Clinical Proof-of-Concept – A Evaluation Method for Pervasive Healthcare Systems

Jakob E. Bardram
IT University of Copenhagen
Rued Langgaards Vej 7, DK-2300 Copenhagen, Denmark
bardram@itu.dk

ABSTRACT

Pervasive Healthcare – i.e. designing pervasive computing technologies for healthcare usage – is an especially promising area within pervasive and ubiquitous computing research. However, it is extremely difficult to evaluate such systems because establishing clinical evidence for medical benefits would require longitudinal, randomized, double-blind, placebo-controlled trials involving a homogeneous patient population and medical condition. This would not only require huge resources in terms of clinical staff and patient participation, but would also require the technology to be fully developed and ready for large scale use. The latter is simply not feasible when doing technological research into new types of pervasive healthcare technologies. In this paper, I suggest the method of 'Clinical Proof-of-Concept' as a method for evaluating pervasive healthcare technologies in order to establish the clinical feasibility of the technology before entering large-scale clinical trials. The method has been applied in a couple of cases and I report on lessons learned from this.

INTRODUCTION

Applying Ubiquitous and Pervasive Computing technologies for healthcare purposes is gaining increasing interest and is growing into a research field of its own called 'Pervasive Healthcare' [3, 1]. The research questions and the technologies being investigated within Pervasive Healthcare are quite diversified ranging from using biomedical sensor technology for patient monitoring and prophylactic treatment, to mobile and context-aware systems inside hospitals. Pervasive Healthcare has a lot in common with established medico-technical research areas like biomedical engineering, medical informatics, and telemedicine, but is distinct in its fundamental approach and goals; pervasive healthcare systems are often designed for patients and not for clinicians, and they embody technologies growing out of the ubiquitous computing research, including sensor technology, context-aware and mobile computing, large interactive displays, etc. Similar to Ubiquitous Computing research, Pervasive Health-

care research is specifically targeted towards technology – i.e. aiming at understanding, designing, building, and testing new types of pervasive computing technologies for healthcare purposes.

A common methodological approach to ubiquitous computing research is to design and implement a technical 'Proof-of-Concect' for a proposed new ubiquitous computing technology or application, and subsequently evaluate this implementation in a limited setup. Marc Weiser defined the concept of a technical Proof-of-Concept as:

> *The construction of working prototypes of the necessary infrastructure in sufficient quality to debug the viability of the system in daily use; ourselves and a few colleagues serving as guinea pigs.* [8].

Looking at the research questions posed by pervasive healthcare, this research approach seems to be lacking some rigor in order to investigate whether the technology solve health related challenges. We would, for example, never be able to understand or evaluate to which degree a technical prototype for elderly people would be successful, if it is only tried out by our colleagues in a research laboratory.

From a medical perspective a technical proof-of-concept is not acceptable for introducing new medical technology or treatment. In most healthcare systems, clear clinical evidence needs to exist before a new medical technology is put into use for patient treatment. *Evidence-based medicine* [7] is the clinical methodological approach for establishing this evidence. Evidence-based medicine categorizes different types of clinical evidence and ranks them according to the strength of their freedom from the various biases that beset medical research. The strongest evidence for therapeutic interventions is provided by systematic review of randomized, double-blind, placebo-controlled trials involving a homogeneous patient population and medical condition. In contrast, patient testimonials, case reports, and even expert opinion have little value as proof because of the placebo effect, the biases inherent in observation and reporting of cases, and difficulties in ascertaining who is an expert.

Such strong evidence is, however, impossible to obtain while we are still in the research and development phase of new technology. So an important question is how we can strike a balance between these two extremes; design and implement limited technical proof-of-concepts which at the same time

are suited to provide sufficient clinical evidence for further research and development

In this paper, I suggest a methodological approach called 'Clinical Proof-of-Concept' which is aimed at creating initial clinical evidence for the medical benefits of a pervasive healthcare technology. This approach has been used in a couple of cases and I will report on these cases, how the clinical proof-of-concept was carried out, and what we learned from these cases.

The contributions of this paper is the presentation of the methodological approach of a Clinical Proof-of-Concept together with specific examples of its use in two cases. By suggesting this approach, it is my aspiration that more pervasive healthcare technologies can be subject to initial evaluation and scrutiny before entering large scale clinical trials, while at the same time actually being put into test in a limited real-world deployment. Using this approach, a more incremental and experimental approach to the construction of pervasive healthcare technologies can be achieved, which in the end will lead to developing more appropriate and usable pervasive healthcare technologies. At the same time, the approach would enable us to reject and dismiss those technologies, which show little clinical promises before large amount of resources are spent on developing the technology and running clinical trials.

CLINICAL PROOF-OF-CONCEPT

To rephrase the definition from Marc Weiser, I am defining a Clinical Proof-of-Concept (CPoC) as:

The construction of working prototypes of the necessary functionality and infrastructure in sufficient quality to investigate evidence for improving health in daily use for a suitable period of time; a limited but relevant set of people serving as subjects.

More specifically, the technology should be a working prototype that is usable (but not necessarily user-friendly), works on its own, and is focused on addressing specific research questions. This technology should be deployed in a real clinical setup, should be used by real users (researchers are hands-off), for a short, but sufficient period of time, which – depending on the research question – may range from 1 day to 3 months.

For example, you may want to test a system for monitoring hypertension and evaluate if users are able to control their own blood pressure over time, thereby reducing hypertension, which again – according to medical literature – have a positive effect on a wide range of heart diseases. In this case, a CPoC would involve a technical prototype which runs on its own and is able to monitor blood pressure, but it may not be particular secure, robust, or integrated in a country-specific healthcare system. It should, however, be able to run with limited interference from the researchers, while some 'Wizard-of-Oz' techniques may be applied. The deployment would include a limited amount of people – e.g. 10 – which is not statistically significant for hard medical evidence, but

sufficient for establishing the viability of the technical setup and its use in a real-world deployment. And the trial may run for a couple of weeks – rather than the months normally required for a clinical trial.

The methods used during this CPoC should be targeted at collecting evidence, which demonstrate that the technology seems promising in addressing its specific goal. It may be relevant to gather initial clinical evidence for the medical benefit of the technology. For this purpose, trying to measure some clinical effects is essential during the CPoC. For example, in order to establish any clinical effect in the monitoring of hypertension, blood pressure data may be compared over the time span of the CPoC and questionnaires regarding the patients' awareness and handling of their blood pressure may be issued and analyzed.

Even though the clinical evidence may be biased by different factors and hence not as strong as would be required in Evidence-Based Medicine, providing initial clinical evidence for the working of the technology is still essential in order to justify further development and evaluation. Furthermore, the Clinical Proof-of-Concept may simultaneously work as a 'dry-run' for testing the data collection methods, which later are to be used during the clinical trial. For example, if a questionnaire is handed out to the participants, this questionnaire and the timing of it may be subject to change based on experiences obtained during the clinical proof-of-concept.

Apart from establishing initial clinical evidence, a core purpose of a CPoC is to investigate the usefulness and usability of the proposed solution. To a certain degree, I would argue that this is the *main* purpose of a CPoC for two reasons. Firstly, the clinical benefit of a pervasive healthcare technology may be significantly diluted if the technology is hard to use for the patient. For example, it is obvious that if the blood monitoring technology is hard to use, then limited effect on hypertension management may be found during the clinical trial. Secondly, it is essential to catch and remedy such usability problems in an early phase before resources are invested in developing the technology, producing it in large numbers, and deploying it for a clinical trial.

These arguments may seem trivial. However, some sort of usability problems are often hard to discover and are often unexpected. By running a CPoC which actually puts the technology to a test in a real-world setting with real users for a certain period of time, many of the more complex usability problem may surface. And often, ideas for changing and improving on the technology arise when seeing it in actual use and by working closely together with the users to find a solution to the problem.

Methods for usability inspection would typically be qualitative in nature, involving observations, questionnaires, and studies of perceived usefulness and usability.

Figure 4 show the temporal progression of research methods as the technology is developed and mature. Time-wise,
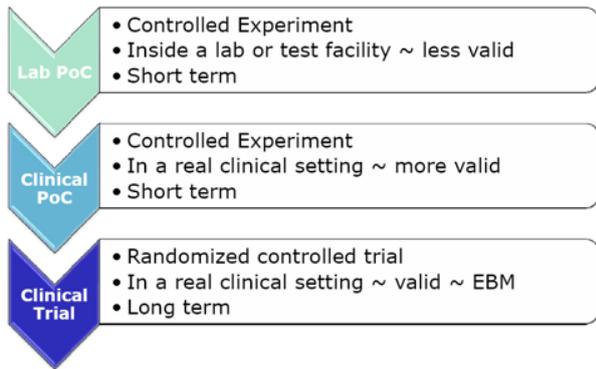
Figure 1. The timing of a Clinical Proof-of-Concept is between a laboratory proof-of-concept and a full clinical trial.

a clinical proof-of-concept lies in between the technical laboratory proof-of-concept and a full-scale clinical trial.

CASES

In order to illustrate how a CPoC can be used in pervasive healthcare research, I will use two cases as examples. The first case is concerned with home-base blood pressure monitoring, and the second case is concerned with developing context-aware technologies for improving patient safety inside the operating room. These two case are very distinct in many respects – technology, users, deployment settings, and goals – but as such, they illustrate the breath of the CPoC approach.

Blood Pressure Monitoring

The first project was concerned with home-based monitoring of hypertensive patients. Hypertension is a direct cause of a number of heart diseases, including congestive heart failure and stroke, and substantial clinical evidence indicates, that frequent blood pressure monitoring helps prevent hypertension [6]. For this reason, many pervasive healthcare projects have addressed hypertension. This project was done in 2002 when state-of-the-art blood pressure monitoring was based on a cuff. Our goal was to deploy the technology in a limited pilot study and perform a CPoC (even though we did not call it that at that time). The technology for home-based monitoring consisted of a suitcase with a traditional blood pressure monitor, a PDA, and a GSM modem as shown in Figure 2.

In this project, the suitcase were given to the patients by their general practitioner (GP). The system had three main features: (i) it allowed the patient to measure the blood pressure several times a day and this data was sent to a central server for the GP to observe; (ii) the GP could prescribe medicines and the patient could indicate that (s)he was complying to the prescription; and (iii) it enabled communication between the patient and the GP, using both text and voice messaging.

During the first months of a longer deployment period, we carried out a series of interviews and field studies of this home-based monitoring and treatment system. This study

Figure 2. A patient using the home-based monitoring system in a briefcase for monitoring her blood pressure.

was focusing on studying issues of medical treatment, division of work, communication, patient self-understanding, and the technology in actual use [2]. Our study – in accordance with most medical studies of home-based monitoring of hypertension – gave evidence that this kind of blood pressure monitoring provides more accurate measurements. Our findings, however, also revealed that the relationship between the GP and the patient changed when this new computer-mediated home-based treatment for hypertension was introduced. More specifically, we found four specific aspects of this transformation caused by pervasive monitoring and treatment technology:

- A new division of work emerged, which transferred the act of monitoring and interpreting the blood pressure data from the GP to the patient.

- The medical treatment of hypertension and the life quality of the patient was improved. However, new demands for monitoring the incoming data and the patient's progression in treatment were inflicted upon the GP.

- The communication pattern between the patient and GP was fundamentally changed from a contextual rich conversation to an asynchronous message exchange.

- Because the patient was more involved in the monitoring and treatment of hypertension, he or she became more self-aware on the nature of high blood pressure and what affects it.

This clinical CPoC was insufficient to establish clinical evidence for improved hypertension treatment of the patients. For this purpose, the time frame of the study was too short, the sort of methods applied insufficient for determining clinical evidence, the number of patients were too small, and no control group was involved. The study, however, were sufficiently large to study, understand, and argue that this kind of home-based monitoring would transform the patient-GP relationship and make the patient capable of managing their own blood pressure in a more efficient way. And since previous clinical studies have shown that regular self-conscious

attention to your blood pressure reduces the risk of hypertension, this was clearly a strong indicator that this kind of technology would be useful.

At the same time, the CPoC revealed a series of usability and deployment issues which needed to be looked into before running larger scale trials involving a larger amount of patients and GPs. The technology were subsequently improved and deployed in a large clinical trial with 10 GPs, 120 patients, and a control group.

Context-aware Patient Safety

The Context-aware Patient Safety and Information System (CAPSIS) monitors what is going on inside the operating room and use this information to show timely medical data to the clinicians, and to issue warnings if any safety issues are detected [4]. CAPSIS monitors events like the status of the operation; the status and location of the patient; location of the clinicians who are part of the operating team; and equipment, medication, and blood bags being used inside the operating room. This information is acquired and handled by a context awareness infrastructure, and a special safety service is used for overall reasoning which actions to take or warnings to issue. The goal is to supplement human safety vigilance with a machine reasoning counterpart.

CAPSIS was deployed and tested in a CPoC where it was used for one day by a full surgical team performing simulated operations inside an operating room. In total, 8 operations were executed during the day, involving both operations with no warnings as well as different types of warnings, including wrong patient, wrong operating table, wrong blood, and incomplete team. In addition, medical records, radiology images, and the operation checklist were presented on displays using the context-aware triggers. A picture from the CPoC is shown in Figure 3. Everything were done exactly as real surgeries, except that no real patients were involved and the acting patients were not sedated or actually cut. The acting patients, were, however treated as any real patient, including being admitted to ambulatory surgery and scheduled in the scheduling system.

The goal was to provide objective measurements on the usefulness and usability of our design while, at the same time, investigate the detailed user reaction to the system and the user interface in a more qualitative fashion. For this purpose, we used a *multi-method evaluation setup* where we (i) asked the users to perform the operations while thinking aloud, (ii) investigated perceived usefulness and usability based on a questionnaire [5], and (iii) made a semi-structured follow-up interview.

Based on this evaluation, the clinicians concluded that the system would be very useful for ensuring patient safety and was very easy to use. Most of the patient safety issues monitored by CAPSIS were found to improve patient safety, and several of the findings resonate with the recommendations from the state-of-the-art regarding patient safety. Moreover, the CPoC revealed a series of usability issues which we had not captured previously, despite several prototyping ses-

Figure 3. The deployment of the system inside the OR; the surgeon and the sterile nurse read medical data on the screen to the right while the scrub nurse interacts with the patient safety system on the screen to the left.

sions. By actually deploying the technology inside the OR, and asking the operating team to use the technology during close to real-world surgeries, a wide range of issues surfaced which would have not been found otherwise. Especially issues regarding the physical working environment of an OR and the tight teamwork taking place during surgery surfaced. Some examples of issues that were discovered during this CPoC include;

- The user interface had to be improved in several place, including issues like coloring, highlighting, and font size on the screen due to the distance from the operating table to the screens.

- The procedures regarding attaching a RFID enabled armband to the patient needed to be scrutinized because patient safety now was depending on that this was done correctly. If the wrong armband was attached to a patient, unpredictable and potential severe safety hazards may occur.

- Better support for handling and registering scanning of blood bags were needed. When moving from a limited test in a lab to a CPoC in the OR where a substantial volume of blood may be needed, the existing method for checking correct blood did not scale to e.g. 10-20 blood bags. The reason for this was due to a number of highly interlinked aspects, ranging from the organizational procedure for ordering and getting blood, to the physical layout of the OR, and the way the RFID technology were working.

- Lack of triangulation, which is the medical safety term for ensuring that a safety check is done by combining the patient, the procedure, and the clinical staff. Even though this was part of the overall systems design, triangulation did not work inside the OR on the individual level.

- The operating team had to change their safety procedures just before surgery in order to leverage the capabilities of the system.

It is important to note, however, that this CPoC setup is not providing 'hard clinical evidence' for improved patient safety inside the OR. We do not know if this system – if build and deployed – would improve on patient safety. This would require a randomized clinical trial over a longer period of time involving a control group, which again would require a full working system ready for large-scale and long-time deployment. Providing such Evidence-Based Medicine is, however, not the purpose of a Clinical Proof-of-Concept; it is rather to investigate the feasibility of the proposed solution for further development. By asking the involved clinicians how they perceive the system's potentials for improving patient safety, we are given sound indications regarding the feasibility and directions for further development.

Beside this indication of potential clinical evidence for improving patient safety, the core benefit from running this CPoC is the different problematic issues regarding the current prototype which must be addressed before making a larger clinical trial. As illustrated above, these issues are a mixture of technical, usability, physical, and team-oriented aspects which need to be addressed in concert. But most importantly, these complex and interrelated issues would probably never have been found without running a CPoC. The next step would be to incorporate the suggestions for improvement and then apply more rigorous clinical methods for evaluating the degree to which the system improves on patient safety. Note, however, that the only reason for such an investment is based on the fact, that the CPoC indicated that the system potentially would improve patient safety.

DISCUSSION

What have we learned from our use of clinical proof-of-concepts so far?

First of all, a CPoC reveals a wide range of technological problems and issues. For example, in the blood pressure monitoring project, the CPoC revealed all sorts of problems with wires, handling software updates, and sustaining power on the devices while not in use. In the patient safety project, the CPoC revealed all sort of issues ranging from the working of the RFID technology to the use of the software on large touch-screens. Hence, a CPoC is useful in determining the sort of technological issues which are related to real-world use by real users for a longer period of time and on a larger deployment scale.

Second, even though a CPoC seldom is done in a way which justify any 'hard' clinical evidence, it is still useful in order to both provide initial clinical evidence for a potential medical effect, as well as providing important information on how this clinical effect should be collected. For example, the clinicians in the patient safety project unanimously agreed that the system had the potential for improving patient safety inside the OR. This do not count as clinical evidence, but nevertheless it encourage further development. At the same time the CPoC revealed that the methods used for evaluating the system were appropriate for judging perceived usefulness, but they were not appropriate for providing clinical evidence for the improvement of patient safety

during surgery. Hence, a new methodological setup is required in any subsequent clinical trial.

Third, because the technology is deployed in a real setting for a non-trivial period of time, a CPoC is well-suited for investigating the usability of a pervasive healthcare system. Especially non-trivial usability problems which arise from complex interaction between different types of technologies, users, real deployment settings, and long-term use may be discovered during a CPoC. For example, the blood-pressure monitoring CPoC revealed that it was hard for some patients to type a message to the GP and this functionality was hence changed to use voice messages instead of text messages. And in the patient safety project, as wide range of usability issues regarding the user interface were found.

Fourth, due to the real-world deployment a CPoC helps reveal and evaluate the physical usage of the technology. The physical aspects of the technology is especially important for pervasive healthcare systems since medical devices and systems are notoriously tied to monitoring or influencing physical properties of a human body or a physical environment in homes or hospitals. For example, a wide range of issues regarding the physical handling of the blood-pressure cuff and the handling of the PDA were revealed during the CPoC. This subsequently lead to the design of a cartoon-like step-by-step instruction card, which were place on the front of the suitcase, as shown in figure 2. In the patient safety project, the physical layout of the OR, the physical handling of patients, blood bags, and instruments turned out to have significant impact on the use of the system.

Finally, often pervasive healthcare systems needs to exist and work in a larger social and organizational context. A CPoC is equally suited for initial investigation of the impact arising from this larger deployment context. Especially in the blood-presure project we found a significant change in the division of work and communication between the GP and the patient, and the CPoC revealed some of the important details of how the technology would potentially influence the way the treatment of hypertension were achieved. In the patient safety project, the CPoC helped judge the fit between the system and the complex and dynamic teamwork taking place inside an OR.

CONCLUSION

In this paper, I have proposed to apply a Clinical Proof-of-Concept as a methodological approach for evaluating pervasive healthcare systems. A CPOC involves a focused study in a real-world deployment setting, involving real patients and users, while being limited in time, scope, and clinical rigorousness in the methods applied. By being a stepping stone in the middle of a laboratory-based evaluation and a full-scale clinical trial, the CPoC is able to provide valuable information regarding the clinical applicability of the system, its usability, and issues regarding the physical and organizational deployment of the system. In this way a CPoC is a more dedicated and cost-effective approach for establishing initial clinical evidence as well as being a invaluable source for improving the technology at a stage before resources are

invested in final development and clinical trials.

REFERENCES

1. J. E. Bardram. Pervasive healthcare as a scientific dicipline. *Methods of Information in Medicine*, 3(47):129–142, 2008.

2. J. E. Bardram, C. Bossen, and A. Thomsen. Designing for transformations in collaboration: a study of the deployment of homecare technology. In *GROUP '05: Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, pages 294–303, New York, NY, USA, 2005. ACM Press.

3. J. E. Bardram, A. Mihailidis, and D. Wan, editors. *Pervasive Healthcare: Research and Applications of Pervasive Computing in Healthcare*. CRC Press, 2006.

4. J. E. Bardram and N. Nørskov. Designing Context-aware Safety Systems for the Operating Room. In *Proceedings of Ubicomp 2008: Ubiquitous Computing*, Seoul, Korea, Sept. 2008.

5. F. D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3):319–339, September 1989.

6. M. A. M. Rogers, D. A. Buchan, D. Small, C. M. Stewart, and B. E.Krenzer. Telemedicine improves diagnosis of essential hypertension compared with usual care. *Journal of Telemedicine and Telecare*, 8:344–349, 2002.

7. D. L. Sackett, W. M. Rosenberg, J. A. Gray, R. B. Haynes, and W. S. Richardson. Evidence based medicine: what it is and what it isn't. *BMJ*, 312(7023):71–2, 1996.

8. M. Weiser. Some computer science issues in ubiquitous computing. *Communications of the ACM*, 36(7):75–84, 1993.