Editorial: Evaluating new interactions in healthcare: Challenges and approaches

Rebecca Randell

Leeds Institute of Molecular Medicine, University of Leeds

Stephanie Wilson

Centre for HCI Design, City University London

Geraldine Fitzpatrick

Institute of Design and Assessment of Technology, Vienna University of Technology

Rebecca Randell

Leeds Institute of Molecular Medicine

University of Leeds

St. James's University Hospital

Leeds, LS9 7TF UK

r.randell@leeds.ac.uk

Introduction

New wireless, networked and sensor-based technologies are increasingly pervasive within healthcare. We are seeing a growth not only in the availability of hardware and software solutions for healthcare, but we are seeing them in a much wider range of settings than before. While previously IT was predominantly based in the consulting room, clinicians now have access to mobile devices such as personal digital assistants (PDAs), computers-on-wheels (COWs), and tablet PCs, and electronic whiteboard are becoming more and more popular within the hospital environment. These technologies mean that clinicians can access electronic information, such as electronic patient records (EPRs) and clinical decision support systems (CDSS) on wards, by the patient's bedside and during ward rounds. Healthcare technologies are also making their way into patients' homes, e.g. as telecare and assistive technology packages, and providing remote physiological and activity monitoring and remote consultations. These technologies do not just influence the distribution of information but allow for the creation of new processes for patient care.

Such changes present interesting opportunities for interaction design but also present new challenges for evaluation. Evaluating new healthcare technologies in a complex context of use such as the hospital or home presents practical challenges but evaluation strategies that fail to do understand how the technology is used *in situ* may not succeed in gauging the true impact of the technology (Wilson, Galliers, & Fone, 2007). While progress has been made in HCI in developing evaluation methods for a variety of challenging settings (e.g. Palen & Salzman, 2002), it is necessary to consider the particular needs of healthcare settings and how evaluation methods can be adapted to meet those needs.

Beyond these practical challenges is the more fundamental question of how to conduct evaluations which provide an assessment of the overall impact of technology in its context of use, while also recognising that any evaluation is undertaken at a point in time of

an ongoing adoption process where the very nature of the work being supported can evolve in relation to the technology as people become more familiar with it. User engagement and acceptance is an enduring problem for the introduction of healthcare technologies, suggesting a need for evaluation techniques that allow us to demonstrate a clear benefit to potential users while also empowering them to appropriate the technology within their own work setting. Within health informatics, clinical trials, particularly randomised controlled trials (RCTs), are often seen as the 'gold standard' and set expectations for very outcome-focussed rigorous measures. While clinical trials can provide important data on the impact on clinical outcomes, the trial protocol separates the technology from the social processes that surround its use and fails to address the issue of whether users will adopt the system into routine practice (Forsythe & Buchanan, 1991; Kaplan, 2001). Attention is restricted to effects anticipated at the time of trial design, so that important unanticipated effects, such as impact on interaction with the patient, may be missed (Moehr, Anglin, Schaafsma, Pantazi, & Grimm, 2006; Stoop & Berg, 2003). For example, previous evaluations of homecare technologies have focused largely on clinical outcomes (Department of Health, 2006) but ignored aspects of the lived experience of the technology and its social acceptability and fit into domestic life (Blythe, Monk, & Doughty, 2005).

Another evaluation approach that is used within HCI is quantitative empirical evaluations but lab based studies are more appropriate for settings with well-known tasks and outcomes; for evaluation of healthcare technologies, where tasks are complex and often involve multiple different healthcare professionals and patients often with multiple unpredictable health issues, how do we develop appropriate tasks and how can we judge their success? Further, new healthcare technologies may result in consequences more subtle than expected and difficult to capture quantitatively (Wilson, Galliers, & Fone, 2006). If findings from a lab based evaluation are to have relevance to the real world context of use, to what

extent is it necessary, or possible, to recreate features of a complex setting such as a hospital ward?

If the results of an evaluation are to inform wider implementation, it is necessary to not only know whether or not an application brings benefit but also to know the nature of the components and the specific context in which it was introduced. Traditionally, the results of clinical trials have failed to provide an understanding of why a system is effective or not, making it difficult to recreate success or prevent similar failures (Heathfield, Pitty, & Hanka, 1998). The introduction of healthcare computing applications involves a number of components – technological, clinical, social, organisational, professional – and these components are introduced into a particular context. It is necessary to understand how these components and the elements of the context impact the process that the technology is trying to support. The components of a healthcare technology intervention include the type of hardware, the functionality provided by the software, particular interface features, the physical configuration of the hardware, the aesthetic design of the device, the training provided and any reorganisation of healthcare provision that accompany the introduction of the technology. The elements of the context that potentially impact the introduction and use of the technology include the skill mix and motivation of the users, the physical setting – whether a hospital ward, a community setting or a patient's home, the acceptability of the technology and new organisation of service provision to both healthcare professionals and patients, the organisational culture, and the range of patients and conditions within the particular setting. How we develop an understanding of the impact of different components of a technology intervention and elements of the context seems like an important priority for the evaluation of new healthcare technologies.

Such understanding is also necessary if the results of an evaluation are to inform subsequent design. While RCTs, and indeed many HCI evaluation studies, are focussed on

outcomes as a summative evaluation, HCI also has a strong tradition of undertaking formative evaluations as part of a user-centred design process to help maximise the possibility of the new technology fitting the needs of the people it is designed to support. For healthcare, this becomes quite complex as clinicians are extremely busy and patients are ill requiring a duty of care with little scope for unproven experimentation even if it were considered ethically acceptable. Nonetheless there is a challenge to find some way to evaluate early design solutions, in as authentic a setting as possible, to help ensure the best technology to deliver the optimal clinical support.

Clearly, the issues are numerous and complex. A workshop at CHI'09 provided an initial opportunity to discuss these issues. Discussions around conducting evaluations in hospital settings highlighted the need to understand what aspects of a technology intervention are having an impact and the need to understand the impact of the technology on everyone within the setting rather than just those who directly interact with the technology, as well as practical challenges such as the difficulty of getting clinicians' time. There was acknowledgement of the benefit of naturalistic studies in hospital settings but also a desire to extend such methods. In discussing evaluations of homecare technologies, again practical concerns were raised, such as how to get participants sufficiently engaged. Questions were raised as to how we measure value for users. The notion of a broad range of users was highlighted again, acknowledging the need for a broad range of participants – not only patients and clinicians but also patients' carers and family. The workshop concluded with a discussion of 'key issues' that require further exploration. These included issues of how to incorporate the different values of the various stakeholders, where best to undertake evaluations and how close simulations need to be to the real world, how to study the longer term impact of healthcare technologies, the relevance of clinical trial methods and how they

might be adapted, how to identify the context in which a healthcare technology could provide most benefit, and whether there are parallels in other domains that can inform our approach.

A way forward

In this special issue, we present five papers that take on some of these issues and provide a fuller exploration of them. The special issue begins with a paper by Favela, Tentori and Gonzalez who, by presenting six evaluation studies across four technologies, provide an overview of methods for evaluating what they term as ubihealth technologies. The evaluation studies cover a range of technologies, settings and users: a context-aware mobile communication system, mobile technologies for recording clinical data for medical and nursing staff, and a persuasive virtual community to support weight loss. They classify the evaluations using two criteria: the extent to which the evaluation is conducted under realistic conditions, i.e. the evaluation's ecological validity, and the degree of integration of the technology with the environment. The studies that they present range from scenario-driven evaluation and theatrical representations to simulations to *in situ* evaluations. The evaluation grid that Favela et al present provides a framework that can be used by researchers to select appropriate techniques as a function of the technological and environmental complexity.

The paper by Dahl, Alsos and Svanæs then explores in more detail the design of simulations for use in evaluation, drawing on research on training simulations to identify elements of the context to be replicated within the simulation. They identify a set of fidelity dimensions and explore the relationship between fidelity configuration and the type of feedback received, drawing on evaluations that they have conducted of technologies to be used on hospital wards. The evaluations explore handheld devices for inputting clinical data and sensor-based technologies for automatic patient identification during medication administration. The authors show how physical features, such as the room layout and furniture, act as cues helping the participants to relate the design solution to their everyday

work. Such features help to explore physical and bodily aspects of usability and ergonomic suitability. They suggest that, rather than incorporating the highest possible fidelity across multiple dimensions, a more feasible approach is to carefully select which aspects of the context to accurately replicate and which aspects to simplify or remove. They provide a framework to guide researchers as they design healthcare simulations for evaluation purposes.

The paper by Tang, Carpendale and Scott also encourages us to think about the relevant elements of the context, although this time for *in situ* evaluations rather than simulations. Based on a study of nurses' information flow in a hospital setting and related literature, they present the InfoFlow framework. This is made up of six interrelated factors that affect information flow: information, personnel, artefacts, spatiality, temporality and communication mode. This framework can be used by researchers to guide the collection and analysis of data for in situ evaluations. The authors describe how the framework was used in the evaluation of a mobile voice communication system. The framework is potentially relevant to the evaluation of a wide range of healthcare technologies, as information flow is the focus of many health IT development projects, such as electronic patient records, tools to support clinical handover, and large displays for supporting awareness.

The paper by Stevenson, Hutchins and Smith presents the results of a pilot trial of a broadband telehealth system. The trial was set up with selected patients and was run as a simulated distributed environment, where the relevant parties were in different locations on the one hospital campus. It has relevance for researchers wishing to design studies for the evaluation of telehealth systems. For example, they emphasise the need for success criteria that are removed from more traditional clinical end-point measures but rather reflect the quality of the supported interaction. However, the paper has broader relevance, encouraging researchers to use a pilot trial to determine success criteria as understood by different groups

of participants involved in use of a system, including the extended networks around the patient, and to consider how use of the technology changes over time. More generally, the study provides an example of how to combine qualitative and quantitative data for the evaluation of healthcare technologies.

In the final paper, by Storni, we move away from the hospital setting to consider the evaluation of self-monitoring technologies, used by patients within their homes. While focusing on the appropriation of home care technologies, Storni raises and begins to tackle important issues that have relevance for the evaluation of new healthcare technologies more generally. For example, he explores the dangers inherent in attempting to separate out the elements of the context that influence how a technology is used rather than seeking to understand the relationship between those elements, and how we move from proving the benefits or otherwise of a technology to actually attempting to improve the technology. Storni explores the issue of who we evaluate for and how we incorporate a range of perspectives within an evaluation study, and again we are reminded of the need to include the extended networks around the patient.

Lessons learned

The papers in this special issue describe the complexities of the settings in which healthcare technologies are used and the particular challenges these present for evaluation. They consider technologies that are used in the home, in the community and in the hospital. The studies of hospital settings consider both inpatient wards and outpatient clinics. A variety of technologies have been considered: mobile tools for accessing and recording patient data and for communicating with colleagues, teleconferencing technologies, and self-monitoring technologies and online tools for patients.

None of the papers stop at describing the challenges for evaluation, instead presenting a range of ways forward. Taken together as a body of work, we see the papers as providing

guidance, many in the form of frameworks, on how to undertake the evaluation of new healthcare technologies. While they describe a whole range of evaluation methodologies, they give guidance on the context in which each method may be best applied and for what purpose. They give guidance on the components of a technology intervention that need to be considered, such as the physical environment, technology and task, whether evaluating technology use *in situ* or *in replica*. They remind us of the need to pay attention to how use changes over time. They also remind us of the different roles that an evaluation study can play, not only determining whether a system should be seen as successful but also identifying success criteria, predicting user acceptance, identifying areas for design improvement, and understanding particular issues such as how perceptions of privacy may influence use of system.

The papers presented here variously point to ways in which the RCT model may be inappropriate or impossible to set up for the evaluation of healthcare technologies. Stevenson et al. describe the difficulties of using a more conventional trial design in evaluating telehealth, such as the logistical problem of 'matching' patients, absence of blindness to the intervention, and the inappropriateness of random allocation and a rigid trial protocol. They also highlight the benefits of the observational data that they collected, allowing them to capture unanticipated uses of the technology and users' reactions to it. Adaptations to the RCT model for the evaluation of complex interventions, where a combination of quantitative and qualitative methods are recommended in order to understand not only the impact on clinical outcomes but also to understand unanticipated effects and to understand the impact of different components of the intervention and elements of the context, certainly have relevance to the evaluation of new healthcare technologies but they do not go far enough. Storni argues that rather than separating out the components of an intervention, we should

seek to understand the relationship between them, acknowledging the complexity rather than trying to break it down.

The papers in this special issue demonstrate the particular contribution that HCI can make to the evaluation of healthcare technologies. We see HCI evaluations of new healthcare technologies as being those that make use of both quantitative and qualitative methods as is appropriate for the aims of the evaluation, that consider the context in which a technology is used whether the evaluation is *in situ* or *in replica*, that consider the perspectives of a broad range of participants surrounding the use of a technology, that are concerned with impact on quality of patient care but which also allow assessment of the overall impact of technology, and that contribute to the refinement of those technologies through design.

We consider that the question of how we can conduct robust evaluations of the impact of healthcare technologies on clinical outcomes while also undertaking an assessment of the overall impact of technology and understanding the reasons for success or failure should be a priority for HCI researchers working within the healthcare domain. The papers presented in this special issue suggest a number of steps towards this but further exploration is needed, particularly in relation to scale and time. The studies presented here are largely smaller-scale and/or early phase evaluations. There is a need to understand how these methods and frameworks scale up to technology deployments across a whole hospital campus or health delivery area. There is the challenge of understanding how new healthcare technologies become integrated with other existing technologies and the impact of increasingly complex technological arrangements, rather than just focussing on the single new system. There is also the challenge of understanding the integration between home and hospital and other settings, as currently many evaluations focus on one setting or the other.

The studies here also represent varying time scales of use. It is well recognised that there is a process of appropriation and change as technologies become embedded into

organisational, spatial, social and professional practices, each of these co-evolving with the technology. Evaluation studies are needed that seek to understand the nature of these longer term appropriation processes and how these relate to ongoing clinical outcomes. More than that, something of a culture change is needed, where the complexity of *in situ* use of healthcare technologies is not only acknowledged but embraced, while at the same time rigorous attempts are made to build an evidence base that demonstrates that new technologies deliver improvements in health outcomes and are acceptable, usable and useful for patients and care providers.

References

Blythe, M., Monk, A. F., & Doughty, K. (2005). Socially dependable design: The challenge of ageing populations for HCI. *Interacting with Computers, 17*(6), 672-689.

Department of Health (2006). *Supporting self care - a practical option: Diagnostic, monitoring and assistive tools, devices, technologies and equipment to support self care*.

Forsythe, D. E., & Buchanan, B. G. (1991). *Broadening Our Approach to Evaluating Medical Information Systems.* Paper presented at the Symposium on Computer Applications in Medical Care.

Heathfield, H., Pitty, D., & Hanka, R. (1998). Evaluating information technology in health care: barriers and challenges. *British Medical Journal, 316*, 1959-1961.

Kaplan, B. (2001). Evaluating informatics applications-some alternative approaches: theory, social interactionism, and call for methodological pluralism. *International Journal of Medical Informatics, 64*, 39-56.

Moehr, J. R., Anglin, C., Schaafsma, J., Pantazi, S., & Grimm, N. (2006). Lest Formalisms Impede Insight and Success: Evaluation in Health Informatics. *Methods of information in medicine, 45*, 67-72.

Palen, L., & Salzman, M. (2002). *Voice-mail diary studies for naturalistic data capture under mobile conditions.* Paper presented at the Proceedings of the 2002 ACM conference on Computer supported cooperative work, New Orleans, Louisiana, USA.

Stoop, A. P., & Berg, M. (2003). Integrating quantitative and qualitative methods in patient care information system evaluation: guidance for the organizational decision maker. *Methods Inf Med, 42*, 458-462.

Wilson, S., Galliers, J., & Fone, J. (2006, November 4-8, 2006). *Not All Sharing Is Equal: The Impact of a Large Display on Small Group Collaborative Work.* Paper presented

at the Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work, Banff, Alberta, Canada.

Wilson, S., Galliers, J., & Fone, J. (2007). Cognitive Artifacts in Support of Medical Shift Handover: An In Use, In Situ Evaluation. *International Journal of Human-Computer Interaction, 22*(1), 59 - 80.