**Title**

Patient Safety Informatics: Recommendations to meet the challenges of emerging digital health

**Authors**

*Name*: Ciarán D. McInerney [CORRESPONDING AUTHOR]

*Degree*: PhD

*Affiliation*: NIHR Yorkshire and Humber Patient Safety Translational Research Centre; University of Leeds

*Full address*: University of Leeds, Woodhouse, Leeds, LS2 9JT

*Email*: [c.mcinerney@leeds.ac.uk](mailto:c.mcinerney@leeds.ac.uk)

*Phone*: n/a

*Name*: Rebecca Randell

*Degree*: PhD

*Affiliation*: University of Bradford

*Full address*: University of Bradford, Richmond Rd, Bradford, BD7 1DP

*Name*: Ibrahim Habli

*Degree*: PhD

*Affiliation*: University of York

*Full address*: University of York Deramore Lane, York, YO10 5GH

*Name*: David Jenkins

*Degree*: PhD

*Affiliation*: NIHR Greater Manchester Patient Safety Translational Research Centre; University of Manchester

*Full address*: University of Manchester, Oxford Rd, Manchester M13 9PL

*Name*:

*Degree*:

*Affiliation*:

*Full address*:

*Name*:

*Degree*:

*Affiliation*:

*Full address*:

*Name*:

*Degree*:

*Affiliation*:

*Full address*:

*Name*: Niels Peek

*Degree*: PhD

*Affiliation*: NIHR Greater Manchester Patient Safety Translational Research Centre; University of Manchester

*Full address*: University of Manchester, Oxford Rd, Manchester M13 9PL

*Name*: Owen Johnson

*Degree*: MSc

NIHR Yorkshire and Humber Patient Safety Translational Research Centre; University of Leeds

*Full address*: University of Leeds, Woodhouse, Leeds, LS2 9JT

*Likely word count for Viewpoint or Perspective article = 2,000 words*

# Introduction

The fourth industrial revolution is based on cyber-physical systems and the connectivity of devices to which healthcare must adapt (Aceto, Persico and Pescapé, 2020). Healthcare is already becoming increasingly digital and connected with moves toward fog computing and the Internet-of-Things (Farahani *et al.*, 2018). Additionally, at the time of writing, the COVID-19 pandemic is occurring and accelerating the conception, design, development and use of digital health technology. Healthcare providers have quickly responded with rapid wide-spread adoption of existing technology like video consultation (Wherton *et al.*, 2020). Other technologies like electronic health records, decision-support tools and handheld medical devices have been developed and used for many years with reported benefits for patient care but also with concerns for patient safety (Sittig *et al.*, 2018). It is currently unclear what the consequences are for patient safety as existing health information technologies become ubiquitous with increasing pace and interact in unforeseen ways (Benbya *et al.*, 2020). There is a need for an improved understanding and praxis of patient safety in relation to information technology.

The Patient Safety Translational Research Centres were set up by the UK National Institute for Health Research to translate patient-safety knowledge into practice. In April 2020, a national, expert collaboration led by the Centres from both Yorkshire and Humber and Greater Manchester was set up to appraise the academic evidence for patient safety in health information systems. Our collaborative intends to host a series of workshops that deliver publications to engage those directly involved in the delivery and study of healthcare, and to provide recommendations to address theoretical and practical challenges for Patient Safety Informatics. Our aim is to define the field of Patient Safety Informatics from a UK perspective and establish a platform of Patient Safety Informatics theory for future research and development.

In Section 1 of this paper, we outline the Patient Safety Informatics domain at the intersection of safety science and health informatics and highlight the need for theory development and research. Section 2 summarises the workshop process. In Section 3, we present the output from the workshop: challenges and patient-safety consequences of emerging changes to digital health and recommendations to address them.

# Section 1: Patient Safety and Patient Safety Informatics

## Patient safety and digital health

Although no consensual definition exists, patient safety can be considered to be the “*avoidance, prevention and amelioration of adverse outcomes or injuries stemming from the process of healthcare*” (Vincent, 2010). It is sometimes conceptualised as a balance between risks of harm, resource use, and improving patient health (Cook and Rasmussen, 2005). Healthcare is a safety-critical industry (Sujan *et al.*, 2016) that must approach safety by concurrently avoiding, managing and embracing risk (Vincent and Amalberti, 2016). This sets healthcare apart from other safety-critical industries that predominantly focus on only one of these approaches.

While the patient-safety perspective on health information technology is not novel (e.g. Gómez-González et al., 2020; Kostkova, 2015), the types of patient-safety challenges and our capacity to address them are constantly in flux. Health information technology is becoming increasingly networked in line with the fourth industrial revolution (Wickramasinghe and Bodendorf, 2020), posing novel safety issues as technologies interact (Paxton and Branca, 2020). This is because when health information technologies interact, they form a health information system (Onik, Fielt and Gable, 2017), the success or failure of which is partly due to emergent rather than planned change resulting from local improvisation (Heeks, 2006). These health information systems are the inevitable structure of how digital health is evolving (Farahani *et al.*, 2018; Aceto, Persico and Pescapé, 2020), and will require a systemic perspective from developers, users and patient-safety researchers to mitigate emergent challenges to patient safety.

Helpfully, Markus (2004) provides a framework to map the ways that digital health could evolve. Markus's (2004) implies a 2x2 model describing the risks associated with both novel and existing technologies and their application (Table x2x2x). *Technochange* refers to the highest-risk of combining *novel* applications of *new* technologies. This high-risk path for digital health is driven by the relationship between vendors who want to be first to market and buyers who want to be seen to innovate. These incentives can encourage high risks for associated large rewards but at the cost of patient safety. It is important to note that health information systems are complex adaptive systems (Johnson, 2019), so whether technology is introduced via familiar or novel applications, it is likely to have unforeseeable consequences.

|  |
| --- |
| **Table x2x2x** Contingency table illustrating the risk categories associated with interactions of novel and existing technology and its application. Adapted from Markus (2004). |
| |  |  |  | | --- | --- | --- | |  | *New technology* | *Existing technology* | | *Novel application* | High risk | Moderate risk | | *Familiar application* | Moderate risk | Low risk | |
|  |

Whether digital health evolves along high, moderate or low risk paths, many challenges posed by increasingly-complex digital health are similar: innovations are unlikely to be equally affordable and available for all (McAuley, 2014; Robinson *et al.*, 2015; Lupton, 2017; Banerjee, 2019); algorithms and models are of transient relevance (Jenkins *et al.*, 2018); there has been a a continued lack of sufficient testing, despite early calls (Leveson, 1986); societal challenges like an aging population (Pilotto, Boi and Petermans, 2018); and legal and political jurisdiction (Wismar *et al.*, 2011). Each of these challenges are associated with known and unknown consequences for patient safety, which need to be addressed for responsible provision of healthcare. Hence, there is a need for rigorous study of the relationship between emerging digital health and patient safety, i.e. a Patient Safety Informatics.

## Patient Safety Informatics

Although there is no official definition of Patient Safety Informatics, the International Medical Informatics Association working group on ‘Health Informatics for Patient Safety’ consider their role as “[promoting] *patient safety of health information systems and their associated medical devices* [and focusing on] *how healthcare information systems can improve patient safety, as well as identifying and rectifying safety issues*” (IMIA WG7, 2018). This scope is exemplified in Singh and Sittig's (2016) Health Information Technology Safety Measurement Framework, which defines three safety domains embedded in a socio-technical work system: safe health information technology, safe use of health information technology, and using health information technology to improve safety. The following paragraphs are in support of our proposed definition of Patient Safety Informatics: the study of patient-safety-related information in healthcare systems.

*Informatics* is the interdisciplinary study of information and its environment (Gammack, Hobbs and Pigott, 2011) with information flow recognised as key to system safety, as a vital resource and as a reflection of safety culture (Westrum, 2014). Safety informatics is a relatively new concept with a proposed definition as a scientific discipline studying safety information and its mechanisms to address the lack of safety information in safety management (Wang and Wu, 2020). Here, safety information refers to safety-related data that shows systems’ safety state and its changes (Huang *et al.*, 2017; Wang and Wu, 2018). There have been calls to bridge the gap between research and practice in the safety of systems (Underwood and Waterson, 2013) but, despite theoretical and practical progress in safety informatics, it has yet to be applied substantially to healthcare and patient safety, in particular.

Bakken, Cimino and Hripcsak (2004) explored how informatics can promote patient safety and provided recommendations like integrating informatics into healthcare curricula and the evaluation of digital health from health-economic, clinical and administrative perspectives. While welcomed, these recommendations, and the challenges they purport to address, concern health information technologies in isolation and their function in promoting patient safety, only. As described earlier, emerging digital health must consider the safety of health information systems and their safe use, too (Singh and Sittig, 2016). A thorough exploration and instantiation of Patient Safety Informatics is thus still lacking.

It is for these reasons that the aim of our workshop series was to develop the theoretical and practical foundations of Patient Safety Informatics by exploring the theory and practice of applying safety informatics to patient safety. Our first workshop focused on the challenges and patient-safety consequences of, and recommendations to address, emerging digital health.

# Section 2: Workshop process

A workshop was convened of 14 health informatics researchers who represent those who develop and evaluate digital health. Collaborators discussed the patient-safety consequences of the challenges posed by emerging digital health, from all four quadrants of table x2x2x. Table xSupplementary1x shows some of the example health information technologies discussed in the workshop, characterised by personalisation, decentralisation, a systemic orientation, and a move toward a user-led/patient-centred experience.

In subsequent meetings, the group collated and synthesised contributions to 1) describe characteristics of new and emerging HITs, 2) describe the challenges posed when HITs are brought together within HISs, 3) describe the patient-safety consequences of these challenges, and 4) recommend approaches to address the patient-safety consequences.

# Section 3: Workshop outputs

During the workshop, we highlighted X challenges that emerging digital health pose, each with consequences for patient safety. Table xC.I.R.x summarises the challenges and our theoretical and practical recommendations to address the safety concerns. Each is only briefly presented below with references for further reading.

## Difficulty conceptualising threats to patient safety

Firstly, much of the innovation is not physical, instead leveraging existing hardware in novel ways. This manifests as software, systems architecture and communication protocols, which lack the tangibility so foundational to trust in digital and robotic systems (Glikson and Woolley, 2020). It is challenging to conceptualise threats to patient safety from these non-physical influences because it requires more-abstract consideration of interactions and effects. This can lead to inadequate consideration of threats to patient safety.

Safety cases might be a useful tool to help map the relationship between abstract influences and consequences. Safety cases are structured arguments supported by evidence that are used to justify why a system or a service in acceptably safety within a particular context (Bishop and Bloomfield, 2000). In safety-critical industries, particularly in the UK, these cases are an established means by which confidence in the safety of the system is communicated to, and scrutinised by, the diverse stakeholders, including users, regulators and policy makers. In the NHS, compliance with the clinical safety standards DCB0129 and DCB0160 requires a safety case for HITs. The process involves an exposition of risk to encourage proactive safety management (Sujan *et al.*, 2016). The preparation of safety cases guides reflexivity that can be insightful when combined with a systems approach to conceptualising risk and safety (Haimes, 2009a; Ravitz *et al.*, 2013). Thus, patient safety might be facilitated by the use of dynamic (Denney, Pai and Habli, 2015), multi-view (Flood and Habli, 2011) safety cases for HIT (Despotou *et al.*, 2012; Habli *et al.*, 2018) and for healthcare services (Sujan *et al.*, 2015).

## Unclear how to integrate and interpret data streams

Secondly, more data of a greater variety can be collected with greater ease and speed, but it is not clear how this data can be sensibly integrated and interpreted (Ranjan *et al.*, 2018). There is a risk that opportunities will be missed to use data to improve safety, and there are risks of inappropriate or biased use of data that threatens patients’ safety.

To mitigate these hazards, safe development and use of middleware will be essential to provide an intermediary “*to abstract* [the] *heterogeneity* [of health information technologies] *… to achieve a seamless integration*” (Díaz, Martín and Rubio, 2016). Proposed solutions include standards for exchanging electronic health records (Saripalle et al., 2019; see Houta et al., 2019 for application in epilepsy data), distributed architectures to integrate electronic health records (Roehrs, André and Righi, 2017; Roehrs, 2019; Roehrs *et al.*, 2019), and 3rd-party infrastructure for linkage and querying of electronic health records, e.g. the CSIRO Health Data Integration tool (Hansen, Pang and Maeder, 2007).

Other contributing solutions include dynamic modelling of the data (Jenkins *et al.*, 2018), which can provide a solution to the transient relevance of predictive models. Similarly, progress continues to be made developing models that respect the latent, data-generating processes underlying the phenomena of interest (Sperrin *et al.*, 2019), which might clarify ‘Big healthcare data’ (Beam and Kohane, 2018). Finally, progress in artificial intelligence (particularly anomaly detection) might help to mitigate problems arising from data errors (Challen *et al.*, 2019; Macrae, 2019). For example, to minimise inappropriate decisions due to poor data quality, Sako et al. (2020) provides a conceptual framework for automated assessment of data quality and information integrity. These methods are examples of how developments in digital health can be used to improve patient safety as well as help with its safe use.

## Reactive regulations and standards

Thirdly, as the pace of innovation accelerates, the current reactive (rather than proactive) regulatory- and standards-based approaches to safety will be increasingly ineffective at assuring patients’ safety (Xplaceholder – OJ and CMc paperX). The consequence is that avoidable harm might be experienced before mitigations are put in place.

We recommend synchronisation of the development and evaluation of health information technologies, similar to the IDEAL framework (Sedrakyan *et al.*, 2016). The IDEAL framework champions gradual approval of medical devices rather than the one-shot approval of CE marking (The European Parliment and The Council of the European Union, 1993), which would allow “*graded, responsible, but earlier patient access*” (Sedrakyan *et al.*, 2016; Hirst *et al.*, 2019). Such frameworks simultaneously address concerns that the increased administrative burden of more-stringent regulations might delay products that are imperfect but practically useful (Oelze, Neeser and Müller, 2019).

We also recommend that regulators and developers of standards adopt a systems approach to conceptualising risk (Haimes, 2009a) to appropriately reflect the complex adaptive nature of healthcare (Plsek and Greenhalgh, 2001). It is hoped that these recommendations might help to manage increased sensitivity to safety during development.

## Trust in opaque and complex systems

Fourthly, trust has long been a part of patient care (Thorne and Robinson, 1988; Song and Zahedi, 2007) but it is threatened by the way that digitisation and complexification of digital health can hinder the intimacy of care. Trust in healthcare is partly a function of inter-personal behaviours (Calnan and Rowe, 2006) with the gatekeeping and competing incentives of actors within a health information system potentially jeopardising this trust (Mechanic and Schlesinger, 1996; Alaszewski, 2003). Without trust in expert and reliable sources, patients’ safety is under threat from misinformation and disinformation from sources more intimate and familiar (Wardle and Derakhshan, 2017).

A further challenge to trust is when an illusion of safety is partly incentivised by short-term returns. In addition to the illusion of patient safety arising from ignorance or lack of engagement (Tomassoni and Simone, 2001; Classen, 2003), ‘safety/security theatre’ describes deliberate safety-related activities intended to provide feelings of improved safety regardless of whether they actually influence safety (Thedell, 2013). With public wariness of technology like artificial intelligence (Siau and Wang, 2018), developers of digital health are incentivised to promote their products persuasively with, for example, focus on short rather than long-term benefits (Glikson and Woolley, 2020). Patients’ safety is threatened directly by misdirection of attention and indirectly from allocation of limited resources to support the distraction.

We suggest that a socio-technical perspective will help all stakeholders in healthcare to acknowledge the systemic nature of digital health systems. Such a perspective can support an awareness and transparency as a foundation of trust (Kim and Lee, 2018), in line with the Transparency for Trust initiative (Wykes and Schueller, 2019). Socio-technical models like the Systems Engineering Initiative for Patient Safety 2.0 (Holden *et al*>, 2014) relate the components of healthcare systems, which also map to the determinants of trusting relationships with technology (cf. Figure 2 in Siau and Wang, 2018).

Practically, developers and vendors of emerging health information technologies could consider supplier declarations of conformity to industry and community co-developed ‘FactSheets’ (Arnold *et al.*, 2019). Such an approach also contributes to proactive, community-led regulation of digital health. A socio-technical perspective has the potential to address all three domains of HIT patient safety as proposed by Singh and Sittig (2016) - safe HIT, safe use of HIT, and using HIT to improve safety – which is perhaps unsurprising given that it is explicitly a socio-technical model.

## Emergent patient-safety consequences

Fifthly, although health information technologies are being developed to leverage health information systems, e.g. the Internet of Things (Wickramasinghe and Bodendorf, 2020), safety considerations are often focused on the technologies in isolation. This reductionist approach leads to a myopic view of the technologies’ effects that does not consider the emergent, patient-safety consequences. Healthcare systems are complex with a diversity of organizational forms, interdependence, and feedback effects (Begun, Zimmerman and Dooley, 2003). They are also holarchical – a nested system of systems – as exemplified by the Heimdall framework of learning health systems (Mclachlan *et al.*, 2018) and Carayon *et al.*'s (2015) model of workplace safety.

We recommend a systemic and holarchical conceptualisation of healthcare processes and patient-safety consequences to complement the complex, holarchical structure of healthcare. To this end, Haimes (2009a) describes a complexity definition of risk, which would be essential to a systems-based discussion of patient safety. Of particular note is a systems-based conceptualisation of resilience as a variable state of a system whose variability is the key performative and protective feature (Haimes, 2009b).

Practically, we recommend the aforementioned safety cases and gradual approval of medical devices as appropriate approaches to handle the limited capacity to predict the behaviour of complex systems (Hilborn, 2004). Such complexity approaches will also be useful in addressing the question of how these new and emerging challenges will interact with the existing challenges alluded to in Section 1.

## Solutionism

Sixthly, and related to the challenge of reductionism, is solutionism, which is an ideology that inappropriately recasts “*complex social situations…as neatly defined problems with definite, computable solutions…if only the right* [technologies] *are in place”* (Morozov, 2013). Examples include diet apps that inappropriately simplify body composition as merely a function of calorie consumption (Maturo, 2014), and the legal and ethical consequences of treatments like deep brain stimulation (Gardner and Warren, 2019). The consequences for patient safety are that digital health interventions might be unfit for the true hazards that they present because of distraction by techno-optimism or technology push. Techno-optimism might arise from differences of perceived risk or perceived capacity for control (Harris and Middleton, 1994), which relate to existing problems of inequality technical education and access to digital health, respectively (Gann, 2019).

In addition to earlier recommendations of adopting socio-technical perspective and a systems approach to conceptualising risk, solutionism can be addressed by adopting a systemic approach to patient safety. Ravit*z et al*. (2013) describe such an approach with a case study on medication infusion pumps and the Systems Engineering Initiative for Patient Safety 2.0 model provides a framework for understanding the structures, processes and outcomes in healthcare, more generally (Holden *et al.*, 2014). These approaches can help to sensitise developers and users of digital health to the relationships within healthcare systems that might facilitate unintended consequences.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table xC.I.R.v1x** Summary of recommendations to address safety concerns posed by the challenges of emerging digital health. Recommendations are tagged as theory development, T, and practical application, P, in line with the foundational aim of the workshop series. The rightmost columns are the domains of safety for HIT, Health Information Technology, as per Singh and Sittig's (2016) Health Information Technology Safety Measurement Framework. | | | | | | |
|  | **Challenges** | **Consequences for patient safety** | **Recommendations** | **Safe**  **HIT** | **Safe use of HIT** | **HIT to improve safety** |
| 1 | Difficult to conceptualise threats to patient safety from non-physical innovations. | Inadequate consideration of threats to patient safety. | Systems approach to conceptualising riskT; Safety casesP; Socio-technical perspectiveT | x | x | - |
| 2 | Unclear how to sensibly integrate and interpret new and voluminous data streams. | Missed opportunities to use data; Inappropriate use of data; Biased use of data. | Dynamic and causal modelling continuously surveilled for performanceP; Middleware for interoperabilityP; Standards for linkage and exchange of healthcare dataP; Automated anomaly detectionP | x | x | x |
| 3 | Reactive regulatory- and standards-based approaches to safety. | Avoidable harm is experienced before mitigations are put in place. | Gradual approval of medical devicesP; Systems approach to conceptualising riskT | x | - | - |
| 4 | Difficult to build and maintain trust in health information systems that are obscure and complex. | Misinformation and disinformation threaten patient safety. | Socio-technical perspectiveT; FactSheetsP | x | x | x |
| 5 | Emergent patient-safety consequences in health information systems. | Hazards cannot be completely foreseen. | Systems approach to conceptualising riskT; Systems approach to patient safetyT; Safety casesP; Socio-technical perspectiveT; Gradual approval of medical devicesP | x | x | - |
| 6 | Solutionism inappropriately simplifies problems and predicaments. | Unfit interventions and assurances are suggested. | Socio-technical perspectiveT; Systems approach to conceptualising riskT | x | - | - |

# Conclusion

To the authors’ knowledge, this is the first paper in the field of Patient Safety Informatics that has provided a definition and explored its rationale. The intention of this article was to begin the process of developing the theoretical and practical foundations of Patient Safety Informatics, answering calls for practical progress in safety science (Rae *et al.*, 2020) and a unifying theory (Swuste *et al.*, 2020). We presented six challenges posed by new and emerging digital health, described the consequences for patient safety, and recommended theoretical and practical mitigations. These challenges, consequences and recommendations were gathered at a workshop of health informatics researchers focused on exploring the theoretical and practical foundations of Patient Safety Informatics. Subsequent workshops in our series will address the consequences of contemporary safety theory for digital innovation, socio-technical evaluation of digital health, and digital health interventions designed to improve patient safety (Johnson *et al.*, 2020).

# References