INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS 81 (2012) I–II

journal homepage: www.ijmijournal.com

Review

Methodologies for assessing telemedicine: A systematic review of reviews

Anne G. Ekeland[a,b,*], *Alison Bowes*[c], *Signe Flottorp*[d,e]

[a] *Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, Tromsø, Norway*
[b] *Department of Clinical Medicine, University of Tromsø, Norway*
[c] *School of Applied Social Science, University of Stirling, Scotland, UK*
[d] *Norwegian Knowledge Centre for the Health Services, Oslo, Norway*
[e] *Department of Public Health and Primary Health Care, University of Bergen, Norway*

ARTICLE INFO

Article history:
Received 1 March 2011
Received in revised form
18 October 2011
Accepted 18 October 2011

Keywords:
Telemedicine
Telecare
Systematic review, evaluation and assessment methodologies

ABSTRACT

Background and objectives: Previous reviews have expressed concerns about the quality of telemedicine studies. There is debate about shortcomings and appropriate methodologies. The aim of this review of systematic reviews of telemedicine is to summarize methodologies used in telemedicine research, discuss knowledge gaps and recommendations and suggest methodological approaches for further research.

Methods: We conducted a review of systematic reviews of telemedicine according to a protocol listing explicit methods, selection criteria, data collection and quality assessment procedures. We included reviews where authors explicitly addressed and made recommendations for assessment methodologies. We did a qualitative analysis of the reviews included, sensitized by two broad methodological positions; positivist and naturalistic approaches. The analysis focused on methodologies used in the primary studies included in the reviews as reported by the review authors, and methodological recommendations made by the review authors.

Results: We identified 1593 titles/abstracts. We included 50 reviews that explicitly addressed assessment methodologies. One group of reviews recommended larger and more rigorously designed controlled studies to assess the impacts of telemedicine; a second group proposed standardisation of populations, and/or interventions and outcome measures to reduce heterogeneity and facilitate meta-analysis; a third group recommended combining quantitative and qualitative research methods; and others applying different naturalistic approaches including methodologies addressing mutual adaptations of services and users; politically driven action research and formative research aimed at collaboration to ensure capacity for improvement of services in natural settings.

* *Corresponding author at*: Norwegian Centre for Integrated Care and Telemedicine, P.O. Box 6060, N-9038 University Hospital of Tromsø, Norway. Tel.: +47 952 66791.
E-mail addresses: anne.granstrom.ekeland@telemed.no (A.G. Ekeland), a.m.bowes@stir.ac.uk (A. Bowes), Signe.Flottorp@kunnskapssenteret.no (S. Flottorp).
URL: http://www.telemed.no (A.G. Ekeland).

1386-5056/$ – see front matter
doi:10.1016/j.ijmedinf.2011.10.009

Larger and more rigorous studies are crucial for the production of evidence of effectiveness of unambiguous telemedicine services for pre defined outcome measures. Summative methodologies acknowledging telemedicine as complex innovations and outcomes as partly contingent on values, meanings and contexts are also important. So are formative, naturalistic methodologies that acknowledge telemedicine as ongoing collaborative achievements and engage with stakeholders, including patients to produce and conceptualise new and effective telemedicine innovations.

Contents

1. Introduction 2
2. Methods 3
3. Results 4
   3.1. Included reviews and methodological traditions – a quantitative summary 4
   3.2. Qualitative summary of results on reported knowledge gaps, results and recommendations 5
      3.2.1. More summative studies with larger samples, preferably RCTs 5
      3.2.2. Improving summative studies through standardisation of interventions, outcome measures, theories and tools 5
      3.2.3. Combining rigorously controlled studies with qualitative inquiry 6
      3.2.4. Complexity and contexts – fitting the objects/subjects of study, assessments and interventions 6
      3.2.5. Action research 6
      3.2.6. Formative assessments – improving emerging services in natural settings 6
      3.2.7. Combining summative and formative assessments 6
4. Discussion 7
   4.1. What are larger studies, rigorous designs and standardisation producing? 7
   4.2. Developing methodologies to affirm subjectivity? 7
   4.3. Methodological approaches as political interventions? 7
   4.4. Crafting effectiveness – constructive alignment of users, producers and evaluators? 7
5. Summary 8
6. Comments on our approach and findings 8
7. Conclusions 8
   Authors' contributions 8
   Competing interests 8
   Acknowledgements 8
   Appendix A. Supplementary data 9
   References 9

1. Introduction

Many systematic reviews and primary studies addressing the impact of telemedicine have been criticized for low methodological quality [1,2]. Previous reviews have also regularly expressed concerns about the quality of telemedicine studies [3–8] and there is continuing debate about shortcomings and appropriate research methodologies. Examples include economic analysis of telemedicine which has not yet met accepted standards [7]; concerns about a relative lack of exploration of the socio-economic impact of telemedicine [9]; lack of evidence on factors promoting uptake of telemedicine [10]; a relatively undeveloped use of qualitative methods [11]; and claims that many existing studies have not been well-designed [3,5,8,12]. The need for simulation modelling has also been expressed, given perceived difficulties of building a robust evidence base for recent innovations [13]. Telemedicine research exists at the crossroads of medical, technological and social/organisational research communities. These may differ in ways that reflect differences in the nature of the topics researched, norms for the conduct of research, and what they and their stakeholders would consider important outcomes [14]. Talmon et al. have developed guidelines for the reporting of evaluations in Health Informatics, independent of the evaluation method used, the 'Statement on Reporting of Evaluation Studies in Health Informatics' (STARE-HI) [15]. They point to the gap between positivist and naturalistic approaches, express concerns that evaluation traditions do not collaborate sufficiently and state that a common language for evaluation is missing [16].

These authors and critics raise questions about the quality of research evidence in terms not only of data collected and analyzed and results produced, but also of the relevance

Table 1 – Contrasting positivist and naturalistic axioms.

Axioms	Positivist paradigm (summative assessments)	Naturalistic paradigm (formative assessments)
The nature of reality	Reality is single, tangible, and fragmentable	Realities are multiple, constructed, and holistic
The relationship of knower and known	The knower and the known are independent, a dualism	The knower and the known are interactive, inseparable
The possibility of generalization	Time- and context-free generalizations are possible	Only time- and context-bound working hypotheses are possible
The possibility of causal linkages	There are real causes, temporally precedent to or simultaneous with their effects	All entities are in a state of mutual simultaneous shaping, so that it is impossible to distinguish causes from effects
The role of values	Inquiry is value-free	Inquiry is value-based

of questions and approaches to assessments and evaluation. Basic axioms and methodologies used may create difficulties in researching all the questions that various stakeholders wish to address.

Methodology is concerned with how we know what we know, and the ways that science can help us to understand the world better. Methodology may refer to a set of methods or procedures, or it may refer to the rationale and the philosophical assumptions that underlie a particular study relative to the scientific method [19]. Positions and axioms on the nature of reality and knowing, which are set out in the philosophy of knowledge, are generally reflected in research questions and approaches.

Positivist and naturalistic approaches and their corresponding summative and formative research models may represent apparently different positions in scientific debates. These positions have been used to sensitize the presentation of qualitative results, the discussion and the conclusions in the paper.

Positivist traditions assert that the only authentic knowledge is that which is based on sense, experience and positive verification. Summative assessments, such as clinical trials and other controlled effect studies are derived from positivism. They address telemedicine as well-defined, preferably singular objects of study or interventions, and predefined outcomes. Controlled experiments, preferably randomized controlled trials (RCTs) are applied to obtain evidence of causal relations and estimates of effects. Rigorous study designs are recommended to minimise risk of bias. Researchers are considered neutral observers of objective facts. Qualitative methods are not considered appropriate to examine effects of interventions, although they may be combined with quantitative evaluations and used for e.g. preliminary exploration, to assess unexpected results or the intervention process [20,21].

Naturalistic traditions view telemedicine as heterogeneous, multiple and developing in interaction with different stakeholders in natural settings. All entities involved in telemedicine development, including technologies, are considered to be in a state of mutual simultaneous influencing. Formative assessments are applied, where researchers interact with the object of study, and knowledge and "evidence" are context bound. That is, objectivity is only partly possible, as values are inherent in all inquiry. Formative assessments often seek to move beyond the objective/subjective dualism and apply both quantitative and qualitative methods. They focus on competing discourses, conflicting scripts, and the socially contingent nature of knowledge [22–24].

Table 1, available as supplementary material, summarizes different axioms for positivist and naturalistic approaches, adapted from Lincoln and Guba [25]. It indicates that they differ in their perspectives on the nature of reality and what it is possible to know, the relationship of knower and known, on generalization and possible causation, and on the role of values in knowledge.

The paper is not concerned with the relative merits of quantitative or qualitative methods. Their value has been widely substantiated in medical, medical informatics and social science research [26–28]. Our aim in this review of systematic reviews of telemedicine is to summarize methodologies used in telemedicine research, analyze knowledge gaps and suggest methodological recommendations for further research. This paper reports on research funded under EU SMART 2008/0064, which sought to review evidence of the effectiveness of telemedicine with reference to both outcomes and methodologies.

2. Methods

We searched for systematic reviews of research on the effects of telemedicine, and reviews describing or summarizing methods used in studies assessing telemedicine. We included reviews published from 2005 and onwards. We searched these databases: British Nursing Index, Cochrane database of systematic reviews (CDSR), Database of reviews of effects (DARE), Health Technology Assessment Database (HTA), Medline, Embase, PsychInfo, Telemedicine Information Exchange (TIE) and other relevant databases. The last search was performed in July 2009.

The review adheres to the PRISMA checklist for preferred reporting items for systematic reviews, except for the items that are not relevant for a systematic review of reviews [17].

The full account of methods, data collection and quality assessment of the reviews is presented in a companion paper which reviewed effectiveness and outcomes [18]. The inclusion criteria are presented in Table 2 and the PRISMA flow chart in Fig. 1. The full protocol (Appendix 1), the search criteria (Appendix 2), the Proforma used to extract data from the reviews (Appendix 3) and the PRISMA checklist (Appendix 4) are available as supplementary material.

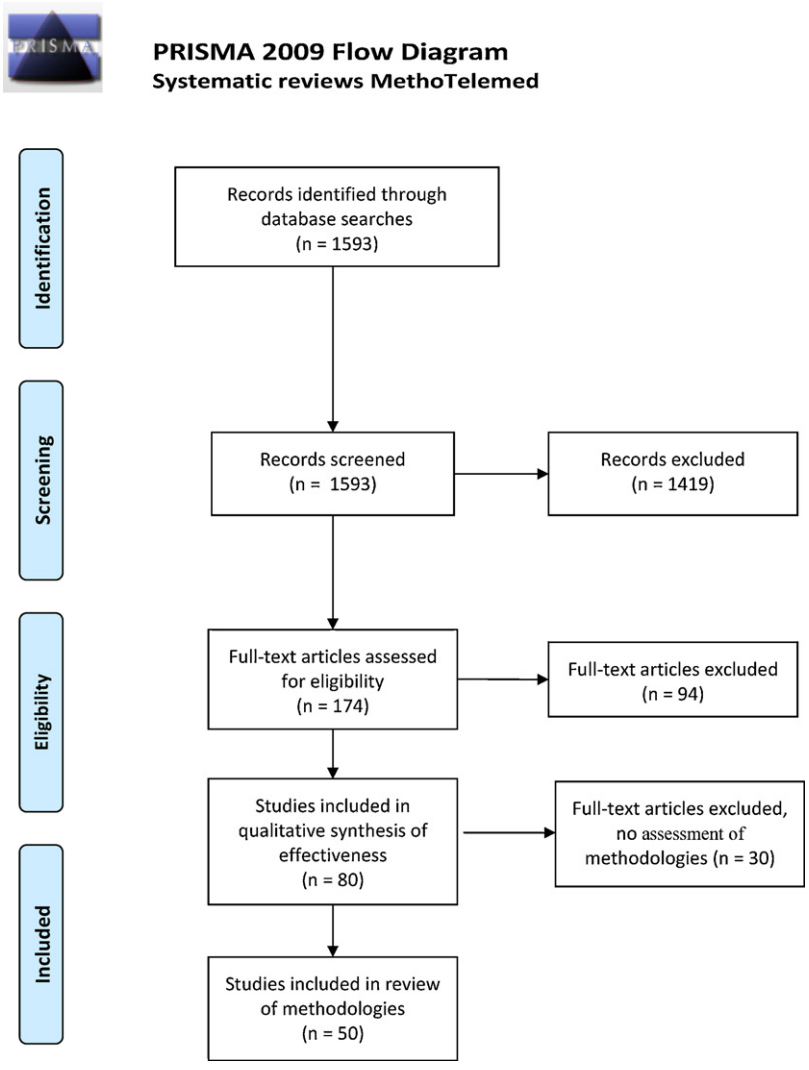

Fig. 1 – PRISMA flow diagram of included reviews.

Table 2 – Inclusion criteria.

Population/participants	Patients and consumers, health professionals and family caregivers, regardless of diagnoses or conditions
Interventions	All e-Health interventions, information and communication technologies (ICT) for communication in health care, Internet-based interventions for diagnosis and treatments, and social care if an important part of health care and in collaboration with health care for patients with chronic conditions
Comparisons	Standard care Another type of care Different e-Health solutions
Outcomes	Health related outcomes (morbidity, mortality, quality of life, patient satisfaction) Process outcomes (quality of care, professional practice, adherence to recommended practice, professional satisfaction) Costs or resource use

A group of scientific experts assessed the systematic reviews identified and extracted data using the proforma developed for this project (Appendix 3). The review team (AGE, AB and SF) checked the data abstraction and assessments made. Based on the data extracted, we performed a qualitative summary and analysis of methodologies used in the studies included in the systematic reviews.

The preliminary results were presented for discussion in two feedback workshops involving service users, providers, policy makers and researchers. We asked the participants to comment on the usefulness of the results.

3. Results

3.1. *Included reviews and methodological traditions – a quantitative summary*

The systematic search identified 1593 records of which 1419 were excluded following screening of titles and abstracts. We assessed 174 articles in full text for eligibility, and excluded 94 of these because they were of low quality or not relevant. The 80 included reviews on evidence of effectiveness and cost

of telemedicine are presented in the companion paper [18]. They are listed as supplementary material in Appendix 5. Of the 80 systematic reviews included in the review on effectiveness, 50 had explicitly commented on methodologies. These are included in this review. The PRISMA flow chart illustrates the selection process: Insert PRISMA flow chart with the logic of the 50 included reviews for this paper

Table 3 presents the 50 reviews according to their reference number in this paper, the methodologies used and the recommendations made. It is available as supplementary material. Of these 50 reviews, 30 had only included primary studies using summative assessment approaches, involving quantitative methods such as RCTs, observational studies, economic evaluations, pre–post studies, surveys and analysis of existing datasets (e.g. census or surveys, statistics, administrative data). In summative traditions, qualitative methods are considered to be suitable for enriching data collected by quantitative methods. In ten of these 30 reviews use of supplementary qualitative methods was reported, such as interviews, self reported perceptions of control and case control studies. Three reviews had considered formative assessments, using various quantitative, qualitative and process evaluation approaches. Eighteen reviews had included studies with both summative and formative approaches. Of these, three had a specific focus on methodologies. The following sections present the qualitative summary of knowledge gaps, methodologies used, and the comments and recommendations made by the authors of these 50 systematic reviews, classified in seven groups.

3.2. *Qualitative summary of results on reported knowledge gaps, results and recommendations*

3.2.1. *More summative studies with larger samples, preferably RCTs*

Eleven systematic reviews proposed conducting larger and more rigorously designed controlled studies. Many of the studies performed to assess the effects of telemedicine were considered too small with short term follow up only, and with suboptimal designs leading to risk of biased effect estimates. The knowledge gap identified as a main argument for this proposal was lack of evidence for clinical effects, due to the fact that telemedicine and e-Health were relatively new services. Evidence on effectiveness was considered crucial for decisions on further implementation. We found these recommendations in the following reviews: virtual reality (VR) in stroke rehabilitation [29]; Internet administered Cognitive Behavioural Therapy (eCBT) [30]; e-Therapy [31]; and electronic decision support systems [32]. A review on technologies for an aging society pointed out the need for evidence not only for clinical, but also technical, ethical, legal, economical and organisational effects [33]. A review comparing costs of telemonitoring and usual care for heart failure underlined the need for evidence of both indirect and direct costs [34]. Better, more and larger studies were also in demand for dementia care [35], for the effects of information technology in health promotion [36], and for the effect of robot-aided therapy on recovery of the hemiparetic arm after stroke [37]. Holland et al. instituted a search for RCTs of multidisciplinary telemedicine interventions for patients with heart failure [38], and Griffiths and Christensen established the absence of a focus on self-help programs for children and called for research to produce evidence for the effects of Internet programs for older people [39].

3.2.2. *Improving summative studies through standardisation of interventions, outcome measures, theories and tools*

Knowledge gaps deriving from heterogeneity of study design, interventions, participants, outcome measures and analysis, which made firm and general conclusions difficult to draw, provided the rationale for proposals to standardise a range of aspects in summative research on telemedicine. Fourteen reviews proposed standardisation of partly overlapping aspects. van den Berg et al. [40], for instance recommended standardisation of all aspects referred to above regarding Internet-based physical activity interventions. Referring to the fact that not all studies reported the same improvements in telemonotoring for heart failure, and several had small sample sizes, Maric et al. recommended controlled, randomized studies directly comparing different modalities and evaluating their success and feasibility when used as part of routine clinical care [41]. A review of communication technology in care improvement strategies for type 2 diabetes called for standardised interventions, outcome measures and evaluations [42]. Gaikwad and Warren assessed the role of home-based interventions in chronic disease management and proposed evidence-based outcome indicators that could provide a basis for meta-analysis [43]. Price et al. called for a standardised format in reporting of service configurations, activity data and safety data in stroke thrombolysis services [44]. Clarke and Thiyagarajan suggested development of a definitive standards-based telemedicine evaluation framework that could be applied systematically to assess and compare telemedicine systems [45]. Rojas and Gagnon argued that evidence for cost effectiveness required consistency in measures and a more encompassing approach, and recommended that the specific context in which projects take place should be taken into account when selecting key indicators in telehomecare assessments [46]. Christensen et al. reasoned that development of a theoretical approach was important, to understand coherence and dropout in Internet interventions for anxiety and depression [47], and similarly Whitten et al. concluded that until the telemedicine field adheres to agreed standards of reporting methodological details, it will be difficult to draw firm conclusions [48]. Koch pointed to a weakness in lack of standards; firstly for combining incompatible information systems, secondly for an evaluation framework considering legal, ethical, organisational, clinical, usability and technical aspects, and thirdly for proper guidelines for practical implementation of home telehealth solutions [49]. Martin et al. argued that an international descriptive terminology was needed for RCTs in assessments of smart home technologies [50]. Jaana et al. recommended study designs that would induce more control of interventions and effects in hypertension, aiming at creating better conditions to investigate long term effects [51,52]. Finally, Verhoeven et al. promoted consistency of evaluation through calibration of ICT methods and intended outcomes in diabetes care [53].

One of Scott et al.'s [54] explicit objectives was to summarize outcome indicators that had been used to evaluate telehealth projects for quality, access, acceptability and cost. They found the situation complex and confusing and proposed a continuing process of telehealth research in order to identify a small number of appropriate outcome indicators, related measures and tools, consistent descriptions and applications in future evaluations.

3.2.3. *Combining rigorously controlled studies with qualitative inquiry*

Seven reviews identified lack of knowledge about subjective meaning and individual responses, in addition to objective measures, in order to understand different effects of similar interventions. One example was diabetes, where motivation and patient participation were considered key elements for the success of telemedicine interventions [55]. Similarly, Bewick et al. proposed further controlled trials to determine effects, as well as studies to determine which elements were keys to which outcomes and to understand if different interventions engaged low and high risk drinkers [56]. In their review, which investigated the effectiveness of web based interventions to decrease alcohol consumption, Christensen et al. also concluded that very few studies formally examined reasons for dropout and proposed that this should be done in addition to controlled studies of effects [47]. Elliot et al. argued that it might be useful to investigate psychological factors that could be theoretically linked to responses to e-interventions for college drinking in future studies [57]. In addition to proposing standardised studies and larger samples in their review of home telemonitoring for patients with diabetes, Jaana et al. also questioned whether this was realistic, taking into account the variation in patient characteristics such as background, ability, medical condition and sample selection [52]. Studies to explore the attitudes of service providers as a clue for understanding success [58] and to measure different consumers' experiences and factors that influenced them, were also called for [59].

3.2.4. *Complexity and contexts – fitting the objects/subjects of study, assessments and interventions*

The needs and wishes of different target populations were identified as complex and under-researched in a number of reviews. For instance, Reger and Gahm showed that a significant number of individuals preferred ICT to walking into a counselling centre [60]. They suggested that assumptions that a therapist was preferred over ICT might be inaccurate for significant subgroups. Their meta-analysis provided preliminary support for the use of Internet and computer-based CBT for the treatment of anxiety and they recommended further research to identify different needs. Further, Weinstein reviewed Internet-based weight loss interventions [61]. She found that the studies predominately included white, educated women and proposed that future research should investigate the applicability of interventions for diverse age, ethnic, and socioeconomic groups to allow for cultural differences. Myung et al. [62] demonstrated the importance of investigating different preferences in programs for smoking cessation and Mo et al. called for research to identify how men and women utilized online support groups differently, as a crucial guide to their ongoing development [63].

3.2.5. *Action research*

Two reviews commented on gaps in knowledge about how assessments could help implement or refine interventions or innovations. For this purpose, controlled action research was proposed. Reger and Gahm [60] stated that ICT could be useful as a tool to explore the essential components of CBT as it would allow researchers to add, modify or delete specific components whilst leaving others unchanged for a comparison group. The intention should be to refine CBT theory and practice. Deshpande et al. [2] suggested that policy makers and researchers could help shape future asynchronous telehealth in Canada by pre-defining pragmatic objectives with consistent outcomes, such as telehealth triage services, that could increase the efficiency of health care and enrich the body of research.

3.2.6. *Formative assessments – improving emerging services in natural settings*

The need to know how and why certain services are naturalised or accepted, and how agency is obtained to craft success or failure, provided the rationale for recommendations for formative assessments. Despite calls for more formative assessment, only three reviews addressed and accomplished this approach. Bonacina et al. [64] addressed how telecardiology successes and pitfalls could be understood [64]. By first designing a literature-based inventory of explanations and then populating the inventory with key aspects of how and why telecardiology was widely diffused or abandoned from a systematic review of scientific papers, they aimed to highlight reasons for success. The reasons were then expected to work as inputs for change. Demiris reviewed and conceptualised virtual communities in health care as emerging practices, then identified challenges on ethical issues and questioned the possibilities of producing evidence for health outcomes [65]. He claimed that attention had turned to using assessments strategically to maximise the chances that a program would be successful, instead of waiting until the final results were available and then assessing its usefulness. Thereby he established that a program's success is contingent on contextual factors and power. Similarly, Oh et al.'s qualitative synthesis in their review of 104 published definitions of e-Health over time was formative in that they assumed that communication and knowledge among the many individuals and organisations that used the term could be improved by knowing the range of meanings encompassed by it [66].

3.2.7. *Combining summative and formative assessments*

The knowledge gap stemming from objectives both to provide evidence of effects of controlled interventions and to understand how interventions can be improved as they develop seemed to underlie ten reviews. They described the potential benefits of combining summative and formative methodologies. Murray et al. proposed that further research on interactive health communication applications (IHCAs) for people with chronic disease should include effect studies with large sample sizes, *and* the establishment of how IHCAs are being used and have their effects for different patient groups

[67]. They considered both approaches as prerequisites for improvement of services. Farmer et al. found few studies on the relationship between transmission of results, analysis, advice and subsequent behaviour change in self-monitoring of blood glucose with or without telemedicine [68]. They suggested that test results were likely to be helpful only when linked to educational or behavioural advice and changes in clinical management, and that future research designs could usefully address these processes. The limits of controlled studies and the need to adopt other approaches were made explicit in one study; Kairy et al. suggested that large scale RCTs might be difficult in a rapidly developing field in which sample sizes could be small and control groups difficult to identify [69]. They argued that if this were so, more case studies, qualitative and process studies should be conducted. In a review of the use and benefits of teleoncology, Hailey et al. requested that general findings should be validated by local studies because of the heterogeneity of contexts [70]. Hersh et al. considered teledermatology a mature intervention and proposed larger and more comprehensive analyses that assessed key patient outcomes [5]. Additionally they argued that in fields such as psychiatry and neurology, RCTs provided valuable information, but that longitudinal observational studies and demonstration projects were also useful. For home-based telemedicine, they argued that the independent contributions of technology and human resources in complex care models for patients with chronic diseases should be studied. Bee et al. suggested that future research in tele-psychotherapy might overcome existing methodological shortcomings by conducting large-scale trials that incorporated both clinical outcome and process-oriented measures [71]. Kaltenthaler et al. called for a variety of approaches in studying computerised cognitive behaviour therapy (CCBT) for depression and anxiety, including comparing CCBT with other therapies, conducting RCTs and exploring reasons for withdrawals [72]. They also recommended qualitative studies of patient preferences to inform further service developments. Linton proposed studies to clarify how patients should be selected for treatment, and of the role of CCBT compared with other treatments [73]. Polisena et al. argued that controlled studies of higher quality were required in home telehealth for diabetes management in order to give more precise insights into clinical effectiveness [74,75]. They added that studies should also include more diverse diabetes populations to increase the external validity of the outcomes, and examine the impact of various clinical approaches to determine optimal telehealth use.

4. Discussion

4.1. *What are larger studies, rigorous designs and standardisation producing?*

The first knowledge gap presented was lack of evidence for effectiveness of telemedicine. RCTs with more rigorous design and less risk of bias were required to get valid estimates of the effectiveness of telemedicine. Whilst many reviews identified large numbers of small studies and argued that larger studies are needed, there was little discussion of how large studies may be conducted in order to produce sufficient clinically relevant evidence for implementation of telemedicine. One might argue that well conducted small studies with low risk of bias may still provide useful and valid evidence, but they will usually lack power to detect clinically important effects of telemedicine. For the second group of reviews, the knowledge gap was also described as lack of evidence, and standardisation of interventions, population and outcome measures were recommended. Relevant for a discussion of methodologies, the authors assumed that control and standardisation were both possible and desirable. Patients as objects of the study of outcomes were implicitly expected to possess equal and measurable attributes possible for researchers to know objectively and generalize.

4.2. *Developing methodologies to affirm subjectivity?*

The possibility of measurable and generalized evidence was challenged as effects were considered to be contingent upon users' participation, motivations, attitudes and knowledge. By drawing attention to patients and stakeholders with subjective attitudes and intentions for their use of services, and therefore producing different responses to similar services, group three and four understood knowledge gaps to be of a more subjective character. Research on reasons for heterogeneous effects and responses were recommended. From a methodological point of view, subjective meaning was highlighted as important for understanding effects of interventions. Thus, the studies implicitly demonstrated that evidence for the effects of services must be expected to be both temporal and contingent.

4.3. *Methodological approaches as political interventions?*

Lack of knowledge that could help mobilize capacities to implement new services was addressed by recommending controlled action research. From a methodological point of view this approach is different from both naturalist and positivist approaches. In naturalist methodologies, different stakeholders and interventions are assumed to co-produce effects and goals through interaction between different interests in specific contexts or situations. The fact that goals here are described as predefined by policy makers and researchers partly described as supporters for service developers and policy makers, implies that positivist studies' assumptions of researchers as objective and autonomous are contradicted. These recommendations reflect a view that research methodologies can be partly political tools.

4.4. *Crafting effectiveness – constructive alignment of users, producers and evaluators?*

Underlying processes concerning production of effects were considered under-researched. Assessments that seek to explore these, to produce new concepts for what counts as quality and hence to identify ways to improve services in natural settings were recommended. Telemedicine was implicitly assumed both to be complex, shaped in dynamic processes

and to produce dynamic effects, partly contingent on context and the meanings that actors attribute to its use. Thereby the possibility of measuring objective effects was questioned. Assessments were recommended that recognise ways that shared actions can modify concepts, understandings, configurations, successes or failures, that is, formative approaches.

There were clear suggestions that controlled studies, whilst crucially important, could not address all current questions needed to understand and produce success or failure, or provide all necessary answers.

Medical technologies do not necessarily undergo strict and controlled assessments before they are used in clinical practice, in contrast with the requirements for the licensing of drugs. In telemedicine, the complexity, experimental and temporal aspects of services, and the communication and collaboration capacities of the technologies are key characteristics. The success and use of these services can therefore not only be depending on the technology per se but also on patients' or medical staffs' motivations, knowledge and extra work efforts. In addition political and economic incentives seem to be crucial for ensuring capacity and crafting agency to produce quality and effectiveness in ongoing collaborative processes. Thus it is important that telemedicine services are recognised not only as controlled and singular interventions, hence necessitating assessment methodologies beyond controlled effect studies.

Table 2 suggests that objectivist and subjectivist methodologies are mutually exclusive. Based on the results from the reviews, our argument is that they are complementary and that both are needed in research on different aspects of telemedicine.

5. Summary

An understanding of the positivist and naturalistic methodological positions and their underlying axioms regarding the nature of reality and knowing, the relationships of knower and known, the possibility of generalization and causal linkage and the role of contexts and values, assisted in making sense of the methodological attributes of the research reported in the reviews. Reviews including positivist and summative methodologies concluded that we need more controlled studies of better quality including standardisation of methodological aspects. Reviews that had included only summative primary studies also recommended that future research should address issues of subjectivity, needs, preferences, attitudes and knowledge of target populations and service providers. Reviews that had included both formative and summative methodologies noted that methodologies should also assume and address telemedicine interventions and achievements as complex and ongoing innovations in natural settings.

The included reviews also demonstrated that combinations of apparently opposing approaches like the positivist summative and naturalistic formative ones, gave valuable insights into the ways recent telemedicine assessments and evaluations can contribute to methodological debate. There is a clearly established need for further development of combinations of positivist and naturalist inquiry in the field of telemedicine, drawing on the limitations and strengths of both.

6. Comments on our approach and findings

The strengths of our review are the rigorous approach, the comprehensive coverage and the critical focus on methodology. The main limitation is that we have conducted a systematic review of reviews, without looking for more recently published studies not included in the reviews. We had to rely on the information available in the reviews, and we have not been able to look more deeply into the studies that had used innovative approaches. The strength of our findings and conclusions is thus limited by the quality of and the information in the included reviews.

We used the two methodological positions, the positivist and the naturalistic, as a guide for assessing the use of methodology in studies of telemedicine. We found this useful in the analyses and presentation of the results, but other approaches might have provided other insights and results.

7. Conclusions

Larger and more rigorous controlled studies, including standardisation of methodological aspects are recommended to produce evidence for the effectiveness of unambiguous telemedicine services on pre defined outcome measures. Summative methodologies that acknowledge telemedicine as complex innovations and outcomes as partly contingent on values meaning and changing contexts are also important. So are formative, naturalistic methodologies that acknowledge telemedicine as ongoing collaborative achievements and engage with stakeholders including patients, to ensure capacities and support processes to co-produce and conceptualise effective telemedicine innovations.

Authors' contributions

AGE managed the review process. All authors collaborated on the protocol and proforma, read abstracts, selected reviews for full text assessment and reviewed papers. AGE wrote the draft of the full manuscript. All authors critically revised the manuscript and approved the final version.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The authors are grateful to the European Commission DG Information Society for funding the one-year study Methodology to Assess Telemedicine Applications (MethoTelemed). We would also like to express gratitude to our MethoTelemed colleagues and the external review experts: Claus Duedal Pedersen, Kristian Kidholm and Niels Rossing, Denmark, David Bell and Petra Wilson, UK, Ellen Rygh, Norway and Paolo

Summary points

What was already known on this topic?

- There is concern about the quality of research on telemedicine
- There is debate about shortcomings and appropriate methodologies

What this study added to our knowledge?

- It is necessary to carry out larger and better summative studies of pre defined effects of controlled telemedicine interventions
- Such studies cannot address all emerging questions
- Assessments of telemedicine as complex interventions and with mixed methodological approaches considering outcome as partly contingent on values and contexts are also recommended by authors of reviews
- Assessments that formatively engage with stakeholders including patients, in natural settings to support effective telemedicine innovations are increasing and also recommended

Zanaboni, Italy. We are also grateful to the research librarians Ingrid Harboe and Marit Johansen at Norwegian Knowledge Centre for the Health Services for doing the systematic literature search, the project officers at the Norwegian Centre for Integrated Care and Telemedicine: Heidi Aasheim-Olsen, Vibeke Hemmingsen, Hilde Gard and Sissel Breivoll, and webdesigner Jarl-Stian Olsen.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijmedinf.2011.10.009.

REFERENCES

[1] A. Deshpande, S. Khoja, A. McKibbon, A.R. Jadad, Real-time (Synchronous) Telehealth in Primary Care: Systematic Review of Systematic Reviews, 2008.
[2] A. Deshpande, S. Khoja, J. Lorca, A. McKibbon, C. Rizo, A.R. Jadad, Asynchronous Telehealth Systematic Review of Analytic Studies and Environmental Scan of Relevant Initiatives, Canadian Agency for Drugs and Technologies in Health, Ottawa, 2008.
[3] D. Hailey, R. Roine, A. Ohinmaa, Systematic review of evidence for the benefits of telemedicine, J. Telemed. Telecare 8 (Suppl. 1) (2002) 1–30.
[4] G. Demiris, D. Tao, An analysis of the specialized literature in the field of telemedicine, J. Telemed. Telecare 11 (6) (2005) 316.
[5] W.R. Hersh, D.H. Hickam, S.M. Severance, T.L. Dana, K.P. Krages, M. Helfand, Telemedicine for the medicare population: update, Evid. Rep. Technol. Assess. (Full Rep.) (February (131)) (2006) 1–41, AHRQ Publication No. 06-E007.
[6] R. Roine, A. Ohinmaa, D. Hailey, Assessing telemedicine: a systematic review of the literature, Can. Med. Assoc. J. 165 (6) (2001) 765.
[7] P. Whitten, F. Mair, A. Haycox, C. May, T. Williams, S. Hellmich, Systematic review of cost effectiveness studies of telemedicine interventions, Br. Med. J. 324 (7351) (2002) 1434.
[8] W. Hersh, M. Helfand, J. Wallace, D. Kraemer, P. Patterson, S. Shapiro, et al., Clinical outcomes resulting from telemedicine interventions: a systematic review, BMC Med. Inform. Decis. Mak. 1 (1) (2001) 5.
[9] P. Jennett, L. Affleck Hall, D. Hailey, A. Ohinmaa, C. Anderson, R. Thomas, et al., The socio-economic impact of telehealth: a systematic review, J. Telemed. Telecare 9 (6) (2003) 311.
[10] E. Johnsen, E. Breivik, R. Myrvang, F. Olsen, Benefits from Telemedicine in Norway, 2006.
[11] E. Murphy, R. Dingwall, D. Greatbatch, S. Parker, P. Watson, Qualitative research methods in health technology assessment: a review of the literature, Health Technol. Assess. 2 (16) (1998), iii–ix, 1–274.
[12] R. Currell, C. Urquhart, P. Wainwright, et al., Telemedicine Versus Face to Face Patient Care: Effects on Professional Practice and Health Outcomes. Cochrane Library The Cochrane Collaboration, Update Software, Oxford, 2002 (Issue 3).
[13] J. Barlow, S. Bayer, B. Castleton, R. Curry, Meeting government objectives for telecare in moving from local implementation to mainstream services, J. Telemed. Telecare 11 (Suppl. 1) (2005) 49–51.
[14] A. Clarke, Evidence-based evaluation in different professional domains: similarities, differences and challenges, in: F. Shaw, GJaMM (Eds.), The SAGE Handbook of Evaluation, SAGE Publications, London, 2000, pp. 559–581.
[15] 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.
[16] E. Ammenwerth, J. Brender, P. Nykanen, 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 (June (6)) (2004) 479–491.
[17] D.L.A. Moher, J. Tetzlaff, D.G. Altman, The PRISMA Group preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement, PLoS Med. 6 (6) (2009) e1000097.
[18] A.G. Ekeland, A. Bowes, S. Flottorp, Effectiveness of telemedicine – a systematic review of reviews, Int. J. Med. Inform. 79 (11) (2010) 736–771.
[19] Oxford Dictionaries Online, Oxford University Press, 2010.
[20] M.Q. Patton, Qualitative Research & Evaluation Methods, Sage Publications, 2001.
[21] P. Craig, P. Dieppe, S. Macintyre, S. Michie, I. Nazareth, M. Petticrew, Developing and evaluating complex interventions: the new Medical Research Council guidance, Br. Med. J. 337 (2008) 979–983, ISSN 0959-8138 http://eprints.gla.ac.uk/42736/.
[22] M. Shriven, The methodology of evaluation, in: M.E. Gredler (Ed.), Program Evaluation, Prentice Hall, New Jersey, 1967.
[23] A. Rip, J. Schot, T. Misa, Constructive Technology Assessment: A New Paradigm for Managing Technology in Society, 1995.
[24] Research Methods Knowledge Base (Database on the Internet), 2006 (cited 23 November 2010). Available from: http://www.socialresearchmethods.net/kb/index.php.
[25] Y. Lincoln, E.G. Guba, Naturalistic Inquiry, Sage Publications, Newbury Park, CA, 1985.
[26] S. Lewin, C. Glenton, A.D. Oxman, Use of qualitative methods alongside randomised controlled trials of complex

healthcare interventions: methodological study, Br. Med. J. 339 (2009) b3496.
[27] C. Friedman, J. Wyatt, J. Ash, Evaluation Methods in Biomedical Informatics, Springer, 2006.
[28] J. Creswell, Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, Sage Publications, 2009.
[29] J.H. Crosbie, S. Lennon, J.R. Basford, S.M. McDonough, Virtual reality in stroke rehabilitation: still more virtual than real, Disabil. Rehabil. 29 (14) (2007 Jul 30) 1139–1146 (discussion 47–52).
[30] P. Cuijpers, A. van Straten, G. Andersson, Internet-administered cognitive behavior therapy for health problems: a systematic review, J. Behav. Med. 31 (April (2)) (2008) 169–177.
[31] M.G. Postel, H.A. de Haan, C.A. De Jong, E-therapy for mental health problems: a systematic review, Telemed. J. E Health 14 (September (7)) (2008) 707–714.
[32] V. Sintchenko, F. Magrabi, S. Tipper, Are we measuring the right end-points? Variables that affect the impact of computerised decision support on patient outcomes: a systematic review, Med. Inform. Internet Med. 32 (September (3)) (2007) 225–240.
[33] G. Demiris, B. Hensel, Technologies for an aging society: a systematic review of "smart home" applications, in: Yearbook of Medical Informatics, 2008, p. 33.
[34] E. Seto, Cost comparison between telemonitoring and usual care of heart failure: a systematic review, Telemed. J. E Health 14 (September (7)) (2008) 679–686.
[35] S. Lauriks, A. Reinersmann, H.G. Van der Roest, F.J. Meiland, R.J. Davies, F. Moelaert, et al., Review of ICT-based services for identified unmet needs in people with dementia, Ageing Res. Rev. 6 (October (3)) (2007) 223–246.
[36] T.P. Lintonen, A.I. Konu, D. Seedhouse, Information technology in health promotion, Health Educ. Res. 23 (June (3)) (2008) 560–566.
[37] G.B. Prange, M.J. Jannink, C.G. Groothuis-Oudshoorn, H.J. Hermens, M.J. Ijzerman, Systematic review of the effect of robot-aided therapy on recovery of the hemiparetic arm after stroke, J. Rehabil. Res. Dev. 43 (March–April (2)) (2006) 171–184.
[38] R. Holland, J. Battersby, I. Harvey, E. Lenaghan, J. Smith, L. Hay, Systematic review of multidisciplinary interventions in heart failure, Heart 91 (July (7)) (2005) 899–906.
[39] K.M. Griffiths, H. Christensen, Review of randomized controlled trials of Internet intervention for mental disorders and related conditions, Clin. Psychol. 10 (1) (2006) 16–29.
[40] M.H. van den Berg, J.W. Schoones, T.P. Vliet Vlieland, Internet-based physical activity interventions: a systematic review of the literature, J. Med. Internet Res. 9 (3) (2007) e26.
[41] B. Maric, A. Kaan, A. Ignaszewski, S.A. Lear, A systematic review of telemonitoring technologies in heart failure, Eur. J. Heart Fail. 11 (May (5)) (2009) 506–517.
[42] A. Mathur, J.C. Kvedar, A.J. Watson, Connected health: a new framework for evaluation of communication technology use in care improvement strategies for type 2 diabetes, Curr. Diabetes Rev. 3 (November (4)) (2007) 229–234.
[43] R. Gaikwad, J. Warren, The role of home-based information and communications technology interventions in chronic disease management: a systematic literature review, Health Inform. J. 15 (June (2)) (2009) 122–146.
[44] C.I. Price, F. Clement, J. Gray, C. Donaldson, G.A. Ford, Systematic review of stroke thrombolysis service configuration, Expert Rev. Neurother. 9 (February (2)) (2009) 211–233.
[45] M. Clarke, C.A. Thiyagarajan, A systematic review of technical evaluation in telemedicine systems, Telemed. J. E Health 14 (March (2)) (2008) 170–183.
[46] S.V. Rojas, M.P. Gagnon, A systematic review of the key indicators for assessing telehomecare cost-effectiveness, Telemed. J. E Health 14 (November (9)) (2008) 896–904.
[47] H. Christensen, K.M. Griffiths, L. Farrer, Adherence in Internet interventions for anxiety and depression, J. Med. Internet Res. 11 (2) (2009) e13.
[48] P. Whitten, L.K. Johannessen, T. Soerensen, D. Gammon, M. Mackert, A systematic review of research methodology in telemedicine studies, J. Telemed. Telecare 13 (5) (2007) 230–235.
[49] S. Koch, Home telehealth—current state and future trends, Int. J. Med. Inform. 75 (August (8)) (2006) 565–576.
[50] S. Martin, G. Kelly, W.G. Kernohan, B. McCreight, C. Nugent, Smart home technologies for health and social care support, Cochrane Database Syst. Rev. (4) (2008), CD006412.
[51] M. Jaana, G. Pare, C. Sicotte, Hypertension home telemonitoring: current evidence and recommendations for future studies, Dis. Manage. Health Outcomes 15 (1) (2007) 19–31.
[52] M. Jaana, G. Pare, Home telemonitoring of patients with diabetes: a systematic assessment of observed effects, J. Eval. Clin. Pract. 13 (April (2)) (2007) 242–253.
[53] F. Verhoeven, L. van Gemert-Pijnen, K. Dijkstra, N. Nijland, E. Seydel, M. Steehouder, The contribution of teleconsultation and videoconferencing to diabetes care: a systematic literature review, J. Med. Internet Res. 9 (5) (2007) e37.
[54] R.E. Scott, F.G. McCarthy, P.A. Jennett, T. Perverseff, D. Lorenzetti, A. Saeed, et al., Telehealth outcomes: a synthesis of the literature and recommendations for outcome indicators, J. Telemed. Telecare 13 (Suppl. 2) (2007) 1–38.
[55] M. Azar, R. Gabbay, Web-based management of diabetes through glucose uploads: has the time come for telemedicine? Diabetes Res. Clin. Pract. 83 (January (1)) (2009) 9–17.
[56] B.M. Bewick, K. Trusler, M. Barkham, A.J. Hill, J. Cahill, B. Mulhern, The effectiveness of web-based interventions designed to decrease alcohol consumption—a systematic review, Prev. Med. 47 (July (1)) (2008) 17–26.
[57] J.C. Elliott, K.B. Carey, J.R. Bolles, Computer-based interventions for college drinking: a qualitative review, Addict. Behav. 33 (August (8)) (2008) 994–1005.
[58] F. Griffiths, A. Lindenmeyer, J. Powell, P. Lowe, M. Thorogood, Why are health care interventions delivered over the Internet? A systematic review of the published literature, J. Med. Internet Res. 8 (2) (2006) e10.
[59] K.M. Akesson, B.I. Saveman, G. Nilsson, Health care consumers' experiences of information communication technology—a summary of literature, Int. J. Med. Inform. 76 (September (9)) (2007) 633–645.
[60] M.A. Reger, G.A. Gahm, A meta-analysis of the effects of Internet- and computer-based cognitive-behavioral treatments for anxiety, J. Clin. Psychol. 65 (January (1)) (2009) 53–75.
[61] P.K. Weinstein, A review of weight loss programs delivered via the Internet, J. Cardiovasc. Nurs. 21 (July–August (4)) (2006) 251–258 (quiz 9–60).
[62] S.K. Myung, D.D. McDonnell, G. Kazinets, H.G. Seo, J.M. Moskowitz, Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials, Arch. Intern. Med. 169 (May (10)) (2009) 929–937.
[63] P.K. Mo, S.H. Malik, N.S. Coulson, Gender differences in computer-mediated communication: a systematic literature review of online health-related support groups, Patient Educ. Couns. 75 (April (1)) (2009) 16–24.
[64] S. Bonacina, L. Draghi, M. Masseroli, F. Pinciroli, Understanding telecardiology success and pitfalls by a systematic review, Stud. Health Technol. Inform. 116 (2005) 373–378.

[65] G. Demiris, The diffusion of virtual communities in health care: concepts and challenges, Patient Educ. Couns. 62 (August (2)) (2006) 178–188.
[66] H. Oh, C. Rizo, M. Enkin, A. Jadad, What is eHealth (3): a systematic review of published definitions, J. Med. Internet Res. 7 (1) (2005) e1.
[67] E. Murray, J. Burns, T.S. See, R. Lai, I. Nazareth, Interactive Health Communication Applications for people with chronic disease, Cochrane Database Syst. Rev. (4) (2005), CD004274.
[68] A. Farmer, O.J. Gibson, L. Tarassenko, A. Neil, A systematic review of telemedicine interventions to support blood glucose self-monitoring in diabetes, Diabet. Med. 22 (October (10)) (2005) 1372–1378.
[69] D. Kairy, P. Lehoux, C. Vincent, M. Visintin, A systematic review of clinical outcomes, clinical process, healthcare utilization and costs associated with telerehabilitation, Disabil. Rehabil. 31 (6) (2009) 427–447.
[70] D. Hailey, M.-J. Paquin, O. Maciejewski, L. Harris, A. Casebeer, G. Fick, et al., The Use and Benefits of Teleoncology, The Institute of Health Economics (IHE), Canada, January 2007.
[71] P.E. Bee, P. Bower, K. Lovell, S. Gilbody, D. Richards, L. Gask, et al., Psychotherapy mediated by remote communication technologies: a meta-analytic review, BMC Psychiatry 8 (2008) 60.
[72] E. Kaltenthaler, J. Brazier, E. De Nigris, I. Tumur, M. Ferriter, C. Beverley, et al., Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation, Health Technol. Assess. 10 (September (33)) (2006), iii, xi–xiv, 1–168.
[73] S. Linton, Datorbaserad kognitiv beteendeterapi vid ångestsyndrom eller depression, 2007.
[74] J. Polisena, K. Tran, K. Cimon, B. Hutton, S. McGill, K. Palmer, Home telehealth for diabetes management: a systematic review and meta-analysis, Diabetes Obes. Metab. 11 (October (10)) (2009) 913–930.
[75] J. Polisena, D. Coyle, K. Coyle, S. McGill, Home telehealth for chronic disease management: a systematic review and an analysis of economic evaluations, Int. J. Technol. Assess. Health Care 25 (July (3)) (2009) 339–349.