelines

This problem of insufficient publication quality is well-known in the medical sciences and several publication guidelines have been developed in the last few years for several clinical study types.

Due to the rising number of publication guidelines in the medical sciences, the EQUATOR network was launched in 2006 [13]. On its website, EQUATOR collects available guidelines and makes them easily accessible. As of June 2015, the website already contained 276 publication guidelines. Many of these guidelines are also of relevance for health IT evaluation publications.

The publication guidelines included in the EQUATOR network have different adoption rates in the scientific community. Some of them are very well known and frequently used. Some of them have even been adopted by major medical journals; submitting authors have to indicate which guideline applies to their submission, and how they follow this guideline. Some of these broadly adopted guidelines include CONSORT for reporting of randomized controlled trials [14], STARD for reporting of diagnostic studies [15], STROBE for reporting of observational studies [16], and PRISMA for systematic reviews [10]. We will give a summary of these guidelines in this section, and discuss their applicability for health IT evaluation studies.

3.1. CONSORT

CONSORT (Consolidated Standards of Reporting Trials) addresses the problems arising from inadequate reporting of randomized controlled trials. The CONSORT Statement is a minimum set of recommendations for reporting randomized trials [9]. The CONSORT 2010 checklist includes 25 items that have to be included in a report of a randomized trial, including information on objectives, design, participants, outcomes, blinding, patient flow, harms, and limitations. A detailed explanation and elaboration paper is available [17]. Several adaptations of the CONSORT statements for specific situations have been published, e.g. for reporting of cluster randomized trials [18] or for reporting of patient-reported outcomes [19]. CONSORT has been endorsed by more than 600 biomedical journals [20]. More details on CONSORT are available at <http://www.consort-statement.org>.

Health IT evaluation studies which use a randomized controlled trial design are recommended to use CONSORT when reporting their results. In addition, a precise description of the health IT system under evaluation and the context in which the intervention is implemented should be provided.

3.2. STARD

The objective of STARD (Standards for Reporting of Diagnostic Accuracy) is to improve the completeness and transparency of reporting of studies of diagnostic accuracy and to allow assessing internal and external validity [15]. The STARD checklist comprises 25 items, including participant recruitment, data collection, study population, estimates of diagnostic accuracy, adverse events, and discussion of clinical applicability. An explanation and elaboration paper is available [21]. STARD has been endorsed by more than 200 biomedical journals [22]. More details on STARD are available at <http://www.stard-statement.org>.

Health IT evaluation studies which focus on diagnostic accuracy, e.g. accuracy of teledermatology systems or clinical decision support systems, are recommended to use STARD when reporting their results.

3.3. STROBE

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement supports the dissemination of observational studies [16]. Observational studies comprise, for example, cohort studies, case-control studies, and cross-sectional studies. The STROBE checklist comprises 22 items, including objectives, study design, setting, participants, variables, statistical methods, outcome data, and key results. An explanation and elaboration paper is available [23]. STROBE has been endorsed by around 200 biomedical journals [24]. More details on STROBE are available at <http://www.strobe-statement.org>.

Many health IT evaluation studies have an observational nature, monitoring for example the effect of an health IT system in a before-after study or in a time series study. These studies are recommended to use STROBE when reporting their results.

3.4. PRISMA

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement is a minimum set of items for reporting of systematic reviews and meta-analyses [10]. The PRISMA checklist comprises 27 items, including objectives, eligibility criteria, search, study selection, data collection, risk of bias, synthesis of results, summary of evidence, and limitations. An explanation and elaboration paper is available [25]. PRISMA has been endorsed by more around 200 biomedical journals [26]. More details on PRISMA are available at <http://www.prisma-statement.org>.

All systematic reviews and meta-analyses on health IT topics are recommended to use PRISMA when reporting their results.

4. Publication guidelines for health IT evaluation studies

We will now look at two publication guidelines specifically developed for health IT evaluation studies.

4.1. STARE-HI

The Statement on Reporting of Evaluation Studies in Health Informatics (STARE-HI) addresses writing and assessing evaluation reports in Health Informatics. Its goal is to improve the quality of published evaluation studies in Health Informatics, and thus to improve the evidence-base of Health Informatics [11]. The STARE-HI checklist comprises 30 items, including objective of the study, organizational setting, system details and system in use, study design, study flow, outcome measures, unexpected observations, and meaning and generalizability of results. An explanation and elaboration paper is available [27] as well as a shortened version for conference paper [28].

STARE-HI has been endorsed by major health informatics journals as well as by the International Medical Informatics Association (IMIA) and the European Federation for Medical Informatics (EFMI) [24]. Furthermore, STARE-HI has been included in the EQUATOR network [13]. More details on STARE-HI are available at <http://iig.umit.at/efmi/starehi.htm>. Table 1 presents the content of STARE-HI in more detail.

Table 1. The STARE-HI principles: Items recommended to be included in health informatics evaluation reports [11].

1 Title
2 Abstract
3 Keywords
4 Introduction
4.1 Scientific background
4.2 Rationale for the study
4.3 Objectives of study
5 Study context
5.1 Organizational setting
5.2 System details and system in use
6 Methods
6.1 Study design
6.2 Theoretical background
6.3 Participants
6.4 Study flow
6.5 Outcome measures or evaluation criteria
6.6 Methods for data acquisition and measurement
6.7 Methods for data analysis
7 Results
7.1 Demographic and other study coverage data
7.2 Unexpected events during the study
7.3 Study findings and outcome data
7.4 Unexpected observations
8 Discussion
8.1 Answers to study questions
8.2 Strengths and weaknesses of the study
8.3 Results in relation to other studies
8.4 Meaning and generalisability of the study
8.5 Unanswered and new questions
9 Conclusion
10 Authors' contribution
11 Competing interests
12 Acknowledgement
13 References
14 Appendices

4.2. CONSORT-eHealth

CONSORT-eHealth aims at improving and standardizing evaluation reports of web-based and mobile health interventions. The authors argue that “RCTs of web-based interventions pose very specific issues and challenges, in particular related to reporting sufficient details of the intervention” [12], and therefore they developed CONSORT-eHealth based on CONSORT. CONSORT-eHealth comprises 53 additional sub-items explaining or enhancing the original CONSORT items, such as type of system, bug fixes and down items, computer literacy of participants, names of sponsors, revisions and updating, level of human involvement, intensity of use, and safety and security

procedures. CONSORT-eHealth has been endorsed by the Journal of Medical Internet Research [12].

5. Discussion and Conclusion

Publication of health IT evaluation studies needs to comprise sufficient information to be understandable and generalizable. Low publication quality as found in many reviews impedes evidence-based health informatics, devalues the work that has been done, reduces the value to the reader eager to learn and apply the knowledge, and compromises the potential value of subsequent wider systematic reviews and meta-analyses.⁴

Medical science has a long history of publication guidelines, the major ones being endorsed by many larger biomedical journals. However, there are specific issues of health IT evaluation studies that are often insufficiently reported, such as details of the health IT intervention; IT related characteristics of the system users; and the technical and organizational setting. STARE-HI and other specific guidelines attempt to address this by offering specific guidance for publishing health IT evaluation studies. They do not replace, but complement, other established guidelines. Health IT evaluation publication should apply these other relevant guidelines where appropriate – for example, CONSORT for randomized trials. The adoption of STARE-HI by larger health informatics journals as well as by international health informatics organizations stresses the importance of ensuring the quality of publication, and thus in turn the evidence base, for health informatics and its applications.

Publication guidelines are one of the three means of improving publication of studies and thus strengthening evidence available in the public domain to enable effective evidence based health informatics (EBHI). Publication guidelines are the easiest of the three to implement, as they are focussed on the reporting of studies which have been undertaken. Improving the quality of studies also requires use of other guidelines which address planning and conduct of studies, such as GEP-HI, the Guidelines for Good Evaluation Practice in Health Informatics.⁵ Publishing all studies clearly and effectively, and not just positive ones, faces a number of challenges [29], and is a problem also being faced in the clinical and pharmaceutical domains, but is a moral duty and should be addressed by all who believe in effective health IT support to health care.

Recommended further readings

1. J. Talmon, A. Ammenwerth, J. Brender, N. de Keizer, P. Nykänen, M. Rigby, STARE-HI - Statement on Reporting of Evaluation Studies in Health Informatics, *Int J Med Inform* **78**(1) (2009), 1-9.

⁴ See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁵ See also: P. Nykänen et al., Quality of health IT evaluations, in: *ibid.*

Food for thought

1. Do you routinely assess the quality of the publication of studies which you use either in evidence for decision-making, or as a basis for designing further studies?
2. Imagine different situations or use contexts where you are reading or writing a paper. Is it possible that in these different situations, the importance of items to be covered in a publication may differ?
3. How could you access whether the quality of publications increased after a journal endorsed a certain guideline? Look for evidence on this question!

References

- [1] L.S. Uman, Systematic reviews and meta-analyses, *J Can Acad Child Adolesc Psychiatry* **20**(1) (2011), 57-9.
- [2] C. Friedman, J. Wyatt, Publication bias in Medical Informatics, *J Am Med Inform Assoc* **8**(2) (2001), 189-191.
- [3] P. Shekelle, S. Morton, M. Keeler, Costs and Benefits of Health Information Technology, *Evid Rep Technol Assess (Full Rep)* **132** (2006), 1-71.
- [4] E. Ammenwerth, P. Schnell-Inderst, U. Siebert, Vision and challenges of Evidence-Based Health Informatics: a case study of a CPOE meta-analysis, *Int J Med Inform* **79**(4) (2010), e83-8.
- [5] A. Jamal, K. McKenzie, M. Clark, The impact of health information technology on the quality of medical and health care: a systematic review, *Him J* **38**(3) (2009), 26-37.
- [6] B. Chaudhry, J. Wang, S. Wu, M. Maglione, W. Mojica, E. Roth, S.C. Morton, P.G. Shekelle, Systematic review: impact of health information technology on quality, efficiency, and costs of medical care, *Ann Intern Med* **144**(10) (2006), 742-52.
- [7] J. Talmon, E. Ammenwerth, T. Geven, The quality of reporting of health informatics evaluation studies: a pilot study, *Stud Health Technol Inform* **129**(Pt 1) (2007), 193-7.
- [8] N.F. de Keizer, E. Ammenwerth, The quality of evidence in health informatics: how did the quality of healthcare IT evaluation publications develop from 1982 to 2005? *Int J Med Inform* **77**(1) (2008), 41-9.
- [9] G. Antes, The new CONSORT statement, *BMJ* **340** (2010), c1432.
- [10] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, P. Group, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, *Int J Surg* **8**(5) (2010), 336-41.
- [11] J. Talmon, A. Ammenwerth, J. Brender, N. de Keizer, P. Nykänen, M. Rigby, STARE-HI - Statement on Reporting of Evaluation Studies in Health Informatics, *Int J Med Inform* **78**(1) (2009), 1-9.
- [12] G. Eysenbach, CONSORT-eHealth: improving and standardizing evaluation reports of Web-based and mobile health interventions, *J Med Internet Res* **13**(4) (2011), e126.
- [13] EQUATOR, *The Equator Network: Enhancing the Quality and Transparency of Health Research*, <http://www.equator-network.org>, last access 11 February 2016.
- [14] K.F. Schulz, D.G. Altman, D. Moher, C. Group, CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials, *BMJ* **340** (2010), c332.
- [15] P.M. Bossuyt, J. Reitsma, D. Bruns, C. Gatsonis, P. Glasziou, L. Irwig, J. Lijmer, D. Moher, D. Rennie, H. de Vet, Towards Complete and Accurate Reporting of Studies of Diagnostic Accuracy: The STARD Initiative, *Ann Int Med* **138**(1) (2003), 40-44.
- [16] E. von Elm, D.G. Altman, M. Egger, S.J. Pocock, P.C. Gotzsche, J.P. Vandebroucke, S. Initiative, The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies, *Lancet* **370**(9596) (2007), 1453-7.
- [17] D. Moher, S. Hopewell, K.F. Schulz, V. Montori, P.C. Gotzsche, P.J. Devereaux, D. Elbourne, M. Egger, D.G. Altman, CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials, *BMJ* **340** (2010), c869.
- [18] M.K. Campbell, G. Piaggio, D.R. Elbourne, D.G. Altman, Consort 2010 statement: extension to cluster randomised trials, *BMJ* **345** (2012), e5661.
- [19] M. Calvert, J. Blazeby, D.G. Altman, D.A. Revicki, D. Moher, M.D. Brundage, Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension, *Jama* **309**(8) (2013), 814-22.
- [20] Consort, *Impact of CONSORT*, <http://www.consort-statement.org/about-consort/impact-of-consort>, last access 11 February 2016.

- [21] P.M. Bossuyt, J.B. Reitsma, D.E. Bruns, C.A. Gatsonis, P.P. Glasziou, L.M. Irwig, D. Moher, D. Rennie, H.C. de Vet, J.G. Lijmer, The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. The Standards for Reporting of Diagnostic Accuracy Group, *Croat Med J* **44**(5) (2003), 639-50.
- [22] STARD, *News about STARD*, <http://www.stard-statement.org>, last access 11 February 2016.
- [23] J.P. Vandebroucke, E. von Elm, D.G. Altman, P.C. Gotzsche, C.D. Mulrow, S.J. Pocock, C. Poole, J.J. Schlesselman, M. Egger, S. Initiative, Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration, *PLoS Med* **4**(10) (2007), e297.
- [24] STROBE, *STROBE Supporting journals & organizations*, <http://www.strobe-statement.org/index.php?id=strobe-endorsement>, last access 11 February 2016.
- [25] A. Liberati, D.G. Altman, J. Tetzlaff, C. Mulrow, P.C. Gotzsche, J.P. Ioannidis, M. Clarke, P.J. Devereaux, J. Kleijnen, D. Moher, The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration, *BMJ* **339** (2009), b2700.
- [26] PRISMA, *PRISMA Endorsers*, <http://www.prisma-statement.org/Endorsement/EndorsePRISMA.aspx>, last access 11 February 2016.
- [27] J. Brender, J. Talmon, N. de Keizer, P. Nykanen, M. Rigby, E. Ammenwerth, STARE-HI - Statement on Reporting of Evaluation Studies in Health Informatics: explanation and elaboration, *Appl Clin Inform* **4**(3) (2013), 331-58.
- [28] N.F. de Keizer, J. Talmon, E. Ammenwerth, J. Brender, M. Rigby, P. Nykanen, Systematic Prioritization of the STARE-HI Reporting Items. An Application to Short Conference Papers on Health Informatics Evaluation, *Methods Inf Med* **51**(2) (2012), 104-11.
- [29] M. Rigby, *Evaluation: 16 Powerful Reasons Why Not to Do It - And 6 Over-Riding Imperatives*, in: V. Patel, R. Rogers, R. Haux, editors. Proceedings of the 10th World Congress on Medical Informatics (Medinfo 2001), IOS Press, Amsterdam, 2001. pp. 1198-202.

Finding, Appraising and Interpreting the Evidence of Health IT

Andrew GEORGIOU^{a,1}

^a*Centre for Health Systems & Safety Research, Australian Institute of Health Innovation, Macquarie University, Sydney, Australia*

Abstract. Evidence-based health informatics can be described as a scientific approach to meeting the multiplicity of tasks involved in the development, implementation and sustainability of health information technologies (IT). The practice of evidence-based health informatics incorporates methods to help find, appraise and utilise research-based knowledge. The aim of this contribution is to describe the steps of finding, appraising and interpreting the evidence of health IT. It lists major sources of literature in the health field, and highlights a number of considerations for undertaking reviews, drawing on some key landmark reviews that have helped to shape the health informatics discipline. It also considers key issues highlighted by these reviews particularly in regard to the validity of findings, their generalisability and their impact on patient outcomes. The contribution also provides suggestions for tackling the challenge of potential publication bias, and how to deal with heterogeneous findings.

Keywords. Evidence-Based Medicine; Evaluation Studies; Information Systems; Medical Informatics; Research Methodology.

1. Introduction

Evidence-based health informatics can be described as a scientific approach to meeting the multiplicity of tasks involved in the development, implementation and sustainability of health information technologies (IT) [1]. The practice of evidence-based health informatics incorporates methods to help find, appraise and utilise research-based knowledge by using the most current literature sources, ranging from large established databases through to government reviews, consultancy reports and industry appraisals. This can encompass a variety of different review methods including: *personal searches* of the literature; traditional *literature reviews*, which provide an overview of the literature without necessarily focusing on outcomes; *scoping reviews* that assess the scope of existing literature; *expert reviews* where an expert or group of experts makes a judgement call on an issue related to the literature; *realist reviews* which concentrate on generating generalisable theories, and *meta-systematic reviews* involving reviews of systematic reviews.

In and of themselves, all forms of literature reviews can be conceived of as a means to make sense of a (usually large) body of literature. The distinguishing feature

¹ Corresponding author: Associate Professor Andrew Georgiou, Centre for Health Systems and Safety Research, Australian Institute of Health Innovation, Macquarie University, NSW 2109 Australia. andrew.georgiou@mq.edu.au.

of a systematic review, however, lies in its *comprehensiveness* and its adoption of formal approaches and transparency leading to its ability to be reproduced [2].

Systematic reviews are thus required to comply with rigorous methods as a means of identifying, appraising and synthesising all the studies relevant to the research question(s) at hand [3]. Generally, this means providing answers about the effectiveness of a healthcare intervention. Systematic reviews can be adopted for a variety of other reasons, including the investigation of the feasibility of an intervention, its appropriateness or even to identify evidence about consumer experiences. Systematic reviews have been used to examine many aspects of the impact of health IT on different areas of the clinical care process (e.g., medications, incidents, preventative care) employing a range of metrics such as care quality, provider productivity, user satisfaction and information quality [4].

The aim of this contribution is to describe the steps of finding, appraising and interpreting the evidence of health IT, and the context in which they may be employed. These steps include:

- Formulating the study question.
- Developing a search strategy.
- Finding the evidence.
- Tackling publication bias.
- Organising the evidence.
- Critically appraising the evidence.
- Being sensitive and aware of context.
- Interpreting the evidence.

2. Formulating the study question

The first thing that a clinician, policy maker, health informatician or researcher needs to consider before undertaking a review is what type of review is required. Systematic reviews of randomised controlled trials are considered to be the highest level of evidence (Level I) [5]², but this does not mean a systematic review is always the most suitable instrument for the topic area. There are many factors to consider. It is usually worthwhile undertaking a systematic review of a health IT system when there is a lack of clarity about its effectiveness or impact. Or it may be necessary to ascertain what the evidence is revealing about the system. Sometimes, there may be a lot of existing research without clear answers about its effect on patient outcome, clinical performance or about its broader benefits. A systematic review can also be used to obtain an overall picture of the existing evidence and its quality in order to direct future research efforts.

Guides about how to undertake a review of the literature emphasise the importance of arriving at a suitable question. Petticrew and Roberts [3] offer some helpful questions to guide the decision-making process: i) Does this review really need to be carried out? ii) Does anyone want the question answered? iii) Who will benefit from the results [3]?

The formulation of a question involves a decision about the problem or issue under investigation. Is the question about the *quality* of the health IT application (e.g., its

² See also: J.C. Wyatt, Evidence-based health informatics and the scientific development of the field, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

functionality, performance or responsiveness); its *utilisation* (e.g., user perception, ease-of-use or user experience); or about its *benefits* (e.g., appropriateness, health outcomes or efficiency). Identifying and appreciating the multiple perspectives involved (e.g., clinicians, health IT vendors, patients, healthcare managers) is also important for arriving at an appropriate question to guide the literature search.

3. Developing a search strategy

Finding the relevant literature necessary to answer the research question is a crucial part of a successful systematic review. Search strategies are required to be as specific as possible in order to maximise the identification of all relevant evidence. The task of undertaking a robust search strategy begins with ensuring that all the different components of the research question are incorporated into the search strategy. Many researchers begin the process by identifying some key papers that meet the inclusion criteria for review. This can help to identify common subject terms and keywords to be incorporated as part of the full search [6].

Scientific and methodological rigour requires the clear and transparent reporting of the methods used to undertake the search. This allows the search to be replicated and updated at another point in time. It is important to report items such as how the evidence was found (e.g., the databases used to find the relevant literature as discussed further below), the dates of the search period and when it was undertaken, the search terms used (e.g., the Medical Subject Headings [MeSH] terms – a controlled vocabulary thesaurus using various levels of specificity [7]), the electronic search strategy, alongside a flow diagram that accounts for all the available references. There are a number of guides and appraisal tools (discussed further below) which provide valuable information about how to deal with different types of evidence (e.g., qualitative or quantitative) across diverse scientific disciplines.

The quality of the search strategy will depend heavily on the accuracy, precision and completeness of the completed tasks. A 2006 study by Sampson and McGowan [8] of the types of errors in reviews identified a range of search errors including:

- Spelling error (e.g., misspelled search operator).
- Missed spelling variant or truncation errors (e.g., failure to accommodate for multiple spellings of the word ‘randomised’).
- Logical operator error (e.g., mistakes using the logical operators “AND” or “OR”).
- Missed MeSH terms (e.g., where a relevant MeSH term was not used).
- Search strategy was not adapted to suit differing databases [8].

4. Finding the evidence

CENTRAL (The Cochrane Central Register of Controlled Trials), MEDLINE and EMBASE are widely recognised as the most valuable sources of literature in the health field [6]. Normally researchers are expected to have considered more than one database. For instance, although MEDLINE is a highly-utilised database that covers biomedical literature, there are many journals (including some in the health informatics and allied health fields) that may not be indexed in MEDLINE and are more likely to be found in

databases such as the Cumulative Index of Nursing and Allied Health Literature (CINAHL) or the Institution of Engineering and “Inspec” database. Table 1 provides a selected list of some of the most widely used databases relevant to health informatics.

Table 1. Selected list of widely-used literature databases.

BIOSIS Biological and biomedical sciences that includes journal articles along with meeting and conference reports, books and patents. http://scientific.thomsonreuters.com/products/bp
Campbell Collaboration A library of systematic reviews in the areas of education, criminal justice and social welfare. http://www.campbellcollaboration.org
Cochrane Central Register of Controlled Trials (CCRCT) Part of the Cochrane Library, includes details of published articles taken from bibliographic databases (notably MEDLINE and EMBASE), and other published and unpublished sources. http://apps1.jhsph.edu/cochrane/central.htm
Cochrane Database of Systematic Reviews (CDSR) Includes systematic reviews and protocols. http://www.cochrane.org
Human-Computer Interaction Resources Human-Computer Interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems. http://www.hcibib.org/
Cumulative Index of Nursing and Allied Health Literature (CINAHL) Nursing and allied health disciplines such as occupational therapy, emergency services, and social services. http://www.ebscohost.com/cinahl
Database of Abstracts of Reviews of Effects (DARE) A database published by the Centre for Reviews and Dissemination. Also part of the Cochrane Library. http://www.crd.york.ac.uk/crdweb
Embase An international pharmacological and biomedical database http://www.embase.com
Health Services/Technology Assessment (HSTAT) Contains clinical guidelines and the publications of the Agency for Healthcare Research and Quality (AHRQ) and the National Institutes of Health (NIH). http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat
IET Inspec Created by the Institution of Engineering and Technology that provides indexing to a wide range of scientific and engineering papers including computing and information technology. http://www.theiet.org/resources/inspec/
MEDLINE A general medical database accessed through service providers such as Ovid or PubMed. http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed
PsycINFO International literature in psychology and related behavioural and social sciences. Includes book chapters and dissertations as well as journal articles. http://www.apa.org/psycinfo
Science Citation Index/Science Citation Index Expanded Articles from approximately 6,000 major scientific, technical and medical journals and links them to the articles in which they have been cited (a feature known as cited reference searching). http://wokinfo.com/

There are a large number of additional international or specialty databases which either contain evidence themselves (e.g., the Evaluation Database, EvalDB, as a specialist resource specific to the field of evaluation in health informatics) or that are helpful for developing and undertaking a robust search strategy (e.g., EQUATOR). Table 2 provides a selected list of these websites.

Table 2. Valuable websites to aid the planning and undertaking of an evidence search.

Bandolier – Bandolier aims to summarise results research studies in a clear concise way. Searches PubMed and the Cochrane Library each month to identify recently published systematic reviews and meta-analyses. http://www.medicine.ox.ac.uk/bandolier/
Evaluation Database – web-based inventory of evaluation studies in medical informatics (EvalDB) EvalDB contains references and structured information of health IT evaluation studies and is part of a collaboration between the European Federation of Medical Informatics Working Group on “Evaluation in Health Information Systems” and the International Medical Informatics Association Working Group on “Technology Assessment and Quality Development in Health Informatics” https://evaldb.umin.at/
The EQUATOR Network: Enhancing the Quality and Transparency of health Research provides useful guidelines and toolkits for authors of systematic reviews www.equator-network.org
The Health InterNetwork access to Research Initiative (HINARI) HINARI provides access to a number of databases including <i>The Cochrane Library</i> and nearly 4,000 major journals for healthcare professionals in local, not-for-profit institutions in over 100 low-income countries.[6] www.who.int/hinari/en/
The McMaster Health Knowledge Refinery encompasses a collection of projects of the Health Information Research Unit (HiRU) related to the retrieval, appraisal, classification, organisation, dissemination and uptake of health care evidence http://hiru.mcmaster.ca/hiru/HIRU_McMaster_HKR.aspx
The New York Academy of Medicine provides evidence-based resources by specialty http://www.nyam.org/fellows-members/benefits-resources/sections/resources-by-specialty.html including “Clinical Queries” – a user-friendly approach to evidence-based searching on Medline. http://www.nyam.org/fellows-members/ebhc/eb_pubmed.html?referrer=https://www.google.com.au/

5. Tackling publication bias

Publication bias is caused by the publication or non-publication of research findings, depending on the nature and direction of the results [9]. It is often associated with the failure to include “grey literature,” or literature that has not been formally published in sources like a book or a journal [6]. The existence of publication bias can affect the validity of the findings and their application as guidelines and best-practice protocols [10]. There are a number of factors that can contribute to publication bias, including pressure from vendors, managers or publishers to publish only positive findings, or even a lack of commitment to the effort to generate a robust evidence base [11]. The issue is important for health informatics [12]. In 2005 the authors of a US Agency for Healthcare Research and Quality (AHRQ)-funded systematic review of the costs and benefits of health IT highlighted the absence of research evidence about how health IT is used, the individuals who are using it, and the environment in which it is used [13]. This absence can have a major effect on the applicability, generalisability and usability of the existing evidence base to broader audiences.

There are many potential sources of publication bias that can weaken or threaten the validity of systematic reviews. One of these relates to *database bias* where relevant journals are not indexed. Another potential source of bias may occur with conference abstracts which contain only truncated information that may be misleading when compared with full publications [10]. There is also potential for *language bias* when non-English language studies are excluded, and *time-lag bias* if publication of relevant findings are delayed because of a failure to get published [14]. The issue of publication bias is hence integrally bound up with key ethical imperatives related to the safety, usability and cost effectiveness of health IT [11].

There has been a growing awareness of the problems associated with publication bias [10]. In line with advice offered by many critical appraisal tools, many journals

now expect authors to undertake rigorous measures to locate research evidence and to include an assessment of the extent of publication bias in their systematic review [15]. One of these methods involves hand-searching – the manual examination (page-by-page) of journal issues, conference proceedings and relevant research literature. This is an important part of the search strategy because there may be relevant findings not included in the electronic database, or even if they are included, they may be difficult to find because they do not contain relevant search terms in their titles and abstracts [6]. This is a problem often encountered in health informatics, which is still a relatively new and developing research discipline with a rapidly evolving vocabulary of concepts, terms, and applications [16].

Other ways to locate grey literature include contacting authors/institutions of key studies, and searching the Internet (e.g., Google Scholar) for dissertations, conference proceedings or reports. Publication bias in meta-analysis can be examined using funnel plots, a simple graphic test involving a scatter plot of the effect estimates against sample size (or function of sample size) [9]. The results should normally show a symmetrical scatter of points around the total overall estimated effect (see Figure 1). If the scatter of points is skewed and asymmetrical in shape, there may be a bias, even though it may not be possible to identify which biases are present [14].

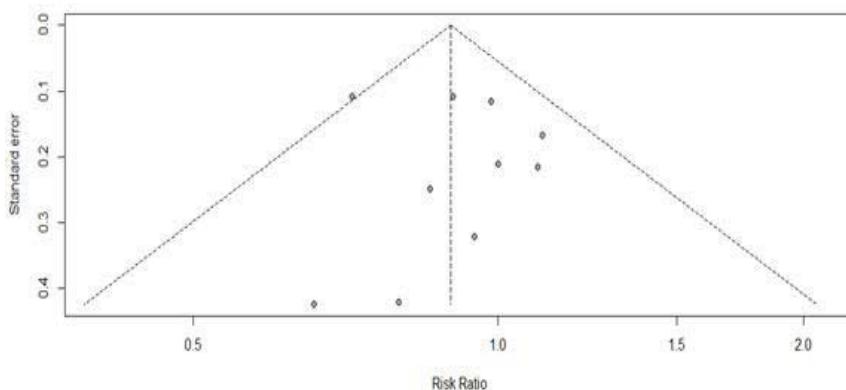


Figure 1. An example of a funnel plot with a symmetrical scatter of points around the total overall estimated effect.

The OPEN Project (Overcome failure to Publish nEgative fiNDings) Consortium has suggested the establishment of a public register which lists all healthcare trials prior to their commencement. This is an idea aimed at detecting and dealing with publication bias [17], and also making research evidence accessible to a wider public. The trial registry idea has had particular resonance in the health informatics field because of its potential to enhance the quality and transparency of health informatics research [18].

6. Organising the evidence

One way to appraise and interpret findings from search results is to present the relevant findings in a table. The process of tabulating the evidence involves identifying logical categories that describe the studies, and then analysing the findings within each of these categories. There is no definitive approach to choosing a set of meaningful categories other than ensuring that they are related to the research question. For instance, a systematic review of the impact of electronic ordering in Emergency Departments may categorise evidence according to the type of electronic order studied, (e.g., laboratory, medical imaging, etc. [19]); according to the different types of study design (e.g., randomised controlled trial, case-control study, etc.); or by outcome (e.g., adverse events rates, drug dosing etc.) [20].

A decision also needs to be made about the characteristics of the findings. This should include an assessment of the heterogeneity of the findings so as to ascertain whether or not the studies are comparable in terms of population, intervention, study characteristics and primary outcomes of interest, and whether a meta-analysis is possible or necessary.³ There are advantages to undertaking meta-analysis including: increased statistical power of results, increased precision for the estimation of the effect of the health IT system, and the ability to explore difference between studies and groups of studies to investigate potential causes and effects [3]. Conversely a meta-analysis may not be suitable when the studies in a systematic review include different hypotheses and measures. In such cases it may be unwise to make conclusions based on inappropriate aggregations across sub-groups, as this may lead to an incorrect conclusion about the benefit (or otherwise) of the system. Different findings may be due to a range of reasons such as the variances in the products themselves, their implementation and their use in different care settings [21].

7. Critically appraising the evidence

Critical appraisal of all the relevant studies is a key component of evidence reviews. It is important to consider the quality of the evidence, its validity and the degree to which the studies are free from study bias.

Common types of study bias in health informatics can come from differences in the comparison groups (*allocation bias*), or as a consequence of systematic differences in those involved in the study (*selection bias*). It is also important to consider the impact of other, extraneous or unintended factors on the study (e.g., *contamination* or *confounders*). Research studies are expected to provide an *adequate sample size* to have the ability (*statistical power*) to detect significant differences between the comparison groups. A lack of a significant effect could be due to the study lacking sufficient numbers rather than the failure of the health IT system. It is also important to consider the independence of participants – a lack of independence is known as the

³ See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

clustering effect (e.g., when participants are from the same GP practice or community group).

Health informatics researchers additionally need to carefully consider the integrity of the health IT implementation, including the independence of the researcher evaluating the system (i.e., is the researcher also the designer of the system?). They should also consider whether the failure or ineffectiveness of a health IT system may have been a consequence of the incomplete delivery and implementation of the system – this type of evaluation error has been described as a *Type III error* [22].⁴

The validity of the conclusions drawn from systematic reviews depends on the quality of the systematic reviews. In 2012 Weir et al. assessed the quality of 13 systematic reviews of the impact of computerised provider order entry systems on clinical outcomes. The authors noted the wide variability in the quality and scope of the reviews. While some content areas including the reporting of search strategies, selection of articles, and description of original studies were robust, others like the diversity of primary studies and the assessment of the scientific quality of studies were less strong [23]. There is wide and growing range of reporting aids, frameworks and critical appraisal tools which can be used to enhance the quality of research and aid the task of critical appraisal. Within the health informatics field the Good Evaluation Practice guideline for Health Informatics (GEP-HI) provides a systematic approach to the design and execution of evaluation studies and the building of a stronger health informatics evidence base [24].⁵ This is supplemented by the Statement on Reporting of Evaluation Studies in Health Informatics (STARE-HI), which was published in 2009 [25] and has been endorsed by the International Medical Informatics Association (IMIA) and the European Federation of Medical Informatics (EFMI) as a guideline for the reporting of evaluation studies [1].⁶

The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) guidelines provide a system for rating the quality of evidence in systematic reviews and the strength of recommendations in guidelines. The GRADE working group website www.gradeworkinggroup.org/index.htm provides a comprehensive toolkit [26]. Other tools which researchers can utilise include instruments that perform an appraisal (e.g., Assessment of Multiple Systematic Reviews [AMSTAR] [27]); provide a reporting checklist (e.g., Preferred Reporting Items for Systematic Reviews and Meta-Analyses [PRISMA] [28] and Critical Appraisal Skills Program [CASP][29]); or a reporting standard (e.g., Meta-Analysis of Observational Studies in Epidemiology [MOOSE] [30]). These instruments vary according to the type of study they target. Instruments for meta-analyses include MOOSE and PRISMA; for qualitative research systematic reviews there is STARLITE (Sampling strategy, Type of study, Approaches, Range of years, Limits, Inclusion and exclusions, Terms used, Electronic sources) [31] and for systematic reviews there is AMSTAR.

⁴ See also: J. Brender, Theoretical basis of health IT evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

⁵ See also: P. Nykänen et al., Quality of health IT evaluations, in: *ibid*.

⁶ See also: E. Ammenwerth et al., Publishing health IT evaluation studies, in: *ibid*.

8. Being sensitive and aware of context

Context can be defined as a set of factors or attributes that can affect a health IT implementation. This may include the organisation's leadership, mission, climate or learning environment [32]. The task of designing, implementing and sustaining health IT systems is complex. It involves a lot of people (often across many settings), who are required to coordinate the storage, management, analysis and display of large amounts of mostly heterogeneous information [33]. In 2011, the US Committee on Patient Safety and Health Information Technology, Institute of Medicine, warned against viewing health IT as a single product which is expected to lead inevitably to improved health care. Rather health IT should be understood as a collection of hardware and software working in conjunction with people, processes, and workflow [21].

Existing reviews of health informatics research have commented on the scarcity of information about contextual domains such as the size, location, academic status and the implementation components of the intervention [34]. The reason contextual information is not always reported is that many implementations are reviewed as if they were the "solution" which if shown to work in one situation will also work elsewhere. Many health informatics researchers have embraced elements from a social science approach to systematic reviews, recommending that the traditional *Population, Intervention, Comparison, Outcomes (PICO)* approach to undertaking systematic reviews should now become *PICO(C)* by including a "C" for context [3].

For instance, Shekelle and Goldzweig's systematic review of the costs and benefits of health information technology published in 2009 included a valuable list of categories for data extraction, which identified factors related to the health IT system's implementation strategy, its penetration, interoperability and sustainability and even its financial context (e.g., managed care or capitation) and long-term cost issues [13]. Such contextual information can improve our understanding about how to maximise value from health IT and deal with any potential negative effects.

9. Interpreting the evidence

There are some basic considerations which contribute to a robust interpretation of the results and the appraisal of the quality of the evidence. This can begin with an assessment of the *overall strength of the evidence*, particularly its *consistency* and *validity*. It is also important to consider the *integrity* and *applicability* of the evidence, as this constitutes an important part of processing the evidence and the resulting recommendations. This may include an assessment of the *reach* of the system, its *sustainability* and *effectiveness*.

In their overview of eHealth and its impact on the quality and safety of healthcare, Car, Black et al. concluded that the volume of research publications in health informatics was poorly collated and of variable quality, which made it difficult to synthesise and interpret [16]. Theoretical approaches can play a valuable part in the interpretation of findings by providing an analytical frame of reference or schema for understanding the significance of research findings. In health informatics, some researchers employ *socio-technical* theoretical approaches to help analyse, and interpret the complex interrelationships between technology (e.g., software, hardware), people

(e.g., clinicians, patients), processes (e.g., workflow), organisation (e.g., decisions about how health IT is applied and incentivised), and the external environment (e.g., regulations, public opinion) [21].⁷

Another area of expanding interest is the role of IT as an enabler of *patient-centred care*, *care coordination* and *shared decision-making*. Despite the enthusiasm about health IT's potential contribution, the evidence up to now has not been conclusive [35]. In some part, this may be due to the failure to adequately explore how electronic systems (e.g., patient portals and personal health records) actually contribute to better patient care [36, 37]. This situation points to the need to explore how greater access to information and evidence, and communication with health care providers can contribute to improved patient care.

10. Conclusion

For many years now, there has been an ongoing concern within the health informatics community, that despite the massive increase in the research literature dedicated to the evaluation of health IT, understanding of how it can be translated into better and more effective health care is still variable [34]. Evidence-based health informatics, with its focus on methodological rigour and transparency, provides an effective means for enhancing the quality of care to meet the needs of patients, clinicians, health care administrators and policy makers, now and into the future.

Recommended further readings

1. G.H. Guyatt, A.D. Oxman, H.J. Schünemann, et al., GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology, *Journal of Clinical Epidemiology* **64**(4) (2011), 380-382.
2. C. Lefebvre, E. Manheimer, J. Glanville, *Searching for studies*, in: Cochrane handbook for systematic reviews of interventions, J. Higgins, S. Green, eds., 2008, Wiley: New York. pp. 95-150.
3. M. Petticrew, H. Roberts, *Systematic Reviews in the Social Sciences*, Blackwell Publishing, Cornwall, 2006.
4. J.A. Sterne, M. Egger, D. Moher, *Addressing reporting biases*, in: Cochrane Handbook for Systematic Reviews, J. Higgins, S. Green, eds., 2008, Wiley, New York. pp. 297-333.

Food for thought

1. How do flaws in search strategy (e.g., databases used, search script) affect the validity and robustness of a health informatics systematic review?
2. Do you think that health informatics has been, or is being affected by publication bias?

⁷ See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

3. What methods can be employed to address publication bias?
4. What possible effect does the context and setting of a health IT system have on the external validity of a study?

References

1. M. Rigby, E. Ammenwerth, M.-C. Beuscart-Zephir, et al., Evidence Based Health Informatics: 10 Years of Efforts to Promote the Principle, *IMIA Yearbook of Medical Informatics* 2013, 34-46.
2. C. Bambra, Real world reviews: A beginner's guide to undertaking systematic reviews of public health policy interventions, *Journal of Epidemiological Community Health* **65** (2011), 14-19.
3. M. Petticrew, H. Roberts, *Systematic Reviews in the Social Sciences*, Blackwell Publishing, Cornwall, 2006.
4. F. Lau, C. Kuziemsky, M. Price, et al., A review on systematic reviews of health information system studies, *Journal of the American Medical Informatics Association* **17**(6) (2010), 637.
5. National Health and Medical Research Council, *NHMRC levels of evidence and grades for recommendations for developers of guidelines*, https://www.nhmrc.gov.au/_files_nhmrc/file/guidelines/developers/nhmrc_levels_grades_evidence_120423.pdf, last access 11 February 2016.
6. C. Lefebvre, E. Manheimer, J. Glanville, *Searching for studies*, in: Cochrane handbook for systematic reviews of interventions, J. Higgins, S. Green, eds., 2008, Wiley, New York, pp. 95-150.
7. U.S. National Library of Medicine, *Medical Subject Headings (MeSH)*, <https://www.nlm.nih.gov/pubs/factsheets/mesh.html>, last access 11 February 2016.
8. M. Sampson, J. McGowan, Errors in search strategies were identified by type and frequency, *Journal of Clinical Epidemiology* **59**(10) (2006), 1057-63.
9. J.A. Sterne, M. Egger, D. Moher, *Addressing reporting biases*, in Cochrane Handbook for Systematic Reviews, J. Higgins, S. Green, eds., 2008, Wiley, New York, pp. 297-333.
10. S. Parekh-Burke, C.S. Kwok, C. Pang, et al., Uptake of methods to deal with publication bias in systematic reviews has increased over time, but there is still much scope for improvement, *Journal of Clinical Epidemiology* **64**(4) (2011), 349-357.
11. M. Rigby, *Evaluation: 16 powerful reasons why not to do it-and 6 over-riding imperatives*, in: Studies in Health Technology and Informatics, V. Patel, R. Rogers, R. Haux, eds., 2001, IOS Press, Amsterdam, pp. 1198-1202.
12. C.P. Friedman, J.C. Wyatt, Publication bias in medical informatics, *Journal of the American Medical Informatics Association* **8**(2) (2001), 189-191.
13. P. Shekelle, C.L. Goldzweig, *Costs and benefits of health information technology: an updated systematic*, Quest for Quality and Improved Performance, The Health Foundation, London, England, 2009.
14. P. Sedgwick, Meta-analysis: testing for reporting bias, *British Medical Journal* **350** (2015), g7857.
15. V. Smith, D. Devane, C.M. Begley, et al., Methodology in conducting a systematic review of systematic reviews of healthcare interventions, *BMC Medical Research Methodology* **11**(1) (2011), 15.
16. J. Car, A. Black, C. Anandan, et al., *The impact of eHealth on the quality & safety of healthcare: a systemic overview & synthesis of the literature*, Imperial College London & The University of Edinburgh, London, 2008.
17. E. Wager, P. Williams, "Hardly worth the effort"? Medical journals' policies and their editors' and publishers' views on trial registration and publication bias: quantitative and qualitative study, *British Medical Journal* **347** (2013), f5248.
18. G. Eysenbach, Tackling publication bias and selective reporting in health informatics research: register your eHealth trials in the International eHealth Studies Registry, *Journal of Medical Internet Research* **6**(3) (2004).
19. A. Georgiou, M. Prgomet, R. Paoloni, et al., The Effect of Computerized Provider Order Entry Systems on Clinical Care and Work Processes in Emergency Departments: A Systematic Review of the Quantitative Literature, *Annals of Emergency Medicine* **61**(6) (2013), 644-653.
20. S. Eslami, N.F. Keizer, A. Abu-Hanna, The impact of computerized physician medication order entry in hospitalized patients—a systematic review, *International Journal of Medical Informatics* **77**(6) (2008), 365-376.
21. Committee on Patient Safety and Health Information Technology, Institute of Medicine, *Health IT and Patient Safety: Building Safer Systems for Better Care*, National Academies Press, Washington DC, USA, 2011.

22. C.E. Basch, E.M. Sliepcevich, R.S. Gold, et al., Avoiding type III errors in health education program evaluations: a case study, *Health Education & Behavior* **12**(3) (1985), 315-331.
23. C.R. Weir, N. Staggers, T. Laukert, Reviewing the impact of computerized provider order entry on clinical outcomes: the quality of systematic reviews, *International Journal of Medical Informatics* **81**(4) (2012), 219-231.
24. P. Nykänen, J. Brender, J. Talmon, et al., Guideline for good evaluation practice in health informatics (GEP-HI), *International Journal of Medical Informatics* **80**(12) (2011), 815-827.
25. J. Talmon, E. Ammenwerth, J. Brender, et al., STARE-HI-Statement on reporting of evaluation studies in Health Informatics, *International Journal of Medical Informatics* **78**(1) (2009), 1-9.
26. G.H. Guyatt, A.D. Oxman, H.J. Schünemann, et al., GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*, *Journal of Clinical Epidemiology*. **64**(4) (2011), 380-382.
27. B.J. Shea, C. Hamel, G.A. Wells, et al., AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews, *Journal of Clinical Epidemiology* **62**(10) (2009), 1013-1020.
28. D. Moher, A. Liberati, J. Tetzlaff, et al., Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, *British Medical Journal* **339**(jul21_1) (2009), b2535.
29. Critical Appraisal Skills Programme (CASP), *Making sense of evidence*, <http://www.casp-uk.net/#/casp-tools-checklists/c18f8>, last access 11 February 2016.
30. D.F. Stroup, J.A. Berlin, S.C. Morton, et al., Meta-analysis of observational studies in epidemiology: a proposal for reporting, *Journal of the American Medical Association* **283**(15) (2000), 2008-2012.
31. A. Booth, "Brimful of STARLITE": toward standards for reporting literature searches, *Journal of the Medical Library Association* **94**(4) (2006), 421.
32. The Health Foundation, *Perspectives on context*, The Health Foundation, London, 2014.
33. A. Georgiou, A. Marks, J. Braithwaite, et al., Gaps, Disconnections, and Discontinuities - The Role of Information Exchange in the Delivery of Quality Long-Term Care, *The Gerontologist* **53**(5) (2013), 770-779.
34. S.S. Jones, R.S. Rudin, T. Perry, et al., Health information technology: an updated systematic review with a focus on meaningful use, *Annals of Internal Medicine* **160**(1) (2014), 48-54.
35. C.L. Goldzweig, G. Orshansky, N.M. Paige, et al., Electronic Patient Portals: Evidence on Health Outcomes, Satisfaction, Efficiency, and Attitudes: A Systematic Review, *Annals of Internal Medicine*. **159** (10) (2013), 677-687.
36. T. Otte-Trojel, A. de Bont, T.G. Rundall, et al., How outcomes are achieved through patient portals: a realist review, *Journal of the American Medical Informatics Association*. **21** (4) (2014), 751-757.
37. M. Rigby, A. Georgiou, H. Hyppönen, et al., Patient Portals as a Means of Information and Communication Technology Support to Patient-Centric Care Coordination-the Missing Evidence and the Challenges of Evaluation. A joint contribution of IMIA WG EVAL and EFMI WG EVAL, *Yearbook of Medical Informatics* 2015, 148-159.

Evaluation of Health IT in Low-Income Countries

Tom OLUOCH^{a,1} and Nicolet F. de KEIZER^b

^aU.S. Centers for Disease Control and Prevention - Division of Global HIV & TB,
Nairobi, Kenya

^bDepartment of Medical Informatics, Academic Medical Center, University of
Amsterdam, Amsterdam, The Netherlands

Abstract. Low and middle income countries (LMICs) bear a disproportionate burden of major global health challenges. Health IT could be a promising solution in these settings but LMICs have the weakest evidence of application of health IT to enhance quality of care. Various systematic reviews show significant challenges in the implementation and evaluation of health IT. Key barriers to implementation include lack of adequate infrastructure, inadequate and poorly trained health workers, lack of appropriate legislation and policies and inadequate financial 333indicating the early state of generation of evidence to demonstrate the effectiveness of health IT in improving health outcomes and processes. The implementation challenges need to be addressed. The introduction of new guidelines such as GEP-HI and STARE-HI, as well as models for evaluation such as SEIPS, and the prioritization of evaluations in eHealth strategies of LMICs provide an opportunity to focus on strategic concepts that transform the demands of a modern integrated health care system into solutions that are secure, efficient and sustainable.

Keywords. Low-income countries, medical informatics, evaluation, eHealth.

1. Background

The last three decades have seen substantial growth in innovation and development of health information systems globally, encompassing both successes and failure [1]. The World Health Organisation (WHO) and the American Institute of Medicine (IOM) report that Health Information Technology (Health IT) has the potential to reduce medical errors and improve patient safety [2;3]. EMRs can improve health care, through among other means better adherence to therapeutic guidelines and protocols, informing clinical decisions, and decreasing medication errors [4]. WHO defines eHealth as the transfer of health resources and healthcare by electronic means [5]. According to WHO, one of the three main components of an eHealth system is the delivery of health information (e.g. patient data) to health professionals through the Internet and telecommunications to improve quality of care through well informed

¹ Corresponding author: Dr. Tom Oluoch, U.S. Centers for Disease Control and Prevention, Nairobi, Kenya, hof0@cdc.gov.

clinical decisions [6]. In this contribution, we use eHealth and health IT interchangeably.

Low and middle income countries (LMICs) have lagged behind in adopting Health IT despite bearing a disproportionately large share of major global public health threats, including maternal and child health, and infectious diseases like HIV, Tuberculosis and Malaria [7-9]. LMICs, in this chapter, are countries described by the World Bank as having a Gross National Income (GNI) per capita of US\$ 12,736 or lower [10]. Fritz *et al* describe LMICs as parts of the world in which resources for healthcare services (e.g. financial and human resources and infrastructure) are scarce [11]. Most LMICs are located in sub-Saharan Africa, South America and South-East Asia. LMICs, like their counterparts in the developed world, need strong Health IT to improve quality of health care. This is not without challenges, as discussed below, but done correctly also provides an opportunity to bring skills and knowledge quickly to currently under-provided areas.

Rigorous evaluation of health IT is essential in ensuring that the interventions are safe, beneficial and cost-efficient, set in the local context [12]. Various studies have shown that implementation and evaluation of Health IT in low income countries are still in early stages as shown by the limited number of published studies [13-15]. In this chapter, we describe the challenges of implementing Health IT in LMICs, the current state of evaluations, and future opportunities. We focus on evaluation of systems used in patient care.

2. Challenges of implementation and evaluation of health IT in LMICs

LMICs experience unique challenges which include infrastructural, human capacity and policy limitations. The early state of implementation of Health IT in LMICs has a direct correlation with the limited evaluation studies conducted and published so far [16]. The consensus statement of the WHO Global eHealth Evaluation Meeting held in Bellagio, Italy, in September 2011 resolved that: *“To improve health and reduce health inequalities, rigorous evaluation of eHealth is necessary to generate evidence and promote the appropriate integration and use of technologies.”* [17].

2.1. Inadequate infrastructure

Lack of basic infrastructure, which includes reliable electric power, adequate computers and related hardware, secure accommodation for computing devices, and stable and fast Internet connectivity, are often a hindrance to the implementation of health IT. Some rural clinics in sub-Saharan Africa experience power outages that last up to eight hours a day. There has been limited investment in innovative, affordable and sustainable technologies such as solar power, on-site backup generators and rechargeable invertors in rural areas. Lack of routine maintenance of hardware and software due to lack of policies or technically qualified personnel often results in extended downtime, which is a waste of valuable resources that lie unused for weeks or months.

Technologies used in telemedicine in LMICs need to be tolerant to low-bandwidth and intermittency of connectivity [16;18]. Although there has been a rapid expansion of cellular networks in many LMICs which has inspired the growth of application of mobile technology solutions in health (mHealth), high initial cost of procuring and

installing telemedicine equipment, high cost of bandwidth, poor network signal and slow data transmission rates in rural areas remain obstacles to efficient use. In areas with intermittent power supply, unreliable Internet connectivity and inadequate infrastructure, relying on servers and computers for radiographic images used in telemedicine involves considerable risks. Piette et al. suggest that for meaningful use of large-scale implementation of picture archiving and communication in clinical care in LMICs, there is need for effective off-site data backup that can be restored in case of data loss [19]. Recent initiatives such as Google's helium-filled balloons to provide nation-wide internet access to Sri Lanka might be a solution to provide cheap or free Wi-Fi to people in remote rural areas around the world [20].

2.2. Limited human capacity

Inadequate and often poorly trained health workers are a key challenge to the implementation of health IT in LMICs. Of the 57 countries classified by WHO as having an acute shortage of health workers, 36 are in sub-Saharan Africa [21]. Health care workers' limited computer skills and frequent transfer of health workers between health facilities also hinder successful implementation of technology based solutions. Many medical schools in LMICs have not yet integrated training on informatics (e.g. basic computer skills as well as specific topics such as EMRs and telemedicine) as part of routine clinical care and this poses a challenge once the doctors graduate and have to use computers to support patient care [22].

Many LMICs experience serious shortages of medical informatics personnel who are well trained and have experience in designing, implementing and evaluating health IT solutions in resource-limited settings [22]. The few highly skilled medical informaticians tend to live and serve hospitals in cities and rarely offer their services in rural areas where majority of the patient population seek healthcare services [23]. Other key cadres such as programmers, network and database administrators and hardware technicians are mainly found in cities.

2.3. Lack of appropriate legislation and policies

Health IT is a relatively new field in many LMICs and the majority of them have not revised the necessary legislation and regulatory policies to ensure appropriate application and practice. Standards and guidelines that are customized for resource-limited settings are now emerging in countries that have recognized the need for well-regulated practice of health informatics. WHO and the International Telecommunication Union (ITU) have recently launched the National eHealth Strategy Toolkit [24] to help countries develop eHealth strategies which prioritize key technology-based interventions that are relevant for their settings. South Africa and Kenya are examples of countries in sub-Saharan Africa that have successfully developed and implemented eHealth strategies, which include evaluation of health IT [25;26].

2.4 Inadequate financial investment

Health IT solutions are often expensive. The initial capital investment in hardware and software, and the recurrent costs of maintenance and ongoing capacity building, make them unaffordable to many LMICs. Despite evidence from developed countries on the benefits of health IT, there has not been adequate investment by country governments and the private sector. Additionally, rigorous evaluation of installed systems to determine their effectiveness on quality of healthcare is often seen as a low-priority activity that can be omitted when available financial resources are not adequate.

Blaya et al. recommend that major funding agencies of health IT in LMICs, such as the US Centers for Disease Control and Prevention (CDC), the US Agency for International Development (USAID) and the World Bank should include resources for evaluation of eHealth systems developed and implemented in LMICs and make it a requirement for future funding [27]. This is consistent with a recommendation by Ammenwerth et al. that rigorous evaluation of health IT is of high importance to decision makers and users [28]. A meeting of the heads of eight global health agencies and the Global Health Information Forum (GHIF) in 2010 recommended an increase of investments to strengthen country health information systems [16].

3. Current status of Health IT evaluation in LMICs

The barriers for health IT implementation need to be addressed before new systems are implemented and their effectiveness evaluated. Furthermore these barriers can be part of the outcome measures used in evaluation studies and/or taken into account when interpreting the results on process and outcome measures. Many LMICs now have eHealth strategies; however, there is still sub-optimal financial investment by national governments in the implementation and evaluation of the strategies. Health systems projects and evaluations are often funded by donors and collaborators who in many cases drive the agenda and the identified evaluation topics may not always align with the top priorities of the country where the evaluations are conducted [23;29]; they may even find the evaluation concept challenging to their decision-making, or a potential source of criticism for apparently diverting investment money from direct service investment [30].

The rapid increase in use of health IT in LMICs is mainly driven by reduced cost of hardware (including digital cameras, videoconferencing units, and medical equipment used in telemedicine), wider coverage of Internet access and availability of affordable mobile technology. However, systematic reviews conducted recently show the pre-mature state of health IT evaluation in these settings [11;19;27;31]. The types of health IT systems assessed in evaluation studies include EMR based clinical decision support systems (CDSS), mHealth and telemedicine. Within mHealth, mobile devices include cellular phones and smartphones, tablets, personal digital assistants, patient monitoring devices, and mobile telemedicine devices.

3.1. Description of studies conducted

The eight published systematic literature reviews on health IT in LMICs that informed this contribution were conducted between 2010 and 2014 in diverse geographical settings including Southeast Asia, South America, the Caribbean and sub-Saharan Africa [11;14;15;19;27;31-33]. Of the eight systematic reviews, five described studies on EHR/EMR and CDSS, four focused on mHealth while three reviewed evaluations in the area of telemedicine. Some reviews described more than one focus area (Table 1).

All the systematic reviews concur that the studies are of varying and overall low quality. Blaya et al. included two articles in which an evaluation was never conducted because the systems implementation was not completed but also noted that studies on unsuccessful systems or those with negative associations between health IT and anticipated health outcomes were rarely reported [27]. Earlier evaluations in low income countries were mainly descriptive studies, but recent ones apply more rigorous quantitative methods including randomised controlled trials [27]. The reviews describe the application of health IT in maternal and child health, communicable (infectious) and non-communicable diseases, and for acute and chronic care. An area with a big gap in evaluation studies is the effect of health IT on maternal and child health; Fritz et al. report that only 2% of the studies included in their systematic review had outcomes related to maternal and child health (9).

The various studies described the effectiveness of health IT on major public health problems in the respective locations LMICs. For example, studies conducted in sub-Saharan Africa were likely to describe the application of health IT on AIDS and HIV, TB or Malaria while those conducted in South America were likely to be on TB and non-communicable diseases such as diabetes, hypertension and cancers. The IT solutions described in the majority of papers were implemented in public clinics and hospitals.

3.2 Study design and outcome measures

The majority of the studies reported in the eight reviews were quasi-experimental although two included some randomized controlled trials (RCTs). The quasi-experimental studies applied some quantitative measure of effectiveness and were either descriptive or used a before-after design. Many of the telehealth studies were descriptive of experiences in implementing telehealth solutions (including technological modalities such as synchronous, real-time teleconsultations and asynchronous technologies). Khanal et al. reported that out of the 46 studies fulfilling their inclusion criteria, 36 had some quantitative measure of effectiveness on process although clinical effectiveness and cost-saving were rare [14]. This is similar to the review by Blaya which reported that 72% of studies were quantitative, of which 40% had some statistical analysis [14;27].

The systematic review by Oluoch et al. showed that very few studies had been conducted in low income countries on the effect of CDSS on HIV care. Of the 12 papers included in the review, seven (~60%) presented descriptive studies while pre-post (n=3), controlled trial (n=1) and qualitative (n=1) designs were also reported. None of the papers described a study based on an RCT. Nearly all the studies described improvements in clinical processes but none demonstrated associations between health IT and a health outcome [15]. Although the review by Piette et al. did not provide a

Table 1. The most common design and outcome measures of reported health IT studies in low income countries.

EHR = Electronic Health Record, EMR = Electronic Medical Record, CDSS = Clinical Decision Support System, RCT = Randomized Controlled Trial

Systematic review	No. of studies (No. of RCTs)	Focus	Region	Included Study Designs	Outcome Measure(s) Identified
Blaya et al.	45 (9)	EHR, lab and pharmacy systems, CDSS	Sub-Saharan Africa, South America, Asia, Eastern Europe	Qualitative, Descriptive, Controlled Trials*, RCT	Staff productivity, patient waiting time, staff satisfaction, data quality, time communicating lab results, time ordering drugs, prescription errors, patient tracing, provider performance, tuberculosis treatment completion rates, cost effectiveness and clinic attendance
Fritz et al.	47 (0)	EMR	Sub-Saharan Africa, South America, Asia	Descriptive	Factors for successful EMR implementation
Hall et al.	76 (4)	mHealth	Sub-Saharan Africa, South America, Asia, Eastern Europe	Controlled Trials*, RCT, Pre-post	Treatment adherence, appointments, data collection, diabetes control, antenatal care, vaccination rates
Kallander et al.	Not stated	mHealth	Sub-Saharan Africa, South America, Asia	Descriptive	Appointments, health behaviour
Khanal et al.	46 (0)	Telemedicine	Sub-Saharan Africa, South America, Asia, Eastern Europe	Descriptive	Cost-effectiveness, clinical effectiveness
Luna et al.	11 (0)	EMR, Telemedicine and mHealth	Developing countries (continents not specified)	Systematic review of reviews	Efficiency in process management, diabetes patients prognosis, data quality
Oluoch et al.	12 (0)	EMR/CDSS	Sub-Saharan Africa, South America	Qualitative, Descriptive, Pre-post and Controlled Trial*	Lab orders, data errors, missed appointments, patient waiting time and barriers to CDSS implementation
Piette et al.	N/A	EMR, mHealth, Telemedicine	Sub-Saharan Africa, South America	Systematic review of reviews	Practitioner performance, guideline adherence, lab ordering, data errors, hospital stay, telemedicine diagnostic accuracy

* - non-randomized controlled trial

breakdown by study design type, the main outcomes were on the effectiveness of health IT on quality of care and healthcare cost [19]. Fritz et al. indicate that only 25% of the papers included in their review were evaluations. The majority were descriptive studies that discussed key areas of successful implementation of EHRs [11].

While there has been a growing number of evaluation studies on mHealth solutions in LMICs, as reported by Hall et al. and Kallander et al., the quality and quantity of the evidence is limited by several factors (e.g. high risk of bias and heterogeneity) and only a few demonstrate impact on clinical outcomes [31;33]. Some key areas which have not been rigorously evaluated and reported include the use of mHealth in clinical decision support, job aids and use of mobile devices in telehealth.

Although the studies mainly reported benefits on patient care and clinical processes, none reported the cost of implementation and maintenance of such systems. The review by Luna et al. did not include specific study designs or outcomes but in their synthesis of the evidence from various studies reiterated that the majority of published reviews concurred that the papers evaluated were generally of poor quality [32].

3.3 Study limitations

The studies reviewed had several limitations. The findings reported in most papers were not generalizable due to the limitations of the study designs, small sample sizes and the statistical analysis methods used which may not have effectively corrected for confounders. In the mHealth systematic reviews, for example, it was unclear whether the reported effects of mobile technologies could have been due to the “novelty” effect resulting from the excitement of use of new technology but which gradually wears off as the users get more accustomed to the mobile devices.

A key limitation noted by Blaya et al. was that evaluations were conducted by the developers of the systems hence potentially introducing bias. Low data quality was also cited as a factor that reduced the validity of findings in some studies. The RCTs reported were based on small pilots with limited sample size and generated evidence that is not easily generalisable. Hall et al. and Kallander et al. both recommend scaling up the use of mHealth solutions in order to strengthen the evidence base [31;33]. Finally, there were rarely studies that triangulated multiple methods, including quantitative and qualitative methods, not only to measure a possible effect but also to understand barriers and facilitators of effective implementation of the health IT intervention.

4. Future opportunities for evaluations of health IT

The systematic reviews included in this contribution (Table 1) demonstrate that the potential of health IT in LMICs remains largely untapped, but equally importantly that the evidence on the best forms of investment and on how to overcome the natural barriers effectively remains minimal due to the lack of investment in objective scientific evaluation. Multilateral and bilateral partnerships, increased investments by country governments, as well as the engagement of the private sector present new opportunities for investing in technology solutions that address the unique challenges in resource-limited settings. Development and implementation of eHealth strategies is increasingly highlighting the relevance and importance of evaluating health IT

solutions and recommending implementation models that are context appropriate. These will see an extension of coverage of health IT (and thus effective healthcare availability) in rural settings which have been previously underserved and not adequately evaluated.

Large randomised trials such as that by Zurovac et al., provide strong evidence of the benefits of health IT [34]. As conducting large RCTs in low income settings might still be rare, Piette et al. recommend the adoption of new approaches to operational research, incorporating qualitative and quantitative methods as well as community-based participatory research and organizational theory to complement RCTs as a way of demonstrating that the benefits of health IT can be adaptive to multiple environments, including resource-limited settings in low income countries [19].

It is important to identify and develop skills and competencies, consistent with low-resources settings and health systems, that will be necessary to achieve the full potential of health IT applications [35]. Synthesizing the expertise of indigenous knowledge and understanding of individual countries or regional groupings, and generic expertise on the potential of eHealth innovations, is necessary to create an informed picture or possibilities or effective evaluation [12;36]. Collaborations and experience sharing between universities and research institutions in LMICs and those in developed countries with mature curricula for post-graduate training and health IT evaluation capacity can do much to help improve the quality of evaluations [37]. Such skills can be cascaded down to lower level health workers and health IT staff to enhance the ability to conduct evaluations in LMICs. Leveraging the research capacity within local universities, research institutes and industry to design and implement evaluation of health IT that informs delivery of appropriate technology is a practical solution.

5. Discussion and conclusion

Implementation of health IT in LMICs needs to grow from its current early phase. Due to the disproportionately large population suffering from major infectious diseases like HIV, tuberculosis and malaria, the increase in reported cases of chronic and non-communicable diseases such as diabetes and cancers, and high maternal and child mortality, the potential benefit of health IT to improve health care by informing clinical decisions, better adherence to therapeutic guidelines and protocols, increasing access to quality healthcare services in rural areas and decreasing medication errors is large. To gain the most benefit from health IT implementations we need robust evidence-based knowledge about antecedents of health IT implementation success in low-resource settings.

Implementation of health IT in LMICs still faces major challenges including weak infrastructure, limited computer skills among health workers and lack of appropriate policies. A recent study by Tilahun et al. [38] used the updated Delone & Maclean model [39] to identify antecedents of EMR success. They concluded that EMR implementers and managers in those settings should give priority to improving service quality of the hospitals like technical support and infrastructure; providing continuous basic computer trainings to health professionals; and paying attention to the system and information quality of the systems they want to implement. There is need to address the barriers to implementation of health IT and partnerships between LMICs and multi-lateral, bilateral organizations as well as the private sector provide an opportunity for

investment in context-appropriate technologies that are sustainable. Universities and research institutions also have an opportunity to integrate training on application and evaluation of health informatics.

Models such as the Software Engineering Initiative for Patient Safety (SEIPS) [40] might be useful before health IT implementation to check whether the five components of the work system (person, tasks, tools and technologies, physical environment, organizational conditions) are ready for implementation. Furthermore this model can be used in the evaluation phase to obtain better understanding of the antecedents of health IT implementation success.

A majority of the systematic reviews on health IT in LMICs mentioned weak study designs and reporting quality of evaluation studies in low income settings hampering evidence-based health informatics. Extending application of existing guidelines such as the Guidelines for Good Evaluation Practice in Health Informatics (GEP-HI)² and Statement on Reporting of Evaluation studies in Health Informatics (STARE-HI) [41;42]³ to fit LMICs' needs is important. The guidelines can be applied to elaborate on how specific barriers to the implementation of health IT in LMICs (e.g. lack of reliable electricity and low computer literacy among health workers) were addressed and how they impacted the evaluation of the effectiveness of technology on healthcare in these settings. Application of such guidelines as part of national eHealth strategies would be an initial step towards having a structured approach to evaluations and reporting of findings.

As technology changes and new health challenges emerge (e.g., the increase in cases of non-communicable diseases), there are new opportunities for implementation and evaluation of context-relevant health IT that demonstrate the ability of technology to improve the quality of care, practitioner performance, clinical processes, cost effectiveness and expanded access to healthcare. Health IT, appropriately designed to the setting, has the potential to bring health knowledge and skills quickly to underserved areas. However; there is need for targeted investment to address infrastructural, IT skills and policies to facilitate focused evidence from evaluations informed by appropriate tools and principles [17].

In conclusion, evaluation of health IT projects has to focus on strategic concepts in order to provide the firm evidence on how to transform the requirements of a modern integrated health and social care system into solutions that are relevant, user-friendly, secure, efficient and sustainable within the context of the LMICs.

Disclaimer: The findings and conclusions in this contribution are those of the authors and do not necessarily represent the official position of the US Centers for Disease Control and Prevention or the Agency for Toxic Substances and Disease Registry.

² See also: P. Nykänen et al., Quality of health IT evaluations, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

³ See also: E. Ammenwerth et al., Publishing health IT evaluation studies, in: *ibid.*

Recommended further readings

1. P. Carayon, H.A. Schoofs, B.T. Karsh, A.P. Gurses, C.J. Alvarado, M. Smith, et al., Work system design for patient safety: the SEIPS model, *Qual Saf Health Care* **15** Supl 1 (2006), i50-i58.
2. P. Nykanen, J. Brender, J. Talmon, N. de Keizer, M. Rigby, M.C. Beuscart-Zephir, et al., Guideline for good evaluation practice in health informatics (GEP-HI), *Int J Med Inform* **80**(12) (2011), 815-27.
3. J. Talmon, E. Ammenwerth, J. Brender, N. de Keizer, P. Nykanen, M. Rigby, STARE-HI - Statement on reporting of evaluation studies in Health Informatics, *Int J Med Inform* **78**(1) (2009), 1-9.
4. World Health Organization, International Telecommunications Union, *National eHealth Strategy Toolkit*, 2012, https://www.itu.int/pub/D-STR-E_HEALTH.05-2012, last access 11 February 2016.
5. B. Tilahun, F. Fritz, Modeling antecedents of electronic medical record system implementation success in low-resource setting hospitals, *BMC Med Inform Decis Mak* **15** (2015): 61.

Food for thought

1. Within the STARE-HI guideline the health IT system and the context in which the system is implemented needs a more detailed description. Based on section 2 and your own thinking what kind of information is essential to report?
2. How would you deal with the barriers mentioned in section 2 before evaluating the health IT interventions? Would the SEIPS model be appropriate to prepare a resource limited setting before implementation of a health IT intervention?
3. With the increased use of health IT, especially in clinical decision support, is there a risk that clinicians may fully rely on the recommendations of the CDSS, thereby compromising their own judgment? This may be more likely in busy and under-staffed clinics.
4. What are the most important pieces of evidence needed in a LMIC work setting in order to enable introduction of effective health IT support to meet unmet health needs? Is such evidence available, or how might it be obtained?

References

- [1] E. Ammenwerth, N.T. Shaw, Bad health informatics can kill--is evaluation the answer? *Methods Inf Med* **44**(1) (2005), 1-3.
- [2] Institute of Medicine. *Health IT and Patient Safety: Building Safer Systems for Better Care*, 2012, <http://www.nap.edu/catalog/13269/health-it-and-patient-safety-building-safer-systems-for-better>, last access 11 February 2016.
- [3] E. Tomasi, L.A. Facchini, M.F. Maia, Health information technology in primary health care in developing countries: a literature review, *Bull World Health Organ* **82**(11) (2004), 867-74.
- [4] P. Campanella, E. Lovato, C. Marone, L. Fallacara, A. Mancuso, W. Ricciardi, et al., The impact of electronic health records on healthcare quality: a systematic review and meta-analysis, *Eur J Public Health* **26**(1) (2016), 60-4.
- [5] World Health Organization, *Trade, foreign policy, diplomacy and health*, <http://www.who.int/trade/glossary/story021/en>, last access 11 February 2016.

- [6] World Health Organization. *eHealth*. <http://www.who.int/topics/ehealth/en>, last access 11 February 2016.
- [7] M. Vitoria, R. Granich, C.F. Gilks, C. Gunneberg, M. Hosseini, W. Were, et al., The global fight against HIV/AIDS, tuberculosis, and malaria: current status and future perspectives, *Am J Clin Pathol* **131**(6) (2009), 844-8.
- [8] World Health Organization, *Millennium Development Goals*, http://www.who.int/topics/millennium_development_goals/en, last access 11 February 2016.
- [9] World Health Organization, International Telecommunications Union, *eHealth and innovation in women's and children's health: A baseline review*, http://www.who.int/goe/publications/baseline_fullreport/en, last access 11 February 2016.
- [10] The World Bank. Country and Lending Groups, 2014. http://data.worldbank.org/about/country-and-lending-groups#Lower_middle_income, last access 11 February 2016.
- [11] F. Fritz, B. Tilahun, M. Dugas, Success criteria for electronic medical record implementations in low-resource settings: a systematic review, *J Am Med Inform Assoc* **22**(2) (2015), 479-88.
- [12] M. Rigby, Impact of telemedicine must be defined in developing countries, *BMJ* **324**(7328) (2002), 47-8.
- [13] S. Agarwal, H.B. Perry, L.A. Long, A.B. Labrique, Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review, *Trop Med Int Health* **20**(8) (2015), 1003-14.
- [14] S. Khanal, J. Burgon, S. Leonard, M. Griffiths, L.A. Eddowes, Recommendations for the Improved Effectiveness and Reporting of Telemedicine Programs in Developing Countries: Results of a Systematic Literature Review, *Telemed J E Health* **21**(11) (2015), 903-15.
- [15] T. Oluoch, X. Santas, D. Kwaro, M. Were, P. Biondich, C. Bailey, et al., The effect of electronic medical record-based clinical decision support on HIV care in resource-constrained settings: a systematic review, *Int J Med Inform* **81**(10) (2012), e83-e92.
- [16] World Health Organization, *Country Health Information Systems: A review of the current situation and trends*, 2011, http://www.who.int/healthmetrics/news/chis_report.pdf, last access 11 February 2016.
- [17] The Bellagio eHealth Evaluation Group, *Call to Action on Global eHealth Evaluation*, 2011, https://www.ghdonline.org/uploads/The_Bellagio_eHealth_Evaluation_Call_to_Action-Release.docx, last access 11 February 2016.
- [18] R. Wootton, L. Bonnardot, Telemedicine in low-resource settings, *Front Public Health* **3** (2015), 3.
- [19] J.D. Piette, K.C. Lun, L.A. Moura, H.S. Fraser, P.N. Mechael, J. Powell, et al., Impacts of e-health on the outcomes of care in low- and middle-income countries: where do we go from here? *Bull World Health Organ* **90**(5) (2012), 365-72.
- [20] Science Alert. *Google's Internet balloons will soon connect all of Sri Lanka with Wi-Fi*. <http://www.sciencealert.com/google-s-internet-balloons-will-soon-connect-all-of-sri-lanka-with-wi-fi>, last access 11 February 2016.
- [21] World Health Organization, *Task shifting to handle health worker shortage*, 2007, http://www.who.int/healthsystems/task_shifting_booklet.pdf, last access 11 February 2016.
- [22] S. Sood, S. Nwabueze, V. Mbarika, N. Prakash, S. Chatterjee, P. Ray, et al., *Electronic Medical Records: A Review Comparing the Challenges in Developed and Developing Countries*, Proceedings of the 41st Hawaii International Conference on System Sciences, 2008.
- [23] World Health Organization, *Everybody's Business: Strengthening Health Systems to Improve Health Outcomes - WHO's Framework for Action*, 2007, http://www.who.int/healthsystems/strategy/everybodys_business.pdf, last access 11 February 2016.
- [24] World Health Organization, International Telecommunications Union, *National eHealth Strategy Toolkit*, 2012, https://www.itu.int/pub/D-STR-E_HEALTH.05-2012, last access 11 February 2016.
- [25] Ministry of Health, *Kenya National eHealth Strategy: 2011 - 2017*, 2011, http://www.istfth.org/files/media/kenya_national_ehealth_strategy_2011-2017.pdf, last access 11 February 2016.
- [26] South Africa Ministry of Health, *eHealth Strategy - South Africa*, 2012, <http://www.health-e.org.za/wp-content/uploads/2014/08/South-Africa-eHealth-Strategy-2012-2017.pdf>, last access 11 February 2016.
- [27] J.A. Blaya, H.S. Fraser, B. Holt, E-health technologies show promise in developing countries, *Health Aff (Millwood)* **29**(2) (2010), 244-51.
- [28] E. Ammenwerth, S. Graber, G. Herrmann, T. Bürkle, J. König, Evaluation of health information systems-problems and challenges, *Int J Med Inform* **71**(2-3) (2003), 125-35.
- [29] K. Ranson, T.J. Law, S. Bennett, Establishing health systems financing research priorities in developing countries using a participatory methodology, *Soc Sci Med* **70**(12) (2010), 1933-42.
- [30] M. Rigby, Evaluation: 16 powerful reasons why not to do it - and 6 over-riding imperatives, *Stud Health Technol Inform* **84**(Pt 2) (2001), 1198-202.

- [31] K. Kallander, J.K. Tibenderana, O.J. Akpogheneta, D.L. Strachan, Z. Hill, A.H. ten Asbroek, et al., Mobile health (mHealth) approaches and lessons for increased performance and retention of community health workers in low- and middle-income countries: a review, *J Med Internet Res* **15**(1) (2013), e17.
- [32] D. Luna, C. Otero, A. Marcelo, Health Informatics in Developing Countries: Systematic Review of Reviews. Contribution of the IMIA Working Group Health Informatics for Development, *Yearb Med Inform* **8** (2013), 28-33.
- [33] C.S. Hall, E. Fottrell, S. Wilkinson, P. Byass, Assessing the impact of mHealth interventions in low- and middle-income countries what has been shown to work? 2014, <http://www.globalhealthaction.net/index.php/gha/article/view/25606>, last access 11 February 2016.
- [34] D. Zurovac, R.K. Sudoi, W.S. Akhwale, M. Ndiritu, D.H. Hamer, A.K. Rowe, et al., The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: a cluster randomised trial, *Lancet* **378**(9793) (2011), 795-803.
- [35] W. Hersh, A. Margolis, F. Quiros, P. Otero, Building a health informatics workforce in developing countries, *Health Aff (Millwood)* **29**(2) (2010), 274-7.
- [36] M. Rigby, *Globalisation or Localisation: Common Truths or Local Knowledge?* Radcliffe Medical Press, Oxford, 2004.
- [37] J. Mantas, E. Ammenwerth, G. Demiris, A. Hasman, R. Haux, W. Hersh, et al., Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics. First Revision, *Methods Inf Med* **49**(2) (2010), 105-20.
- [38] B. Tilahun, F. Fritz, Modeling antecedents of electronic medical record system implementation success in low-resource setting hospitals, *BMC Med Inform Decis Mak* **15** (2015), 61.
- [39] W.H. DeLone, E.R. McLean, The DeLone and McLean Model of Information Systems Success: A Ten-Year Update, *Journal of Management Information Systems* **19**(4) (2003), 9-30.
- [40] P. Carayon, H.A. Schoofs, B.T. Karsh, A.P. Gurses, C.J. Alvarado, M. Smith, et al., Work system design for patient safety: the SEIPS model, *Qual Saf Health Care* **15** Suppl 1 (2006), i50-i58.
- [41] P. Nykanen, J. Brender, J. Talmon, N. de Keizer, M. Rigby, M.C. Beuscart-Zephir, et al., Guideline for good evaluation practice in health informatics (GEP-HI), *Int J Med Inform* **80**(12) (2011), 815-27.
- [42] J. Talmon, E. Ammenwerth, J. Brender, N. de Keizer, P. Nykanen, M. Rigby, STARE-HI - Statement on reporting of evaluation studies in Health Informatics, *Int J Med Inform* **78**(1) (2009), 1-9.

Learning, Training and Teaching of Health Informatics and its Evidence for Informaticians and Clinical Practice

Evelyn HOVENGA^{a,1}, Heather GRAIN^a

^a*eHealth Education Pty Ltd, Melbourne, Australia*

Abstract. A globally agreed well structured framework representing the health informatics discipline's body of knowledge is yet to emerge. Considerable progress has been made towards describing this over the fifty or so years of the discipline's evolution. This contribution explains the need for such a structured body of knowledge from an educational and workforce capacity building perspective. Some examples of how education and training has been provided to date by a few key stakeholders/leaders are given and critical reviews of guideline and competency developments and their applications are presented. This is followed by an explanation of the need for linking health informatics research with education, learning and training strategies and desired future directions to overcome the identified health workforce knowledge and skills gaps are explored. Given the increasingly important role of health IT in health care, and the significant investment being made into Health IT systems and infrastructure, it is illogical not to seriously invest in health workforce capacity building.

Keywords. Medical Informatics, clinical informatics, competency-based education, Continuing education, professional education, health personnel, health informatics.

1. Introduction

As a professional discipline, health informatics is not well understood. Commonly used terms to describe this discipline are Health Informatics, or Medical Informatics or Biomedical Informatics or eHealth [1-4]. A number of authors have mapped publication trends or undertaken a knowledge domain analysis, or a scoping exercise as ways to define this domain [5-8].

As an emerging scientific discipline in most jurisdictions around the globe, it has been difficult to establish and sustain formal educational programs to suitably prepare the health workforce and improve the health workforce capacity. Amongst the lessons learned is that the health informatics discipline needs to remain cognizant of, and involved in, the aims and activities of health care itself. The benefits of using information and communication technologies to support health service delivery and management, as well as the ability to demonstrate such benefits to others, and avoid compromising patient/client safety, are increasingly becoming compulsory. Significant

¹ Corresponding author: Dr Evelyn J.S Hovenga, eHealth Education Pty Ltd RTO 32279, Melbourne, Australia, e.hovenga@ehe.edu.au.

personal, organisational and national benefits are common expectations following substantial investments in training and education.

2. Learning, Educating and Training in Health Informatics

Learning is ideally student focused. The terms education and training are often used interchangeably although there is a difference in meaning. Education is about acquiring and reinforcing knowledge, whereas training is more focused on applying such knowledge to undertake tasks, and in informatics is usually linked to implementation or use of a specific system. Training is primarily about skill acquisition. Teaching covers both, it refers to the process of facilitating learning to take place. The scope of health informatics teaching covers three different audiences which are distinct, although each audience needs to be aware of the needs and priorities of one another. These groups are:

1. technical staff who develop, choose, implement or maintain systems and who need to know both its applied science and relevant engineering or technical discipline components as well as an overview of user interests;
2. end users (usually health professionals and their support staff) who use health informatics systems (often involuntarily) as part of undertaking their daily care delivery practice; and
3. managerial and policy staff who determine health IT policy and investment, as well as overseeing derived and secondary use of data.

Educational processes are guided by learner, organisational or industry interests, motivation and projected or established workforce knowledge, skill and behavioural needs requirements. This discipline's significant breadth and depth provide numerous educational options. Teaching strategies need to make use of well established educational theories and build on their students' foundational knowledge and skills to be effective. From a vocational perspective, learning outcomes are ideally linked to position or job roles that describe the required performance.

Higher education providers are focused on research and on contributing to the development and progression of a discipline's specific body of knowledge though unfortunately there is often a gulf between ICT research and teaching staff and health research and health professional education. Health informatics research outcomes are always about the computational and informational aspects of medicine and healthcare [9] within the context of any component that is relevant to the health industry. Evidence of the benefits to be achieved from health informatics education makes this attractive to students and other investors, increasing demand and making it financially viable for education and training providers. Unfortunately many are unaware of such benefits. The health workforce generally appears to have little or no appreciation of the need to improve their understanding of the health informatics discipline as evidenced by a common reluctance to address this need. Many stakeholders are unable to differentiate between IT skills and health informatics skills and knowledge, nor do they appreciate the importance of maintaining data safety and integrity or facilitating semantic interoperability, what each is and how they are best achieved. Many health service managers and policy makers do not appreciate the power and potential usefulness of health related information, the many technologies now available [10] and benefits of optimal use, nor indeed the related treatment or organisational risks of

adverse incidents. Nor do they understand the potential savings which could be achieved if health informatics expertise was leveraged in projects and planning. This is evident from the many system failures resulting from poor decisions made regarding acquisition, implementation and staff training/education support [11-12]. Bringing about a change in these perceptions requires the value, benefits and effectiveness of health informatics training and teaching to be demonstrated. A focus on the associated risks and additional costs incurred of not investing in health informatics training and teaching is another way of examining this issue to identify skill requirements.

2.1 Evaluating Benefits of Health Informatics Education, Training and Teaching

Educational evaluation studies tend to focus on graduate employment outcomes or learning effectiveness relative to various delivery methods. There is a dearth of evidence that demonstrates the benefits or return on investment of health informatics education from the perspective of improved health workforce capability and capacity relative to risk management, patient safety, quality of care delivered, patient outcomes, organisational performance effectiveness or efficiency. It is difficult to differentiate between system design, support or care deliverer usage as the cause of good or adverse system impacts as all of these factors interact with each other to process all types of data, information and knowledge. A systematic review of studies undertaken to evaluate the effectiveness of health related information skills training found that the majority of these were undertaken in academic settings rather than in hospital libraries or on practicing clinicians [13]. A literature review that aimed to identify attributes that lead to successful health information systems education and training within the healthcare context, revealed no explicit factors leading to successful health information systems education and training. The educational impact on information system usage was seldom explored or measured [14]. Studies have been undertaken to establish new skills required by the health workforce to enable them to function effectively in this digital age.

An opportunity for staff to acquire basic IT skills resulted in staff saving an average of 38 minutes a day because they were no longer struggling with IT; only 5% of staff who had successfully completed this course now required to call on IT support compared to 71% who did so regularly previously [10]. The acquisition of basic IT skills enabled them to learn to work with technology more quickly and more efficiently [15]. When preparing health information system users to safely manage health data, there needs to be a strong focus on risk management, legal and ethical compliance. A Healthcare unit (NHS Health) was developed by an international expert group convened by the ECDL Foundation and added to this ECDL portfolio in 2007 to meet this need. This study module is independently accredited by the ECDL Foundation, which has an accreditation partner in each country [16-17]. Subsequent studies have resulted in the development and adoption of a Health Informatics Career Framework (HICF) [18].

A similar career based focus was adopted by the Canadian Information and Communications Technology Council (ICTC), a not-for-profit national centre of expertise for the digital economy. A situational analysis of eHealth use was undertaken in a study of Health Informatics workforce requirements. This formed the basis for the development of their eHealth competency profile. Details are not publically available so these could not be compared with the UK career framework. Each profile is stated as consisting of:

- detailed descriptions regarding occupations, key activities and tasks, technical, business and interpersonal competencies where key activities represent desired learning outcomes or the skills and level of competency needed by someone to carry out a role in the workplace [19].
- a number of career clusters made up of work streams that share common competencies.

Career frameworks can be used as a basis for the development of workforce capacity building strategies. As health informatics is playing an ever increasing role supporting the delivery of health services, it is crucial that such use does not compromise the quality of care provided or become a catalyst for errors and adverse events. This issue was explored by the Institute of Medicine [20]. Their report explains the potential benefits and risks of health informatics. This committee found that the information needed for an objective analysis and assessment of the safety of health informatics and its use was not available. It was found that safety is the product of interactions within the larger sociotechnical system.² This includes technology, networks, people, processes, internal and external organisational structures, decisions regarding health informatics acquisition, application and incentives. The committee concluded that safer systems require efforts to be made by all stakeholders. This requires research, training and education of safe practices, including the need to identify measures that relate to the design, implementation, usability, and safe use of computational and informational processes by all users, including patients as well as the potential benefits of adopting new disruptive technologies. Workforce capacity building requires the identification of knowledge and skill requirements as these are used as a foundation from which all educational activities are developed.

2.2 The Health Informatics Domain (Body of Knowledge)

Professions that relate to health informatics, such as software engineers [21], computer scientists, information and communication technologists [22], health information managers [23], clinicians, biomedical scientists, and others representing a number of different professions have each defined their own body of knowledge that describes their specific knowledge and skills domain. Due to the extensive breadth and depth as well as the overlaps between and blurring of the boundaries of a number of these well established knowledge domains, it is difficult to gain consensus regarding a unique body of knowledge for the health informatics domain.

The International Medical Informatics Association (IMIA) has undertaken such a development task that began with a 'think tank' of experts and resulted in the identification of fourteen distinct topics representing a cognitive map of the health informatics discipline. This was followed up by the use of an extensive data extraction method that identified the most commonly used keywords published in the health informatics literature. This was followed by a consensus method to produce a final framework and knowledge base [24-25]. The resulting spreadsheet shows fourteen themes, each with numerous sub-themes, was endorsed by IMIA and complements its educational guidelines.

² See also: B. Kaplan, Evaluation of people and organizational Issues – Sociotechnical ethnographic evaluation, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

Established disciplines are viewed as consisting of 'silos' of knowledge and skills, yet the health informatics body of knowledge needs to be fully integrated within all of these traditional roles as and where appropriate. A Health informatics body of knowledge may be viewed as an umbrella structure that accommodates, respects and calls on specialist contributions as and when required. The only way to overcome the currently perpetuating mismatch of objectives and values is to develop strong linkages via multidisciplinary teamwork.

An invited international group of experts in biomedical informatics and related disciplines agreed that 'biomedical informatics is an interdisciplinary field of study where researchers with different scientific backgrounds alone or in combination carry out research' when reflecting on this discipline, and that it is 'a very broad scientific field and still expanding, yet comprised of a constructive aspect (designing and building systems)' [26]. This focus on 'informatics' relative to all the disciplines concerned with and applied to the health industry in the broadest sense, is what essentially sets the Health informatics domain apart from all others. It is about the applied research and practice of informatics across the clinical, public health, health service management and health policy domains, covering numerous theories, methodologies and technological approaches within human, social, cultural and ethical contexts.

A formally documented body of knowledge is one that permits its use for purposes such as the development and accreditation of academically sound educational courses and programs, certification of specialists or for professional licensing. It sets the standard for professional practice, endorsement and accreditation criteria. It promotes the advancement of both the theory and practice for those who wish to specialise in any aspect contained within this domain. The body of knowledge needs to be underpinned by the scientific foundations for the domain. It is highly desirable to adopt a high level framework that encompasses this continuously changing body of knowledge with a focus on the processing of data, information and knowledge, and the technologies and people interactions used to achieve this within the health industry.

A globally recognized health informatics body of knowledge needs to be described using a structured format, yet it also needs to have sufficient flexibility to enable the inclusion of new knowledge in a timely fashion. Such flexibility is required for the development of innovative educational programs and delivery strategies to meet the educational needs of diverse student cohorts who need to focus on specific specialisations associated with certain roles or disciplines. Its objectives are to:

- Promote a consistent view of the health informatics body of knowledge worldwide, including the core (what needs to be known by the health workforce as well as health informaticians).
- Specify its scope and clarify its place with respect to other related disciplines and bodies of knowledge.
- Be publically accessible.
- Enable the identification of role specific competencies from which position descriptions can be developed and associated essential skills, knowledge and attributes identified to suit the many different types of health care organisation.
- Provide a foundation for health informatics course and curriculum design, development, accreditation and professional development program endorsements.

Many studies have been undertaken to not only describe this domain but primarily as the means to identify new knowledge and skills required by those working with these new technologies [27-29]. Such requirements are commonly expressed as 'competencies'.

2.3 Health Informatics Competency Studies and Frameworks

'Competence' describes the ability of an individual to successfully and/or efficiently perform a set of tasks within a role or function, in accordance with essential and desired requirements. Competency standards define these requirements and may be used as criteria against which learning is measured. Such standards need to specify not only the educational level it applies to, but also the learning topics that collectively constitute the standard. Each topic needs a list of performance criteria (what the student will be able to do in the workplace) or learning outcomes (what the student will have learned as a result), and prerequisite foundational knowledge and skills required to enable successful learning to take place. Ideally it also contains assessment requirements that stipulate the evidence required to demonstrate competence. Educational target groups may be defined in very general terms as:

1. end users - the entire workforce associated with the health industry in some capacity;
2. health informaticians - specialists in any area within the health or health informatics domain;
3. policy makers and policy implementers - decision makers regarding resource acquisition or distribution;
4. ICT professionals who design, develop, implement and maintain systems for the health industry.

Each of these groups and their individual members have very different educational needs depending on the role they need to perform. A consensus regarding commonly occurring role definitions (occupation standards) for any of the above is useful for educators and workforce planners [30]. The recognised need for health informatics capacity building has over many years resulted in numerous studies being undertaken [31-37] for a variety of purposes including specialist applications to suit various clinical specialties [38-40]. These plus technology advances and experiences of the IMIA (International Medical Informatics Association) education working group members who had made use of its guidelines, resulted in a revision and update of the IMIA guidelines on education in biomedical and health informatics in 2010 [41]. Each study has its own focus and purpose.

An AMIA white paper focused on identifying the foundations of biomedical informatics as a scientific discipline and details core competencies for graduate study [42]. A needs assessment for training the biomedical informatics workforce in Latin America was undertaken by Quipu: The Andean Global Health Informatics Research and Training Center, across eleven countries [43]. The online survey questions were provided by local and international experts and included the opportunity to name additional courses. They were sent to 330 medical informatics and biomedical infomratics (MI-BI) related professionals. The results based on 142 surveys received, provided a consensus that the top four courses to be included are the introduction to biomedical informatics, data representation and databases, mobile health and courses

that address issues of security, confidentiality and privacy; a further 28 topics from the health informatics domain were identified as well as ten research priorities.

The Canadian HIP® competency framework, first developed in 2007 and updated in 2012 [44], details a core set of competencies as well as other more specialised competencies categorized according to Health Sciences (Canadian Health system and Clinical and Health services), Information Sciences (Information Technology and Information Management) and Management Sciences (Project Management, Organisational and Behavioural Management, Analysis and Evaluation) topics [45]. These core competencies have formed the basis for a more comprehensive HIP® program; version 3.0 includes a career matrix, role profiles and a credentialing process [46]. The Canadian Association of Schools of Nursing published its set of 'Entry to Practice' Nursing Informatics competencies for Registered Nurses in 2012 [47]. The Royal College of Physicians and Surgeons of Canada [48] has developed a set of recommended eHealth competencies for their members relative to seven roles they may occupy at any time throughout their career path. These have the potential to be applied to any other healthcare delivery related profession.

The US based Technology Informatics Guiding Education Reform (TIGER) initiative focuses on education reform and inter-professional community development to maximize the integration of technology and informatics into seamless practice, education and research resource development [49]. It has published Informatics Competencies for Every Practicing Nurse [50], developed a Virtual Learning Environment available to anyone at minimal cost, and it provides further educationally valuable resources to its international community. Their competencies model is based on basic computer competencies, information literacy and information management for which they recommend the use of existing standards such as the Information Literacy competency standards developed by the American Library Association [51], the Electronic Health Record Functional Model – Clinical Care Components, an ANSI standard developed by Health Level Seven (HL7) [52] and the European Computer Driving License [53]. The ECDL/ICDL Health Supplement module wasn't included in the list of recommended modules to be undertaken despite its successful 2006 US version trial [54].

Work in the UK undertaken by its Council for Health Informatics Professions (UKCHIP) has resulted in a registration scheme for three levels of health informatics professional using standards and an agreed code of conduct [55]. These standards were developed from a number of different sources and previous work. NHS informatics workforce development colleagues in England and Wales have worked together to develop a Career Framework for the Health Informatics profession (HICF) [18] last updated in 2011. Their document provides a diagrammatic representation of a number of other frameworks, including UKCHIP, and how these are linked to the HICF.

The Global Health Workforce Council [56] undertook a major project from a health information management perspective to provide a resource for academic programs across health information professions. This was a global attempt to amalgamate the work of these many and varied projects and to make use of these experiences and findings. Many overlaps between these three health information professional roles used for this study were encountered. Their draft publication is a well written educationally sound document. Its focus did not include clinical and other workforce users. Specialisations were not considered but will be considered for future work. This development work was the result of a transparent, consensus-based process.

Comments subsequently received noted that the work is based on traditional care models and practical experiences from well developed countries (USA, UK, Canada, Australia and elsewhere in Europe) [57]. Other comments received noted the need to identify a basic set of required competencies for all categories making up the health workforce and sets of competencies relative to existing health professional/workforce roles to ensure that all new health professional graduates are suitably work ready. All of these domain topics relate in various ways to overarching critical concepts such as the need to ensure patient safety, maintain confidentiality, data protection, and basic IT use relevant to specific job roles. Educators need to analyse these topics to identify and specify required knowledge, skills and behaviours for their educational programs.

A review of the many published competency statements and associated roles based on skill need studies revealed that required professional competencies in the health informatics domain [58] vary based on the many and varied perspectives and dimensions used to underpin these studies designed for a variety of different purposes, as demonstrated in Table 1. In addition most individual competency statements reviewed consisted of multiple concepts such as topic plus level of responsibility or role context in any one statement.

Table 1. Health Informatics Domain topics used as the primary focus skill and competency development studies selected to demonstrate differences.

Canadian HIP® competency framework – domain topics [44]	Canadian Association of Schools of Nursing – domain topics [47]	T.I.G.E.R – domain topics [49-50]	Global Health Workforce Council – Domain topics [56-57]
Health sciences: • Canadian Health System • Clinical and Health Services	Information and Knowledge Management		Health Informatics
Information sciences: • Information Technology Information Management	Information and Communication Technologies	Information literacy Information management	Health Information management
Managements sciences: • Project Management • Organisational and Behavioural Management • Analysis and Evaluation	Professional and Regulatory Accountability	Basic computer competencies	Health Information and Communication Technologies

2.4 Uses of Health Informatics Competency Frameworks

The AMIA's (American Medical Informatics Association) competency framework and definition of the Clinical Informatics sub-specialty has formed the basis for the American Board of Medical Specialty (ABMS) to create an approved certification process. A number of Clinical Informatics Fellowship Programs are now accredited by the Accreditation Council for Graduate Medical Education [59]. AMIA initiated their 10x10 program in recognition of an increasing need for a larger and better trained workforce in medical informatics. This took the form of an introductory medical informatics course (one subject). It provides a direct pathway to further informatics education. A number of Universities have partnered with AMIA to enable delivery of this course nationally via multiple methods to maximize the impact [60].

Not only do health professionals need to fill a knowledge gap, the same situation applies to ICT professionals working in the health industry. Successful application of health informatics requires knowledge about the business of providing health services.

Whilst there is some recognition (as in the examples described above) of health domain needs, there is little recognition by the IT community of their own knowledge gaps. This was recognized by the Computing Technology Industry Association (CompTIA) [61] who in 1992 introduced vendor-neutral IT certifications recognized globally. Their more recently developed CompTIA Healthcare IT Technician certification covers the knowledge and skills required to implement, deploy, and support healthcare IT systems in U.S clinical settings. An exam guide for this certification was published in 2013 [62].

Despite its development focus to suit the US market, much of the content is applicable to all ICT professionals and others working in or for the health industry. The Health Level 7 organisation also provides a certification service for the use of its standards. It could be argued that our focus for role definitions needs to be on data, their acquisition, secure, effective and timely transmission, and seamless exchange within and between health systems as well as its use [63]. All data processing requires the use of various health information technologies including compliance with data standards. Effective data processing generates knowledge that in turn also needs to be managed in a useful manner. New technologies enabling effective knowledge management continue to be developed and used.

A major challenge encountered during the development of health informatics competency frameworks is that new health informatics roles are emerging and are yet to be clearly defined. A competency framework needs to be able to identify various career path options from job role definitions. Emerging roles also encompass or are closely associated with existing professional roles, thus compounding this challenge. Career paths and educational pathways undertaken by current health informaticians are many and varied.

The 2010 edition of the IMIA recommendations on Education in Biomedical and Health Informatics [42] represents the most recent global framework available as these identify the need to differentiate between desired educational outcomes relative to a variety of job roles. They also meet recognised qualification requirements as these relate to any national educational framework and a range of health informatics positions. They are flexible and not prescriptive. Neither the IMIA framework, nor the IMIA Knowledge base is able to accommodate all of these concepts in a logical and more useful manner such as the Skills Framework for the Information Age (SFIA), a framework of professional skills needed by IT professionals [64]. CompTIA, a group described previously, has mapped its requirements to SFIA. This is one linkage framework identified by the UK's HCIF [18].

2.4.1 The Skills Framework for the Information Age (SFIA)

SFIA's success is demonstrated by its widely accepted global use [65]. The SFIA framework was developed collaboratively and first published by IT professionals and their employers, namely people with real practical experience of skills management in corporate and educational environments, in 2003. It provides a common language, is regularly updated, is now in its 6th edition, and is used in many contexts by educators, human resource managers (employers), professional organisations and individuals for career planning purposes in most countries around the world. It provides a common reference model incorporating unambiguous and clear definitions of IT based technical skills as well professional skills (totaling 96), along with definitions for up to seven

generic levels of attainment detailing autonomy, influence, complexity and business skills role requirements as detailed in table 2.

Table 2. Multiple cross referencing axial topics used in the SFIA Framework Structure [66].

High Level Topic groups	Levels of responsibility	Generic skills defined for each level
• Strategy and architecture – incl. governance, planning, consulting	1. Follow	Autonomy: Has authority and responsibility for all aspects of...
• Business change – incl. staff development, project management	2. Assist	Influence: Makes decisions critical to organizational success...
• Solution development & implementation – incl. socio-technical, data/system integration	3. Apply	
• Service Management – all operational functions	4. Enable	
• Procurement & Management support – incl. supply chain, compliance, risk & quality management	5. Ensure/advise	Complexity: Leads on the formulation....
• Client interface – incl. sales, client support, user interaction	6. Initiate/ influence	Business skills: Has a full range of strategic management and.....
	7. Set strategy, inspire, mobilise	

A mapping of health informatics competencies to SFIA revealed that this framework is not well suited for the health informatics body of knowledge and its applications, although the SFIA logical structure can be replicated. Health informatics requires formal naming and definitions of the concepts and fields represented within its domain together with clear definitions. The SFIA framework structure enables its use as a management tool as well as enabling the identification of suitable codes for the inclusion into a Standard Occupational Classification system. This is useful for the purpose of workforce planning and associated activities.

2.5 Health Informatics Curriculum Development

Educational program curricula ideally are designed according to job roles new graduates are likely to occupy. Learning outcome statements need to be specific, measurable and realistic in terms of a student's ability to successfully acquire the required knowledge, skills and attributes within the educational program's timeframe. Any educational program design needs to be undertaken in a manner that enables the evaluation of the effectiveness of such programs. Assessment guidelines, if available, are useful for this purpose. In summary, the health informatics curriculum development process requires the following factors to be considered:

- Industry/enterprise/workplace contexts and requirements –determine desired outcomes
- Desired training outcome – effects the choice of education/training delivery methods.
- Organisation or workplace goals – determine learning activities to be deployed for student/participant engagement to ensure outcomes reflect workplace readiness.
- Workplace application - determine practical placement and research opportunities

- Participant characteristics – determine the learners' starting points
- Learning styles of the participants – effect how individuals learn best and the choice of activity or educational delivery type and styles educators need to employ.
- Available learning resources or facilities – determine delivery options
- Equipment and consumable resources needed – determine delivery cost
- Topics, their depth and breadth to be covered – determine required resources, time and learning activities.
- Dimensions of competency required – determine educational level to be employed ranging from novice to expert or qualification type to be awarded.
- Qualification type – fits with relevant national education framework level

2.6 Health Informatics Educational Program Delivery

Whilst it is highly desirable to have health informatics content integrated in all preparatory health professional education [10][34][67], it is imperative that members of the health workforce are able to engage in lifelong learning and keep up with new developments [68]. This is particularly relevant due to rapidly changing health informatics advances. Training undertaken to learn how to make use of a new application implemented in the workplace is in itself insufficient.

Clinical informatics is an increasingly influential part of the working environment of all clinical staff [69]. The European Universities that offered early Medical Informatics programs did include clinical informatics for medical students. In Germany this topic became compulsory in 1978 [9] and has remained so. Yet even today the inclusion of clinical informatics is still considered to be a rarity in many countries. Where offered this is usually as an elective or optional course. Attempts have been made to introduce and include the use of applications or medical informatics topics in general as core components of undergraduate medical or other health professional education [70]. Murphy et al. [67] noted that the most important factor holding up progress was the lack of staff with the knowledge and skills to provide academic leadership. This situation may in part be due to a lag in professional development curriculum accreditation requirements [71-73]. Ideally educators have the opportunity to make use of applications, such as electronic health records, as educational tools. Simulated systems could be made use of in skill laboratories to support the development of practical clinical skills.

Educational providers in many countries deliver an increasing number of health informatics programs [74] at various levels of complexity resulting in qualifications ranging from Bachelor degrees to PhDs or equivalent according to the prevailing national qualifications framework. Such formal University, or other Higher Education Providers' educational programs need to be combined with continuing professional educational programs that can be provided on an ad hoc basis for just in time learning, online, in the workplace or via seminars, workshops or via more formal short courses. The delivery methods will vary and may consist of any combination of coursework, online self directed study, practical work experience and research. It may be based on practical experiences, and/or consist of reading, assignment work, discussions, self assessment quizzes, project work and multidisciplinary problem solving activities.

2.7 Health Informatics Educational Program Accreditation

Course or program accreditation refers to a process for approval of a learning program leading to a specified qualification. Accreditation committees representing an accrediting organisation such as national government entities, Universities or Professional organisations make use of the quality assurance standards applicable to the relevant accrediting authority. Education providers need to identify the relevant accrediting authority and obtain their standards and accrediting guidelines to ensure curriculum compliance. The IMIA Education working group has developed such standards together with an accreditation protocol that may be used by health informatics educational providers in the absence of a relevant local accrediting authority [9][75-76]. Such providers are visited by IMIA representatives following the provision of a self assessment report that answers the following six main questions.

1. What are the goals of the program for which the institute asks for accreditation?
2. How are the goals implemented in a curriculum?
3. What is the size and quality of the staff?
4. Which facilities for teaching are available?
5. How does the institute guarantee the quality of the program?
6. Are the goals routinely achieved?

The IMIA accreditation procedure is based on the general higher education procedure in use by the Netherlands and Belgium and was tested on six health informatics programs, including a four year Biomedical Informatics Technologist program provided by a vocational technical educational provider [77]. The writing of the initial self assessment report was found to be beneficial for the management of the program itself as it provided a better insight into the quality of the program submitted for accreditation [76].

2.8 Government Initiatives Impacting on Health Informatics Training and Teaching

Governments have a leadership role to play by enacting legislation, appropriate regulations, including the need for standards compliance, and by providing suitable policy initiatives and funding. Some do this better and more comprehensively than others. From a health informatics education provider perspective it means that curricula need to include such national details. A survey paper found that usable IT systems do improve patient care. It explained the impact of recent regulations and patient safety initiatives (EU, US and Canada) based on findings from human factors usability studies and research that focused on Health Information Technology. [78]. Educators need to make use of such findings when updating their educational programs as they reveal workforce knowledge and skill gaps.

Health professionals, health software vendors and consumers need to be educationally prepared to enable them to effectively participate in the development of solutions to identified challenges encountered when Government, system or organisational initiatives are being implemented. Such initiatives establish new training needs, influence educational program development and may provide new health informatics training and teaching opportunities [79-80]. Most commonly new system implementations simply make provision for system usage skills development of staff.

2.9 Professional Initiatives

Most of the competency framework studies discussed previously were initiated and/or undertaken by professional organisations. Some received Government funding and/or considerable in kind support. Many relied on voluntary academic input. This review has found that Canada [45] and the United Kingdom [18] now have very sophisticated career matrices and defined roles as shown in table 3. These have been used by educational providers to develop and implement new educational programs as well as by employers to effectively deploy the health informatics workforce and by individuals for career planning purposes.

Various possible organisational models were explored to enable IMIA education workgroup members to 'teach globally and learn locally' to overcome the identified dearth of qualified health informatics educators during 1997-2004 [81]. Varied arrangements regarding credit transfers within qualifications, funding arrangements and national educational frameworks were obstacles it was unable to overcome although some student exchange programs are in place. Such desirable collaboration tends to be more achievable nationally or regionally. Web 3.0 now available is capable of transforming the Internet to a 'read, write and collaborative web' with the potential of promoting learning and enabling students and teachers to come closer to 'anytime anywhere' learning [82]. Many streamed health informatics lectures are now also widely available via YouTube and TED Talks. IMIA now has 47 academic institutional members making up its education working group.

Table 3. Professional Health Informatics Role high level comparisons.

Royal College of Physicians and Surgeons of Canada - Roles[48]	UK Health Informatics Career Framework: Roles [18] Total = 84 roles	COACH HIP® Role Profiles Health Informatics Professional Career Matrix: Roles [45, 83] Total = 65 roles
• Medical Expert	Clinical Informatics Staff: 13 different roles at 7 levels of seniority	Clinical & Health Sciences 6 roles at 5 levels of seniority
• Communicator/ • Collaborator	Information Management Staff: 14 different roles at 7 levels of seniority Health Records and Patient Administration Staff: 11 different roles at 6 levels of seniority	Canadian Health System 9 roles at 5 levels of seniority Information Management 13 roles at 5 levels of seniority
• Manager (now Leader)	Project and Programme Management Staff: 12 different roles at 7 levels of seniority	Project Management 6 roles at 5 levels of seniority
• Health advocate	Knowledge Management Staff: 7 different roles at 5 levels of seniority	Organisational and Behavioural Management 10 roles at 5 levels of seniority
• Scholar	HI Educators and Trainers: 10 roles at 6 levels of seniority	Analysis & Evaluation 8 roles at 5 levels of seniority
• Professional	ICT staff: 17 roles at 7 levels of seniority	Information Technology 13 roles at 5 levels of seniority

3 Discussion and Future Directions

Enabling the health workforce to make effective and safe use of available and emerging health informatics technologies and developments is a complex task. Formal recognition of the health informatics discipline, plus an ability of each healthcare organisation to develop their own required health informatics workforce competency

requirements, enables better workforce planning and education strategy developments to build health workforce capacity.

Organisational workforce frameworks enable the development of position descriptions together with an identification of relevant certification requirements. Collectively such frameworks could be used to develop an inventory of possible job roles to assist health informatics educators with the identification of knowledge, professional, technical and behavioural competency needs along with required experience and qualification levels. Individuals may also find this a useful resource for career planning. Adopting a standard approach will assist all of the above activities.

Competency statements derived from the many studies reviewed were found to be inconsistent concerning multiple learning topics reflecting differences regarding discipline specific professional profiles, potential roles and work environments. Adopting an ontological approach for the development of a competency framework enables a better mix and match of concepts for the generation of curricula development to suit well defined graduate outcomes. Such development is expected to provide agreed descriptions of a specific set of knowledge, skills and behaviours that collectively define the health informatics domain as a whole.

The SFIA framework structure provides a useful example for the provision of a flexible resource that can meet the needs of multiple users for various purposes. This differs from the UK and Canadian career matrices as it enables the compilation of unique individual job roles rather than matching to a previously defined job role. This is particularly useful as it enables the identification of any knowledge or skill combination to suit any healthcare organization's workload relative to each function. For example small regional healthcare facilities have a greater need to combine job roles/functions, such as nursing plus informatics, for individual positions. The IMIA Educational guidelines combined with the IMIA knowledge base provides a solid foundation for such a structured framework.

The health informatics domain is constantly changing as we learn more about new technologies and how and why the many current technologies in use fail to or are successful in meeting the needs of organisational or national health service delivery needs [84-85]. Such developments need to be able to be accommodated in the Health Informatics Competency Framework; they also need to be monitored by educators so that their curricula and teaching practices can reflect these changes.

4 Conclusion

The Health Informatics discipline continues to be regarded as an emerging one in numerous locations around the globe. Progress in health informatics education is continuing in a relatively small number of well developed 'western' nations. A number of initiatives are underway to address interdisciplinary conflicts occurring due to the nature of the health informatics knowledge domain, to overcome a dearth of well qualified health informatics educators and to develop the integration of health informatics into more traditional discipline based curricula. Recognition of this the health informatics discipline as a formally identifiable occupational category is slowly being addressed.

For as long as the different groupings of the workforce involved with health informatics remain untrained, systems will not be optimally designed or used, health informatics support will not achieve its optimum role in supporting health care delivery,

and there will be real risks to patients and to data safety and integrity. Given the increasingly important role of health IT in health care, and the significant investment being made into Health IT systems and infrastructure, this is anachronistic and illogical. To date the professional organisations have been active advocates to improve this situation with some success in a small number of countries.

This contribution has focused on various aspects concerning the learning and teaching of health informatics, the knowledge domain itself and the many studies that have been undertaken to identify required competencies. Competency statements need to complement career focused frameworks, and both are required as foundations for all types of educational program development and delivery. It is argued that making use of the globally endorsed SFIA structured framework as a model for developing a similar framework to suit the health informatics knowledge domain based on the IMIA educational guidelines and knowledge base would be beneficial. Once such a standard framework is available it is imperative that it is used not only by educators but also by organisations to establish their own workforce capacity needs profile, by health workforce recruiters who need to demand required skills and knowledge to meet workforce requirements and by individuals for career planning purposes.

Recommended further readings

1. E. Coiera, *Guide to Health Informatics*, 3rd Ed, CRC Press, Taylor Francis Group, Boca Raton 2015.
2. V.K.Saba, K.A McCormick (Eds), *Essentials of Nursing Informatics*, 6th Ed, McGraw-Hill Education.

Food for thought

1. Do you consider the SFIA example as a useful example to be made use of for the development of a health informatics competency framework?
2. Which unique high level health informatics concepts need to be made use of as axis for a multi-axial competency framework?
3. Are you able to identify and list health informatics concepts that need to be described for use in a health informatics competency framework under any of the high level concepts or topics?

References

- [1] E.H. Shortliffe, The science of biomedical computing, *Med Inform* **9** (1984), 185-93.
- [2] J.H. Van Bemmel, The structure of medical informatics, *Med Inform* **9** (1984), 175-80.
- [3] R.A. Greenes, E.H. Shortliffe, Medical Informatics: An Emerging Discipline with Academic and Institutional Perspectives, *Journal of the American Medical Association* **263**(8) (1990), 1114-1120.
- [4] W. Hersh, Medical Informatics – improving health care through information, *Journal of the American Medical Association* **288**(16) (2002), 1955-1958.
- [5] M.J. Schuemie, J.L. Talmon, P.W. Moorman, J.A. Kors, Mapping the domain of medical informatics. *Methods Inf Med* **48**(1) (2009), 76-83.
- [6] Fan Li, Min Li, Peng Guan, Shuang Ma, Lei Cui, Mapping Publication Trends and identifying Hot Spots of Research on Internet Health Information Seeking Behaviour: A Quantitative and Co-Word Biclustering Analysis, *J Med Internet Res* **17**(3) (2015):e81.

- [7] M. Synnestvedt, C. Chen, Visualizing AMIA: A Medical Informatics Knowledge Domain Analysis, *AMIA 2003 Symposium Proceedings*. p.1024.
- [8] C. Pagliari, D. Sloan, P. Gregor, F. Sullivan, D. Detmer, J.P. Kahan, W. Ootwijn, S. MacGillivray, What is eHealth (4): A Scoping Exercise to Map the Field, *J Med Internet Res* **7**(1) (2005), e9.
- [9] A. Hasman, J. Mantas, T. Zarubina, An abridged history of medical informatics education in Europe, *Acta Inform Med* **22**(1) (2014), 25–36.
- [10] J. Murphy, *Gaining support from health disciplines and other stakeholders*, in: E.J.S Hovenga, J. Mantas (Eds) Global Health Informatics Education, *Studies in Health Technology and Informatics* **109** IOS Press, Amsterdam 2004, pp. 36-52.
- [11] J. Doyle, Victorian Auditor-General's Report, *Clinical ICT Systems in the Victorian Public Health Sector*, 30 October 2013, Victorian Government Printer, Melbourne, Australia, http://audit.vic.gov.au/reports_and_publications/latest_reports/2013-14/20131030-clinical-ict-systems.aspx, last access 11 February 2016.
- [12] J. Leviss (Ed), *H.I.T or Miss; Lessons learned from Health Information Technology Implementations*, AHIMA Press, 2010.
- [13] A. Brettle, Evaluating information skills training in health libraries: a systematic review, *Health Information and Libraries Journal* **24** (Suppl1) (2007), 18-37.
- [14] K. Lemmetty, T. Kuusela, K. Saranto, A. Ensio, Education and Training of Health Information Systems – A literature review, in: H.A Park et al. (Eds), Consumer-Centred Computer-Supported Care for Healthy People: Proceedings of NI2006, *Studies in Health Technology and Informatics* **122**, IOS Press, Amsterdam 2006, pp. 176-80.
- [15] NHS Learning & Performance Institute, *European Computer Driving License (ECDL)*, http://www.training.iphis.nhs.uk/New_Site/ecdl/ecdl.shtml, last access 11 February 2016.
- [16] M. Rigby, Protecting the patient by promoting end-user competence in health informatics systems – moves towards a generic health computer user 'driving license', *International Journal of Medical Informatics* **3** (2004), 151-156.
- [17] British Computer Society, *Healthcare added to ECDL portfolio*, *Health Informatics now*, Sept 2007. <http://www.bcs.org/content/conWebDoc/14286> or <http://www.bcs.org/upload/pdf/hinow-sep07.pdf>, last access 11 February 2016.
- [18] UK Health Informatics Career Framework (HICF), <https://www.hicf.org.uk>, last access 11 February 2016.
- [19] Information and Communications Technology Council, *eHealth Competency Profiles*, <http://www.ictc-ctic.ca/what-we-do/programs/standards-and-certification-2/competency-profiles/ehealth-competency-profiles>, last access 11 February 2016.
- [20] Institute of Medicine (IOM), *Health IT and Patient Safety: Building safer systems for better care*, National Academies Press, Washington DC, 2012.
- [21] IEEE, *Guide to Software Engineering Body of Knowledge*, <http://www.computer.org/web/swebok>, last access 11 February 2016.
- [22] European Commission, *Foundational ICT Body of Knowledge*, Feb 2015, http://www.ictbok.eu/images/EU_FoundationalaICTBOK_final.pdf, last access 11 February 2016.
- [23] AHIMA: American Health Information Managers Association, *HIM Body of Knowledge*, <http://bok.ahima.org>, last access 11 February 2016..
- [24] G. Wright, The Development of the IMIA Knowledge Base, *Journal of Information Management* **13**(1) (2011), 45.
- [25] IMIA Endorsed Documents: *IMIA Knowledge Base*, <http://www.imia-medinfo.org/new2/node/8>, last access 11 February 2016.
- [26] A. Hasman, E. Ammenwerth, H. Dickhaus, P. Knaup, C. Lovis, J. Mantas, V. Maojo, F. J. Martin-Sanchez, M. Musen, V.L. Patel, G. Surjan, J.L. Talmon, I.N. Sarkar, Biomedical Informatics – A confluence of disciplines, *Methods Inf Med* **6** (2011), 508-524.
- [27] American Medical Informatics Association (AMIA), *The Science of Informatics*, <https://www.amia.org/about-amia/science-informatics>, last access 11 February 2016.
- [28] R. Haux, Medical informatics: Past, present, future, *Int J Med Inf* **79** (2010), 599-610.
- [29] J. Liaskos, A. Frigas, K. Antypas, D. Zikos, M. Diomidous, J. Mantas, Promoting interprofessional education in health sector within the European International Education Network, *International Journal of Medical Informatics* **78** (Suppl 1) (2009), S43-S47.
- [30] IOM (Institute of Medicine), *Health professions education: A bridge to quality*, Washington DC, The National Academies Press, 2003.
- [31] H.D. Covvey, D. Zitner, R.M. Bernstein (Eds), *Pointing the Way: Competencies and Curricula in Health Informatics*, University of Waterloo, Dalhousie University, University of Ottawa, <http://medicine.dal.ca/content/dam/dalhousie/pdf/faculty/medicine/departments/core-units/DME-medical-informatics/PointingtheWay.pdf>, last access 11 February 2016.

- [32] N. Staggers, C.A. Gassert, C. Curran, Informatics competencies for nurses at four levels of practice, *Journal of Nursing Education* **40**(7) (2001), 303-316.
- [33] *Certified Health Informatician Australasia (CHIA) Hub*, <http://www.healthinformaticscertification.com/index.html>, last access 11 February 2016.
- [34] S.T. Liaw, K. Gray, *Clinical Health Informatics Education for a 21st Century World*, in: E.J.S Hovenga et al. (Eds), *Health informatics, Studies in Health Technology and Informatics* 151, IOS Press, Amsterdam 2010. pp. 479-91.
- [35] Australian Health Informatics Education Council, *Health Informatics Scope, Careers, and Competencies*, Version 1.9, November 2011, http://www.ahiec.org.au/docs/AHIEC_HI_Scope_Careers_and_Competencies_V1-9.pdf, last access 11 February 2016.
- [36] *EU-US Workforce Initiative* in 2014 and their compilation of curricula competencies that also included competencies from the AHIMA-AMIA Matrix Tool; Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics (First Revision); Texas Health Information Technology Higher Education Inventory Report; and Curriculum developed for Health IT training in US, <http://wiki.siframework.org/Workforce+Development+ Work+Group>, last access 11 February 2016.
- [37] *Canadian Health Information Management Association's Learning Outcomes for Health Information Management Diploma/Degree Programs* (2010) and European e-Competence Framework 3.0, <http://www.ecompetences.eu>, last access 11 February 2016.
- [38] Academy of Medical Royal Colleges, *The Scottish Government eHealth Competency Framework*, 2011, http://www.aomrc.org.uk/doc_view/9403-ehealth-competency-framework-defining-the-role-of-the-expert-clinician, last access 11 February 2016.
- [39] A. Barakat, R.D. Woolrych, A. Sixsmith, W.D. Kearns, H.S. Kort, eHealth Technology Competencies for Health Professionals Working in Home Care to Support Older Adults to Age in Place: Outcomes of a Two Day Collaborative Workshop, *Med 2.0* **2**(2) (2013), e10.
- [40] Centers for Disease Control and Prevention and University of Washington's Center for Public Health Informatics, *Competencies for Public Health Informaticians*, Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention. 2009, <http://www.cdc.gov/InformaticsCompetencies>, last access 11 February 2016.
- [41] J. Mantas, E. Ammenwerth, G. Demiris, A. Hasman, R. Haux, W. Hersh, E. Hovenga, K.C. Lun, H. Marin, F. Martin-Sanchez, G. Wright, IMIA Recommendations on Education Task Force. Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics. First Revision, *Methods Inf Med* **49**(2) (2010), 105-120.
- [42] C.A. Kulikowski, E.H. Shortliffe, L.M. Currie, P.L. Elkin, L.E. Hunter, T.R. Johnson, I.J. Kalet, L.A. Lenert, M.A. Musen, J.G. Ozbolt, J.W. Smith, P.C. Tarczy-Hornoch, J.J. Williamson, AMIA Board white paper: definition of biomedical informatics and specification of core competencies for graduate education in the discipline, *J Am Med Inform Assoc* **19**(6) (2012), 931-8.
- [43] M.M. Blas, W.H. Curioso, P.J. Garcia, M. Zimic, C.P. Carcamo, J.M. Castagnetto, A.G. Lescano, D.M. Lopez, Training the Biomedical Informatics Workforce in Latin America; Results of a needs assessment, *British Medical Journal Open* **1**(2) (2011), e000233.
- [44] COACH, *HIP® Core Competencies v.3.0*, <https://www.coachorg.com/en/resourcecentre/resources/Health-Informatics-Core-Competencies.pdf>, last access 11 February 2016.
- [45] COACH, *HIP® Core Competencies 3.0, Career Matrix*, <https://www.coachorg.com/en/professionaldevelopment/HIP-Core-Competencies.asp>, last access 11 February 2016.
- [46] A. Gardner, *HIP® Core Competencies 3.0: Thinking Ahead to Future Expectations for HI Professionals*, COACH, 2012, <https://www.coachorg.com/en/resourcecentre/resources/HIP-Core-Competencies-3.pdf>, last access 11 February 2016.
- [47] Canadian Association of Schools of Nursing, *Nursing Informatics: Entry-to-Practice Competencies for Registered Nurses*, Ottawa 2012, <http://www.casn.ca/2014/12/nursing-informatics-entry-practice-competencies-registered-nurses-2>, last access 11 February 2016.
- [48] K. Ho, R. Ellaway, J. Littleford, R. Hayward, K. Hurley, *The CanMEDS 2015 eHealth Expert Working Group Report*, Ottawa: The Royal College of Physicians and Surgeons of Canada, 2014, http://www.royalcollege.ca/portal/page/portal/rc/common/documents/canmeds/framework/ehealth_ewg_report_e.pdf, last access 11 February 2016.
- [49] *The TIGER Initiative*, <http://www.thetigerinitiative.org>, last access 11 February 2016.
- [50] The TIGER Initiative, *TIGER Informatics Competencies for Every Practicing Nurse*: Recommendations from the TIGER Collaborative, http://tigercompetencies.pbworks.com/f/TICC_Final.pdf, last access 11 February 2016.
- [51] American Library Association, <http://www.ala.org>, last access 11 February 2016.
- [52] ANSI/HL7 EHR, R2-2014 (ISO/HL7 10781), *EHR-System Functional Model, R2*, Health Level Seven (HL7), www.hl7.org, last access 11 February 2016.

- [53] European Computer Driving License, <http://www.ecdl.org>, last access 11 February 2016.
- [54] M.J. Rigby, C. Hulm, D. Detmer, L. Buccoliero. Enabling the Safe and Effective Implementation of Health Informatics Systems - Validating and Rolling Out the ECDL/ICDL Health Supplement, in: K.A. Kuhn, J.R. Warren, T.-Y. Leong (eds.), MEDINFO 2007 - Proceedings of the 12th World Congress on Health (Medical) Informatics, *Studies in Health Technology and Informatics* **129** (2007), IOS Press, Amsterdam, pp. 1347-1351.
- [55] UK Council of Health Informatics Professions, *BCS Standards for Registration*, <http://www.ukchip.org/wp-content/uploads/UKCHIP-Registration-Standards-v1.0.pdf>, last access 11 February 2016.
- [56] Global Health Workforce Council, *Global Academic Curricula Competencies for Health Information Professionals*, draft for public comment November 2014, <http://ifhima.org/global-health-information-curricula-competencies>, last access 11 February 2016.
- [57] Global Health Workforce Council, *Response to public comments*, received by 15 Jan 2015, <http://www.ahima.org/~media/AHIMA/Files/AHIMA-and-Our-Work/FinalConsolidatedResponses4415.ashx?la=en>, last access 11 February 2016.
- [58] J. Kampov-Polevoi, B.M. Hemminger, A curricula-based comparison of biomedical and health informatics programs in the USA, *Journal of the American Medical Informatics Association* **18**(2) (2011), 195-202.
- [59] American Medical Informatics Association (AMIA), *Education; History of the Clinical Informatics Subspecialty*, <https://www.amia.org/education>, last access 11 February 2016.
- [60] W. Hersh, J. Williamson, Educating 10,000 informaticians by 2010: the AMIA 10x10 program, *International Journal of Medical Informatics* **6** (2007), 37-382.
- [61] Computing Technology Industry Association (CompTIA), *Healthcare IT Technician*, <http://certification.comptia.org/getCertified/certifications/hittech.aspx>, last access 11 February 2016.
- [62] K.A. McCormick, B. Gugerty, *Healthcare Information Technology Exam Guide for CompTIA® Healthcare IT Technician and HIT Pro™ Certifications*, McGraw Hill, New York.
- [63] E.J.S Hovenga, *Health workforce competencies needed for digital world*, in: E.J.S Hovenga, H. Grain (Eds), *Health Information Governance in a Digital Environment*, *Studies in Health Technology and Informatics* 193, IOS Press, Amsterdam, 2013.
- [64] SFIA Foundation, <http://www.sfia-online.org/en/about-us>, last access 11 February 2016.
- [65] SFIA – Skills for the Information Age, <http://www.sfia-online.org/en>, last access 11 February 2016.
- [66] SFIA Framework Reference Guide, <http://www.sfia-online.org/en/get-sfia>, last access 11 February 2016.
- [67] J. Murphy, K. Stramer, S. Clamp, P. Grubb, J. Gosland, S. Davis, Health Informatics education for clinicians and managers – What's holding up progress? *International Journal of Medical Informatics* **3**(2) (2004), 205-213.
- [68] J.C. Wyatt, F. Sullivan, Keeping up: learning in the workplace, *British Medical Journal* **331**(7525) (2005), 1129-1132.
- [69] K. Gray, J. Sim, Factors in the development of clinical informatics competence in early career health sciences professionals in Australia: a qualitative study, *Advances in Health Sciences Education* **16**(1) (2010), 31-46.
- [70] A. Otto, A. Kushniruk, Incorporation of Medical Informatics and Information Technology as Core Components of Undergraduate Medical Education – Time for Change! In: J.G. McDaniel (Ed), *Advances in Information Technology and Communication in Health*, *Studies in Health Technology and Informatics* **143** (2009), IOS Press, Amsterdam.
- [71] P. Graham-Jones, S.H. Jain, C.P. Friedman, L. Marcotte, D. Blumenthal, The need to incorporate health information technology into physicians' education and professional development, *Health Affairs March* **31**(3) (2012), 41-487.
- [72] C.R. Keenan, H.H. Nguyen, M. Srinivasan, Electronic medical records and their impact on resident and medical student education, *Acad Psychiatry* **30**(6) (2006), 522-7.
- [73] K. Gray, A. Dattakumar, A. Maeder, H. Chenery, Educating future clinicians about Clinical Informatics: A review of implementation and evaluation cases, *European Journal of Biomedical Informatics* **2** (2011), en48-en57.
- [74] J.M. Brittain, A.C. Norris, Delivery of health informatics education and training, *Health Libraries Review* **17** (2000), 11-128.
- [75] A. Hasman, E. Ammenwerth, H. Dickhaus, P. Knaup, C. Lovis, J. Mantas, V. Maojo, F.J. Martin-Sanchez, M. Musen, V.L. Patel, G. Surjan, J.L. Talmon, I.N. Sarkar, Biomedical Informatics – A confluence of disciplines, *Methods Inf Med* **6** (2011), 508-524.
- [76] A. Hasman, J. Mantas, IMIA Accreditation of Health Informatics programs, *Healthc Inform Res* **19**(3) (2013), 154-161.

- [77] DuocUC Professional Institute, School of Health, Biomedical Informatics Technologist Program, personal communication with Jorge A Gonzalez, Director, August 2015, see also <http://www.duoc.cl/carrera/informatica-biomedica>, last access 11 February 2016.
- [78] K. Lemmetty, T. Kuusela, K. Saranto, A. Ensio, *Education and Training of Health Information Systems – A literature review*, in: H.A Park et al. (Eds), Consumer-Centred Computer-Supported Care for Healthy People: Proceedings of NI2006 Studies in Health Technology and Informatics 122, IOS Press, Amsterdam, 2006.
- [79] D. Blumenthal, Launching HITECH, *New England Journal of Medicine* **362** (2010), 382-385.
- [80] Federal Register No 171, Department of Health and Human Services, Office of the secretary, 45 CFR Part 10 Health Information Technology: Standards, Implementation Specifications, and Certification Criteria for Electronic Health Record Technology, 2014, Revisions to the Permanent Certification Program for Health Information Technology, <http://www.gpo.gov/fdsys/pkg/FR-2012-09-04/pdf/2012-20982.pdf>, last access 11 February 2016.
- [81] E.J.S.Hovenga, *Virtual University Governance*, in: E.J.S.Hovenga, J.Mantas Global Health Informatics Education, Studies in Health Technology and Informatics 109, IOSPress, Amsterdam, 2004.
- [82] P. Otero, W. Hersh, Education in Biomedical and Health Informatics in the Web 3.0 Era: Standards for Data, Curricula, and Activities, in: IMIA Yearbook of Medical Informatics **6** (2011), 139-41.
- [83] COACH: HIP® 2013 Role Profiles, <http://www.coachorg.com/en/resourcecentre/hip-role-profiles.asp>, last access 11 February 2016.
- [84] *openEHR Foundation*, <http://www.openehr.org/programs/specification>, last access 11 February 2016.
- [85] L. Michelsen, S.S. Pedersen, H.B. Tilma, S.K. Andersen, *Comparing different approaches to two-level modelling of electronic health records*, in: R. Engelbrecht et al. (Eds), Proceedings of MIE2005, Studies in Health Technology and Informatics **116** (2005), 113-8.

Forward Outlook: The Need for Evidence and for Action in Health Informatics

Michael RIGBY^{a,1}, Elske AMMENWERTH^b, Jan TALMON^c

^a*Keele University, Keele, Staffordshire, ST5 5BG, U.K.*

^b*UMIT – University for Health Sciences, Medical Informatics, Hall in Tirol, Austria*

^c*Hi-Way, Health Informatics Advisor, St. Odiliënberg, The Netherlands*

Abstract. Evidence-based health informatics (EBHI) is a key concept in the development and deployment of IT systems and applications in an ethical, efficient and effective health system. There is an ever growing body of knowledge to guide IT-related decision making, but further growth of this body of knowledge is required as the health IT domain and technologies are continuously evolving, leading to new functionalities and applications. However, EBHI should not wait until retrospective evidence is available – increasingly policy makers should draw on available prior and external evidence to influence design and development processes so as to ensure that health IT is devised in response to a delivery process need and not as an IT driven goal. EBHI should thereby seek to move forward significantly the metric that only some 20% of IT investment is truly successful.

Keywords. Evidence, evaluation, policy, medical informatics.

1. Introduction

The comprehensive book on "Evidence Based Health Informatics" [1] has presented a case, and related knowledge and evidence, for a radical change in the attitude to acquisition and use of health information technology (health IT) applications, from one of marketing-based promotion and optimistic investment, to one of evidence-based policy grounded firmly in evaluation science. This is logical on two fundamental grounds: first, that any intervention that can have an effect on patient care should be based on scientific principles; and second, that patient safety and optimal use of healthcare resources are paramount.

The half-century honeymoon period of health informatics in which it considered itself aloof from such universal core values present elsewhere in the health sector, should cease. This is not an invented problem. Not only has it been demonstrated that health informatics systems can be harmful and even can kill [2,3], but the optimistic claims of at least some sector proponents [4], and lack of precision of at least some vendors' claims [5], have also been identified.

¹ Corresponding author: Professor Michael Rigby, Lavender Hill, 6 Carrighill Lower, Calverstown, Kilcullen, Co. Kildare, Ireland, m.j.rigby@keele.ac.uk.

2. The Evidence Contributed by this Book

The book [1] has provided state of the art evidence from contributors around the globe on all the relevant key issues related to current mature health IT applications and use. This has been done through three lenses.

Part I established the context of evidence-based health informatics (EBHI) and defined its importance for effective and safe healthcare from three complementary viewpoints. First, the editors introduced the central importance of the evidence-based and scientific approach, which – it is argued – should now be seen as an ethical as well as a practical imperative [6]. Wyatt followed this with an appraisal of scientific developments in this field of evidence and evaluation [7], while Magrabi et al addressed the key issues of health IT for patient safety and argues for improving the safety of health IT [8].

Part II then presented nineteen individual chapters on specific methodological topics on the necessary techniques of health IT evaluation, starting with Brender's overall theoretical basis of health IT evaluation [9], then covering a wide range of issues such as stakeholder perspectives in evaluation, study designs, mixed methods, simulations, ethnographic evaluation, evidence-based usability, economic evaluation, evaluation of health IT implementation, health technology assessment, systematic reviews and meta-narrative reviews.

Finally, but just as important, Part III covered how to ensure the relevance and application of evidence including assessment of the quality of studies, application in low income countries, and training and education.

3. The Need for Continuous New Methodologies and Evidence

However, while the book seeks to present a comprehensive set of current research-based guidance from across the globe, this is only a step in the journey. The health care sector, the health informatics discipline, and the enabling technologies will all continue to evolve. Health informatics evaluation, health informatics evidence, and thus the framing and execution of policy to ensure safe, efficient and effective use, will of necessity have to evolve continuously too.

3.1. Proactive Translational Approaches

Health IT systems, whether commercial, bespoke in response to tender, or built in house, will continue to evolve. Indeed, a logical consequence of evaluation is that weaknesses should be resolved, while developing technology will give opportunity for enhanced functionality. At the same time increased user experience will lead to innovation and extension of use. This means that as a system is commissioned or develops it may well lead the field in its own specific context, but this does not mean that there is no relevant evidence which could and should inform its safe and effective design and implementation.

Instead, it leads to the need to use intuition. While any health system situation can claim to be unique, this is no excuse for failing to look for relevant evidence which can be provide valuable lessons which can be projected across. Insight can be gained from many external analyses. Not finding directly comparable evidence from identical

situations is no excuse for ignoring strong messages from elsewhere, or for failing even to look for such external but apposite evidence.

Evidence from other services, other health settings, and indeed from outside the health sector, should be scrutinized and assessed as to how its lessons apply. This is the translational skill – taking evidence from one context and translating it to another setting without loss or corruption of meaning. Not only should this be more cost-effective than making avoidable mistakes, but it is safer, and more respectful towards the ultimate users.

3.2. Product and Application Evolution

Updates or improvements to any system are welcome, but there are countervailing risks that some established good effects will be compromised, or that legitimate benchmarks will change, especially with increasing complexity. This generates the difficult conundrum as to what degree evaluation results can be equally applicable to later variants of an application. There comes a point at which the next generation of an application has to be considered to be a new entity, but in the interim it is necessary to be selective and discerning, as such change to applications is often incremental rather than being a major step change.

This consideration of assessing the continuing validity of evidence from one version's evaluation to evolved versions is not unique in health care. It applies to medical devices which get accreditation, but which may then be updated and do not require reaccreditation under current rules if the manufacturer decrees that they are substantially the same but improved. However, critics argue that some improvements may involve for instance changed materials with allegations of iatrogenic effects for some patients. Similarly a generic drug may be approved on the basis of its active ingredients being identical to an earlier proprietary version, but there may be claims that changed filler compounds or edible colours may have adverse effects which are missed by the lack of new trials.

This issue of when an update is a variant with new effects, as opposed to being simply an improvement, is one of those subjects needing further research.

3.3. New Types of Application

The world of health IT is not standing still – nor will it ever do so. New technologies are ever arriving, ranging from new kinds of social media and Internet sites and repositories to wearable devices like activity trackers, and to the Internet of Things. New technologies will include new access and data input (and output) methods, but not limited to various forms of biometric recognition, and reliable voice input.

The boundary between formal and personal information systems will blur. For instance, a Bring Your Own Device (BYOD) policy will enable professional staff to embed their work IT into their personal time- and process-management communication and data systems, but probably without an analysis of whether there is any impairment or corruption of function on a different platform. Even web-based access to applications may not be a safe approach as a (generated) web-page may not render the same on different browsers.

Conversely patients may be lean systems or enabled to embed provider health software and applications into their own systems, and personal and formal data may be exchanged or co-recorded using patient portals. In all these respects new evidence will

be needed, based on new paradigms of investigation, and on new metrics of process, resource use, and outcome [10].

In all this, the need to avoid new health inequalities emerging will not diminish, and equity must be monitored. However, the most important thing will be the constant moving forward of evaluation methodologies and evidence repositories, coupled with recognition that the evidence-basis is constantly evolving at the frontiers.

In 2011 the IMIA Working Group on Technology Assessment discussed twice the next areas for development of the techniques of evaluation [11,12]. This was in the context of increasingly dispersed health IT systems, cross-border care delivery, and health care systems where end-users as agents of co-production of information are integrated within the formal care delivery system. Telemedicine also gives new opportunities and challenges, not least in measuring quality, access and user views when the system is dispersed beyond institutional boundaries.

Finally, a whole new dimension is opening up in health care with the development of mobile technologies, the mHealth scene, supporting increasingly the provision of vital sign and health function analysis either as part of the health care provision or as consumer oriented apps. How those mHealth applications affect the health care system in terms of effectiveness and efficiency as well as their safety for their users is still poorly understood [13]. The explosive expansion of the domain of health ‘apps’ is just one dimension, with exponential increase in availability and little evaluation of the value of integrating them into health care delivery, with the WHO European Regional Office reporting that only a quarter of countries in the European region regulate mobile health devices and software for quality, safety and reliability [14], and that “73% of [European] States (33 countries) do not have an entity that is responsible for the regulatory oversight of mobile health apps for quality, safety and reliability, despite widespread use of such technology. This presents a potential risk for countries and is an area in need of incentives, guidance and oversight” [15].

3.4. Enhanced User Expectations

A further dimension of the need for continuous evolution of evaluation methods, and of the underlying values and expectations, comes out of the ever-growing sophistication and expectation of citizens as users of IT in general and out of their increasing expectations for ‘joined up’ health care delivery. Arguably there are two dimensions to this. One is that increasingly citizens are becoming more empowered and emboldened as consumers of healthcare. The attitudes of previous centuries, of the submissive patient grateful for any service, are passing or indeed have passed. Instead, and fueled by increasingly sophisticated consumer programmes in commerce, patients have high expectations of service quality and delivery. This extends into expectations of personalization and of service access, and thus of the effectiveness of IT in delivering this.

Linked to this is the increase in citizen education. This increase in literacy, and in health literacy, means that expectations of the use of IT to deliver services, and as a bridge between the health system and the patient, have increased. Such IT could provide services like making appointments and email or chat functionality to consult health care providers, but also access to a consumer-tailored version of the electronic health record (EHR), and opportunities to keep a health diary on-line as part of the care process; and increasingly to be interactive in their care management using consumer-

held informatics embedded into formal services, as early initiatives are rolled out to mainstream services, e.g. [16,17,18].

However, the reverse side of this is also important and should be the subject of evaluation and evidence – such health IT applications will become more central to care delivery and should be accessible to all citizens. Access to consumer IT, and IT literacy, should however not be prerequisites of healthcare access and appropriate education and support are provided for those who do not have them and even then alternative routes for healthcare access should be provided. Evidence on the expectations of patients and citizens and on the impact of these patient-centred tools on health care delivery is needed and will help to shape the future organization of health care delivery.

4. Not Waiting for Adverse Outcomes - Anticipatory Use of Evidence

One of the misunderstandings – and resultant risks of evidence-based work and its relationship to evaluation – is the tendency to wait for evaluations to show weaknesses as well as successes, and then understandably to highlight the difficulty of making change to an established operational system – albeit one with faults. This may be a necessary situation in the early stages of a new domain, but one which can and should be avoided once a discipline is mature – as is the case in health informatics and the applied study of health IT.

In the early stages of the domain of health informatics, systems and applications have been developed and put in operation under the assumption that this would improve the efficiency and quality of health care. Only after the implementation did any evaluation of the impact take place – and often evaluations were not performed. This has lead to unintended consequences. Since the health informatics discipline has matured over time such strategies are not acceptable anymore, yet still too often we can see politically driven aspirational solutions, adoption of applications from one setting into another setting without realization of the effect of context on function and use, or – as is often the case in developing countries – donor driven systems being introduced devoid of understanding of the local needs in a recipient community. Current developments should be based on existing evidence and experience. Methods should be developed to deal with gaps in prior knowledge so as to anticipate and counteract any mismatch between intended use and outcome and the actual application of health IT.

What is necessary, and appropriate, is anticipatory use of preceding knowledge. There is already enough knowledge about health informatics systems, and indeed about large IT projects more generally, to be able to act more intelligently than is often the case at the design and commissioning stages, using prior generic knowledge.

Clegg and Shepherd made these points graphically in a paper in 2007, when they correctly forecast the demise of the English NHS's National Programme for IT (NPfIT) [19]. Writing as generic organizational psychologists and IT experts rather than as health IT specialists, they drew widely from generic enterprise IT experience, and from a joint UK Royal Academy of Engineering and British Computer Society report which identified that 'a mere 16% of IT projects can be considered successful' [20], and they cited several cases of unsuccessful systems implementations. They identified that 'IT push' systems have the poorest track record, and that 'user pull' is far more successful – which is precisely the opposite of the NPfIT objective or philosophy. Their predictions were uncannily accurate.

Clegg and Shepherd came to the conclusion that around 40% of large IT systems investments are complete failures, 40% partial successes, and only up to 20% can be considered successes. Yet in the health sector too often this evidence is ignored – because extrapolation of external prior evidence, and non-health evidence, is not considered. Their analysis from wide evidence is that IT systems not only necessitate process change, but will only work if the stimulus is for positive change in delivery process and that change is of a type which can only be engendered with IT support. They also argue that ownership, and metrics of success, should be bedded in achievement of the new service processes and not with an IT department and implementation criteria.

So not only should the health sector be looking outside at such examples of failure, and more importantly at examples of success, but also there should be far more study of prior work in health IT, or specific eHealth domains, and analysis of successes and failures. The external prior evidence should be used rather than naively believing that design intentions will work efficiently in practice without deeper health process and business psychology evidence. Translation science, and then implementation science, should be given their rightful places.

Waiting for the evidence should not be an acceptable process. What is appropriate is a three phase approach – finding and translating prior external evidence; formative local evaluation as design, selection, purchase and installation progresses; and summative evaluation as the application matures. All are evidence-based, but the source and nature of the evidence changes as the process proceeds.

5. Ongoing Collaborative Work

The work on promoting Evidence-based Health Informatics has been progressed not simply by individual academics and experts, but in particular by the International Medical Informatics Association (IMIA) and its Working Group on Technology Assessment, together with the European Federation of Medical Informatics (EFMI) Working Group on Assessment of Health Information Systems and work of the Working Group on Evaluation of the American Medical Informatics Association (AMIA). The US Agency for Healthcare Research and Quality's National Resource Center for Health Information Technology has also been a contributor to the domain.

This collaboration should continue, but should also be extended as this is more an academic or scientific issue. In particular it is important to seek synergy with the work of the World Health Organisation into informatics and eHealth, and to link with the health quality agenda as to be able to translate the existing evidence to places where it can have the biggest impact. Given the focus on evidence, collaboration with the Cochrane Collaboration should be seen as a further goal.

6. Developed Education

Evidence-based health informatics will only solidly be integrated in health IT practice when health IT professionals as well as health IT decision-makers have sufficient background knowledge and skills to understand and apply evidence and to design and perform health IT evaluation studies. In particular, health IT professionals need to have strong skills in searching, assessing, understanding and applying available evidence.

These skills are related to basic scientific skills and need to be part of health IT education already on the undergraduate level.

Health IT professionals who are responsible for health IT evaluations need additional skills in designing, conducting and analyzing evaluation studies. Recommendations on how to integrate these skills in health informatics curricula are being discussed at the time of writing by the IMIA and EFMI working groups on health IT evaluation. Also, health care managers and health policy makers need to understand the benefit of evidence-based decisions, thus content on collection and appreciation of evidence should be included in health management and related programs.

To reach all these different groups, health informatics organizations such as IMIA, EFMI or AMIA should consider developing online courses on evidence-based health informatics, comprising well-defined core content and additional specialized content depending on the target audience.

7. Amassing and Accessing the Evidence

A further issue which would benefit from cohesive global action is the collation of evidence into one location. Some moves have been made, but these are far less than the subject deserves, and far less than the making available of impartial validated evidence on clinical topics.

The University for Health Sciences, Medical Informatics (UMIT) in Austria, in conjunction with Amsterdam Medical Centre developed and makes available online a searchable database of Health IT Evaluations, updated periodically [21]. Being a database of evaluations it has a strong scientific underpinning. It covers evaluations rather than the lessons learned derived from those evaluation studies. It has been updated systematically since 2003 and now contains also systematic reviews and researcher submissions, and has over 1,800 entries.

The bibliographic databases are another rich source of material, but are fragmented and require the searcher to understand the different bibliographies both as to content and as to structure and access. PubMed contains the health-based literature, but not pure computing literature – such as the paper by Clegg and Shepherd cited above. Literature in the nursing press – and nursing is the largest profession with most direct patient contact – together with that from the allied health professions appears in the Cumulative Index to Nursing and Allied Health Literature (CINAHL) [22]. Social science and commerce bibliographies may also hold key material. So the very subject which seeks to be the uniting influence of healthcare itself has its literature fragmented. This too seems to indicate a need for action.

Finally, but probably most significantly, the WHO Global eHealth Evaluation Meeting held in Bellagio in September 2011 developed a consensus statement entitled ‘Call to Action on Global eHealth Evaluation’. The WHO-assembled expert participants concluded with a call to “Create a multi-stakeholder web-based platform for constructive sharing, publication and learning from successes and failures. Include a registry of eHealth evaluation studies and results, and a repository of evidence-based eHealth best principles and practices” [23].

Yet despite the expert and intellectual power of the WHO, the global reach of IMIA and the resources of several health IT interested global NGOs including the one that sponsored the Bellagio meeting, and the huge resources including promotional bodies of the health IT industry, this fundamental evidence-collating action has not

been progressed in the intervening years. Given the potential health gains as well as economic gains potentially available to be unlocked, this seems to represent a palpable failure of the responsible commercial and professional organisations globally. Without such a global, integrated resource of experience, there is a high risk that the huge waste of resources of imperfect implementations will continue.

8. Conclusion – A Beginning not the End

The book [1] has sought to promote the concept of evidence-based health informatics, its ethical and practical importance to healthcare and thus to patients, and the sources of evidence to support this. It has introduced a wide range of topics, concepts and experience. But, like so many scientific endeavors, this is not the end of the road – at best it is the end of the beginning.

While there is a lot of core material in this volume which should stand the test of time, the world is moving forward, as are the technologies, the healthcare processes and consumer expectations. Therefore, the processes of evaluation, of evidence gathering, and of application of that knowledge need to evolve continually; matched to new needs and new techniques. This may be the beginning of the road, and it is hoped that others will be motivated to continue mapping and navigating the journey. Moreover, since knowledge of Health IT failures is clear and the benefits are under-realized, the Precautionary Principle is still European Commission policy and commended elsewhere [24,25]. The importance of business ethics is increasingly recognized [26, 27] in addition to healthcare ethics [28], it seems imperative that global action is to be taken to assemble as much evidence as possible into one location, readily accessible to scientists, policy makers, developers, end users, and representatives of the public whose health is to benefit from proper health IT. It would be most encouraging if IMIA, the WHO, the Cochrane Collaboration, and other key bodies could come together to effect this modest cost global initiative for the benefit of the health and health systems of the global population.

References

- [1] E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics - Promoting Safety and Efficiency through Scientific Methods and Ethical Policy*, Stud Health Technol Inform 122, IOS Press, Amsterdam, 2016.
- [2] E. Ammenwerth, N.T. Shaw, Bad health informatics can kill - is evaluation the answer? *Methods Inf Med* **44**(1) (2005), 1-3.
- [3] Web list “Bad Health Informatics can Kill” (accessible via the Bad Health Informatics link at <https://iig.umin.at/efmi/index.htm>) (accessed 12 February 2016).
- [4] A.L. Kellermann, S.S. Jones, What it will take to achieve the as-yet-unfulfilled promises of health information technology, *Health Affairs* **32**(1) (2013), 63-68.
- [5] N.K. Yeung, A.R. Jadad, A. Shachak, What do electronic health record vendors reveal about their products: an analysis of vendor websites, *J Med Internet Res* **15**(2) (2013), e36.
- [6] M. Rigby, E. Ammenwerth, *The need for evidence in health informatics*, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 122, IOS Press, Amsterdam, 2016.
- [7] J.C. Wyatt, *Evidence-based health informatics and the scientific development of the field*, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 122, IOS Press, Amsterdam, 2016.
- [8] F. Magrabi, M.-S. Ong, E. Coiera, *Health IT for patient safety and improving the safety of health IT*, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 122, IOS Press, Amsterdam, 2016.

- [9] J. Brender, *Theoretical basis of health IT evaluation*, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 122, IOS Press, Amsterdam, 2016.
- [10] M. Rigby, A. Georgiou, H. Hyppönen, E. Ammenwerth, N de Keizer, F. Magrabi, P. Scott. *Patient Portals as a Means of Information and Communication Technology Support to Patient-Centric Care Coordination - the Missing Evidence and the Challenges of Evaluation*. A joint contribution of IMIA WG EVAL and EFMI WG EVAL, in: A. Geissbuhler, C.A. Kulikowski (editors), IMIA Yearbook of Medical Informatics 2015, Schattauer, Stuttgart, pp. 148-59.
- [11] M. Rigby, E. Ammenwerth, J. Talmon, P. Nykänen, J. Brender, N de Keizer, *Health Informatics 3.0 and other Increasingly Dispersed Technologies Require Even Greater Trust: Promoting Safe Evidence-based Health Informatics*, in: A. Geissbuhler, C.A. Kulikowski (editors), IMIA Yearbook of Medical Informatics 2011: Towards Health Informatics 3.0, Schattauer, Stuttgart, 2011, pp. 105-111.
- [12] M. Rigby, J. Brender, M.-C. Beuscart-Zephir, H. Hypponen, P. Nykänen, J. Talmon, N. de Keizer, E. Ammenwerth, *Next Steps in Evaluation and Evidence – from Generic to Context-Related*, in A. Moen, S.K. Andersen, J. Aarts, P. Hurlen (Eds.), User Centred Networked Health Care – Proceedings of MIE 2011, IOS Press, 2011, pp. 208-212.
- [13] B. Vallespin, J. Cornet, A. Kotzeva, Ensuring evidence-based safe and effective mHealth applications, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 122, IOS Press, Amsterdam, 2016.
- [14] World Health Organisation Regional Office for Europe. *From innovation to Implementation – ehealth in the WHO European Region*; WHO, Copenhagen, 2016.
- [15] World Health Organisation Regional Office for Europe Media Centre. *E-health – when, not if*, WHO, Copenhagen, <http://www.euro.who.int/en/media-centre/sections/press-releases/2016/03/e-health-when,-not-if> (accessed 14 March 2016).
- [16] B. Parmanto, G. Pramana, D.X. Yu, A.D. Fairman, B.E. Dicianno, M.P. McCue, iMHere: A Novel mHealth System for Supporting Self-Care in Management of Complex and Chronic Conditions, *JMIR Mhealth Uhealth* 1(2) (2013), e10.
- [17] S. Marceglia, P. Fontelo, E. Rossi, M.J. Ackerman, A Standards-Based Architecture Proposal for Integrating Patient mHealth Apps to Electronic Health Record Systems, *Appl Clin Inform* 6(3) (2015), 488-505.
- [18] *United4Health*, <http://www.united4health.eu> (accessed 16 March 2015).
- [19] C. Clegg, C. Shepherd, 'The Biggest Computer Programme in the World .. Ever!': time for a change of mindset?, *Journal of Information Technology* 22 (2007), 212-221.
- [20] Royal Academy of Engineering and British Computer Society. *The Challenge of Complex IT Projects* – The report of a working group from Royal Academy of Engineering and British Computer Society, London, 2004.
- [21] *Health IT Evaluation Inventory*, <https://evaldb.umit.at/> (accessed 15 February 2016).
- [22] *CINAHL Database*, <https://health.ebsco.com/products/the-cinahl-database> (accessed 15 February 2016).
- [23] Call to action on global eHealth evaluation: consensus statement of the WHO Global eHealth Evaluation Meeting, Bellagio, September 2011, in: *eHealth evaluation evidence*, The Bellagio eHealth Evaluation Group, 2011, http://www.ghdonline.org/uploads/The_Bellagio_eHealth_Evaluation_Call_to_Action-Release.docx (accessed 15 February 2016).
- [24] European Commission, Communication from the Commission on the Precautionary Principle, COM(2000) 1, Brussels, 2000.
- [25] E Fisher, J Jones, R von Schomberg (eds), *Implementing the Precautionary Principle: Perspectives and Prospects*, Edward Elgar, Cheltenham, UK and Northampton, MA, 2006.
- [26] Business Ethics, *The Magazine of Corporate Responsibility*, <http://business-ethics.com> (accessed 15 February 2016).
- [27] *Institute of Business Ethics*, <http://www.ibe.org.uk> (accessed 15 February 2016).
- [28] T. Beauchamp, J.F. Childress, *Principles of Biomedical Ethics*, 6th ed, Oxford University Press, New York, 2009.

This page intentionally left blank

Appendix

This page intentionally left blank

Appendix: Resources on Health IT Evaluation

This Appendix lists some resources on health IT evaluation that the reader might find beneficial for further study. A list of relevant literature databases can be found in A. Georgiou, Finding, appraising and interpreting the evidence, in: E. Ammenwerth, M. Rigby (eds.), *Evidence-Based Health Informatics*, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

International working groups dealing with health IT evaluation

- Working Group for Assessment of Health Information Systems of the European Federation for Medical Informatics (EFMI), <http://iig.umit.at/efmi>
- Working Group Technology Assessment and Quality Development in Health Care of the International Medical Informatics Association (IMIA), <http://iig.umit.at/efmi>
- Working Group Evaluation of the American Medical Informatics Association (AMIA), <https://www.amia.org/programs/working-groups/evaluation>

International conferences also covering health IT evaluation topics

- Medinfo conference series, held every second year, organized by the International Medical Informatics Association (IMIA), www.imia.org, last access 11 February 2016.
- Medical Informatics Europe (MIE) conferences series, held every year, organized by the European Federation for Medical Informatics (EFMI), www.efmi.org, last access 11 February 2016.
- AMIA Annual Symposium, annually organized by the American Medical Informatics Association (AMIA), www.amia.org, last access 11 February 2016.
- European Conference on Information System Management an Evaluation (ECIME), held annually, <http://academic-conferences.org/ecime/ecime-home.htm>, last access 11 February 2016.

Databases related to health IT evaluation:

- EvalDB: database containing abstracts of more than 1.800 published health IT evaluation studies; <http://evaldb.umit.at>, last access 11 February 2016.
- Cochrane Library: Contains systematic reviews on various topics, including health IT; <http://www.cochranelibrary.com>, last access 11 February 2016.

On-line toolkits related to health IT evaluation:

- Health IT evaluation toolkit: Designed to help project teams develop an evaluation plan of their health IT project; <http://healthit.ahrq.gov/health-it-tools-and-resources/health-it-evaluation-toolkit-and-evaluation-measures-quick-reference>, last access 11 February 2016.
- Health IT survey compendium: Provides health IT surveys that may be used in the evaluation of health IT projects; <http://healthit.ahrq.gov/health-it-tools-and-resources/health-it-survey-compendium>, last access 11 February 2016.

Adverse Incident Reporting Systems and Databases

- European Federation of Medical Informatics “Bad Health Informatics can Kill” database (accessible via the Bad Health Informatics link at <http://iig.umit.at/efmi>, last access 11 February 2016)
- US Food and Drug Administration (FDA) Manufacturer and User Facility Device Experience (MAUDE) database (all medical devices) incident reporting system - <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/PostmarketRequirements/ReportingAdverseEvents/ucm127891.htm>, last access 11 February 2016.

Author Index

Adler-Milstein, J.	181	Lee, L.	53
Ammenwerth, E.	vii, 3, 304, 355	Luzi, D.	165
Bates, D.W.	195	Magrabi, F.	25
Beuscart-Zephir, M.-C.	126	Marceglia, S.	63
Bonacina, S.	63	Marcilly, R.	126
Brender McNair, J.	39	Nøhr, C.	139
Coiera, E.	25	Novak, L.L.	77, 237
Cornet, J.	248	Nykänen, P.	291
Craven, C.K.	77	Oluoch, T.	324
Cresswell, K.	206	Ong, M.-S.	25
Currell, R.	262	Otte-Trojel, T.	275
de Keizer, N.F.	304, 324	Pecoraro, F.	165
Doebbeling, B.	77	Peute, L.	126
Doupi, P.	220	Pincioli, F.	63
Ferrante, S.	63	Pozzi, G.	63
Furniss, D.	77	Rigby, M.	vii, 3, 355
Georgiou, A.	312	Ronchi, E.	181
Grain, H.	336	Savino, C.	63
Haux, R.	v	Scott, P.J.	102
Holden, R.J.	77, 237	Seidling, H.M.	195
Hovenga, E.	336	Sheikh, A.	53
Hyppönen, H.	181	Talmon, J.	355
Jensen, S.	152	Tamburis, O.	165
Kaipio, J.	291	Unertl, K.M.	237
Kaplan, B.	114	Urquhart, C.	262
Kotzeva, A.	248	Vallespin, B.	248
Kushniruk, A.	139	Weir, C.R.	90
Lasorsa, I.	63	Wong, G.	275
Lau, F.	77	Wyatt, J.C.	14

This page intentionally left blank

This page intentionally left blank

This page intentionally left blank