Using Smart Mobile Devices for Collecting Structured Data in Clinical Trials: Results From a Large-Scale Case Study

Johannes Schobel, Rüdiger Pryss, Manfred Reichert
Institute of Databases and Information Systems
Ulm University, Germany
{*johannes.schobel, ruediger.pryss, manfred.reichert*}@*uni-ulm.de*

***Abstract*—In future, more and more clinical trials will rely on smart mobile devices for collecting structured data from subjects during trial execution. Although there have been many projects demonstrating the benefits of mobile digital questionnaires, the scenarios considered in literature have been rather limited so far. In particular, the number of subjects is rather low in respective studies and a well controllable infrastructure is usually presumed, which not always applies in practice. This paper gives insights into the lessons learned in a clinical psychology trial when using tablets for mobile data collection. In particular, more than 1.700 subjects have participated so far, providing us with valuable feedback on collecting trial data with smart mobile devices in the large scale. Furthermore, issues related to an insufficient infrastructure (e.g., unstable Internet connections) have been addressed as well. Overall, the paper provides valuable insights gained during trial execution. In future, electronic questionnaires executable on smart mobile devices will replace paper-based ones.**

***Keywords*-Mobile Data Collection; Electronic Questionnaires; Smart Mobile Device; Clinical Trial**

I. Introduction

Paper-based questionnaires constitute an established as well as cheap method for collecting data in clinical trials. However, before analyzing the collected data, an error-prone manual transcription to electronic worksheets is required. Besides transcription errors, other data quality issues emerge like, for example, incomplete or inconsistent data.

Due to the increasing dissemination of smart mobile devices (e.g., smartphones or tablets), therefore, clinical trials will increasingly rely on mobile electronic questionnaires for collecting trial data as well as for addressing the aforementioned issues. To better understand the requirements that must be met when using electronic questionnaires on smart mobile devices, we developed several mobile applications and used them in the context of different psychological trials [1], [2], [3].

Specifically, this paper refers to a clinical psychology trial conducted in Burundi. For many years, Burundi has been a staging ground for conflicts whose survivors were often left behind with post-traumatic illness. To prove the need for treating post-traumatic related symptoms (e.g., PTSD), we assisted clinical psychologists in conducting a large-scale trial in Burundi during the last years. Particularly, we implemented a mobile application for standardized data collection running on smart mobile devices.

This mobile application as well as its underlying technology have shown a high reliability during trial execution. Furthermore, collecting data with electronic questionnaires on smart mobile devices saved huge efforts on the side of the researchers, while at the same time increasing data quality.

Moreover, the paper provides insights into the use of smart mobile devices in the context of a large-scale trial. We strongly believe that respective information is useful when developing mobile data collection applications for clinical trials in general. Altogether, the contributions of the paper are as follows:

- The paper presents a large-scale trial from the field of clinical psychology, in which trial data was collected with tablets (i.e., iPads) in a very challenging environment.
- The paper discusses the lessons learned from the use of smart mobile devices and refers to challenging issues that emerged during the project.

The remainder of the paper is structured as follows: Section II discusses related work. Section III first presents the large-scale trial from clinical psychology and then introduces the smart mobile application we developed and used in the context of this trial. Section IV discusses the lessons we learned when developing, deploying and running mobile applications. Further, it presents relevant issues for mobile electronic questionnaires applied in clinical trials. Finally, Section V concludes the paper and presents an outlook on further research topics.

II. Related Work

The use of tablet computers for recruiting and screening elderly osteoporosis patients is discussed in [4]. In particular, this work compares the use of a mobile application with a voice response system (for phones). Out of 160 subjects, 93 were assigned to the tablet application and 67 patients to the voice response system. When presenting screening questions (e.g., asking whether the patient received an X-ray examination before), the patients using the tablet application were able to complete the questionnaire on their own (100%).

However, only 46 patients (69%) using the voice response system were able to complete the questionnaire.

In a similar setting, researchers compared a smartphone application with an SMS text-only implementation for screening patients [5]. The overall results were similar as in the aforementioned study. However, most of the participants emphasized that the smart mobile application was much easier to use.

Another case study dealing with the use of a mobile health application that runs on Android devices is presented in [6]. The goal of this application was to replace the paper-based charting during ward rounds. The study was conducted in a hospital, where 8 nurses were equipped with mobile devices. All nurses had 10+ years working experience and were familiar with common features of smart mobile devices. After using the application for 3 weeks, a survey was handed out to obtain feedback. In turn, survey results indicate that the application was easy to use. Furthermore, medical staff reported on problems related to the graph-based visualizations. The diagrams were perceived as too large, including too many data-points; i.e., users had to scroll and zoom to find specific values. Obviously, these problems are related to the small screen sizes compared to regular computers displays.

The project 10.000 Steps uses smartphones to monitor the activity of their members [7]. Overall, 50 subjects had to track their steps using either the website or the provided mobile application. As a conclusion, the researchers emphasize that the mobile application provides a good way to collect data. They further discuss limitations of the study design, especially regarding the number of participants.

The study presented in [8] compares self-monitoring with coaching mobile applications for tracking exercise plans of patients during their rehabilitation.

All case studies emphasized that the mobile applications were well accepted by the subjects; i.e., (elderly) patients or medical staff. However, these studies are limited in respect to the number of participating subjects compared to the trial described in this paper. Besides this, all these studies took place in a well controllable environment and were based on a stable infrastructure (e.g., a stable Internet connection and a support team nearby). When targeting at mobile data collection in real-world scenarios, however, the lack of a stable infrastructure as well as staff untrained in the use of mobile applications are common. In addition, most mobile applications described in literature were solely designed for research purposes and not intended for being actually used in a real-world settings (e.g., in clinical ward rounds [9]).

III. Case Study

This section presents a real-world trial from clinical psychology that deals with traumatic events. In addition, we introduce the iOS tablet application we developed and used in this trial for collecting clinical data. In this context, we also discuss important aspects that emerged when implementing the mobile application.

A. Burundi Case Study

We refer to a trial our partners from clinical psychology conducted in Burundi (i.e., Africa). In this trial, clinical psychologists interviewed ex-combatants and soldiers to explore traumatic events before and after military missions. This trial was planned to run over several years. Hence, it has been divided into 5 consecutive phases, which are summarized in Figure 1. Note that these phases not only differ in respect to the questionnaires used for collecting data, but also the participating subjects (e.g., veterans or soldiers).

German psychologists started with planning the trial in January 2012. To investigate the Post-Traumatic Stress Disorder (PTSD) of the subjects (ex-combatants as well as AMISOM soldiers), psychologists had designed a questionnaire by merging various psychological questionnaires (e.g., PSS-I [10] or AAS [11]). During this initial planning phase, the idea to support trial data collection with smart mobile devices arised. Note that this was mainly motivated to the numerous problems coming with paper-based questionnaires in comparable studies of the same group of researchers.

We implemented a mobile application running on an iPad tablet. It covers the questionnaires of the psychologists as required for the different phases of the trial. The development of the first prototype took about 3 weeks (one programmer). This prototype already met the most relevant requirements, like data encryption or easy to use interfaces (cf. Section IV for additional details).

During the first phase of the trial about 460 subjects (i.e., ex-combatants and veterans) were interviewed. Due to the large number of questions (i.e., up to 500 questions per phase) as well as the rather large number of subjects, a team of 12 international interviewers was formed (5 German psychologists, 1 Burundian psychologist, 6 Burundian students). Per average, filling in the questionnaire during an interview with a subject took about 2-3 hours.

The second phase of the trial involved 550 AMISOM soldiers (African Union Mission in Somalia; a peace-keeping mission approved by the UN) at the time before their military mission started. Note that the original questionnaire had to be customized for these subjects. Over 3 months, 16 interviewers (military psychologists from Burundi also joined the team) were collecting data based on the developed electronic questionnaire and the tablets it runs on.

The third and fourth phase of the trial involved 690 AMISOM soldiers after their military mission (160 soldiers were interviewed between July and August 2013; 530 between April and September 2014). Again the questionnaire had to be customized for each of these two trial phases. For example, questions were added to detect PTSD after the mission. Furthermore, valid questionnaires related to

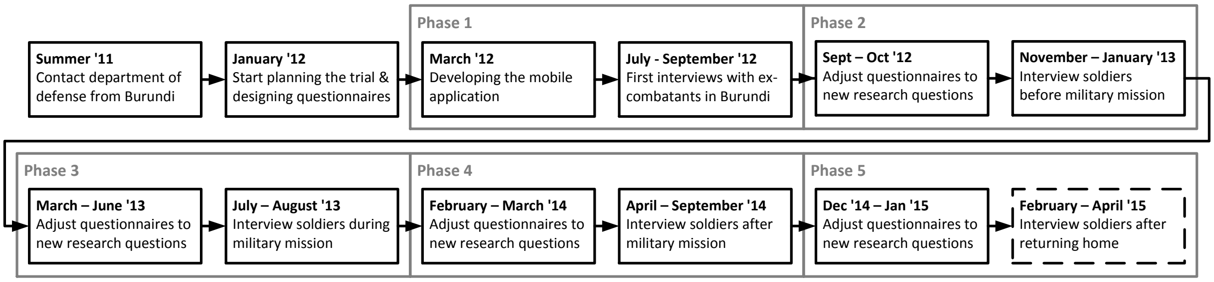

Figure 1. Course of the Trial

Dimension	Till Now	Expectation
Duration (in years)	3.5	4
App Variants	4	5
Subjects	1.700	2.200
Estimated Pages	85.000	110.000
Paper Weight (in kg)	425	550
Paper Height (in m)	8.5	11
Interview Time (in h)	5.100	6.600
	Expression	
Languages	3 (English, French, Kirundi)	
Team	17 Clinical Psychologists	
Team	6 Computer Scientists	

Table I
FIGURES OF THE PRESENTED CASE STUDY

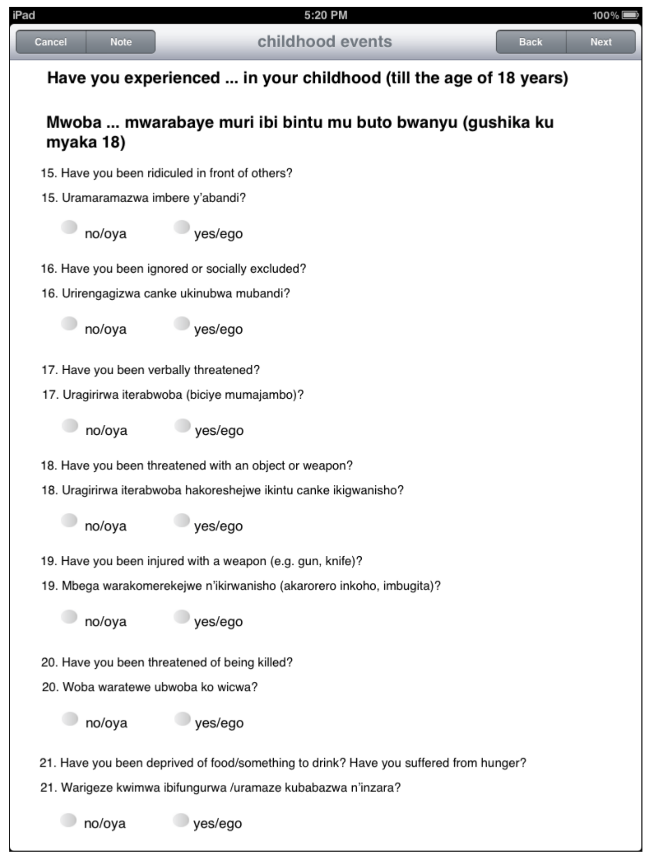

Figure 2. Screenshot of the Originally Developed Application

traumatic events during childhood were removed as they were already asked in a previous phase of the trial.

Currently, the psychologists are running the last (i.e., the fifth) phase of the trial, which will document PTSD from soldiers after being home for several months. The psychologists expect about 500 subjects to complete the last phase (see [3] for details).

By end of January 2015, about 1.700 subjects were interviewed based on different variants of the original electronic questionnaire, which runs on iPads. A questionnaire instance related to a particular subject consists of 450-500 questions per phase. If the psychologists had used paper-based questionnaires instead, one questionnaire instance would have required 40-50 pages. Consequently, this trial would have generated a paper stack of about 8.5m in height and 425kg in weight. Table I provides a quantitative summary of the trial.

When relying on paper-based questionnaires, logistic issues would have to be considered as well, e.g., regarding the challenge to transport large numbers of paper-based questionnaires to rural areas in the Western part of Burundi, where the interviews took place. Furthermore, the psychologists had to deal with privacy and security issues, as sensitive data from soldiers in military missions was processed.

B. Electronic Questionnaire Application

This section describes the application developed to enable mobile data collection. The mobile application runs on an iPad with iOS 5.x and covers the requirements of the already discussed trial.

Figure 2 shows a screen example of the corresponding user interface. More precisely, it shows one page of the electronic questionnaire. Note that the trial was carried out in interview mode, i.e., the interviewer was talking to the subject, while filling in the questionnaire concurrently.

During the interview, the psychologist may add notes to the answers given by the subject using the *Note* button. The latter is located in the top navigation bar of the application. Furthermore, interviewers may navigate through the questionnaire, i.e., they are flexible in jumping to different parts of the questionnaire depending on the respective interview and its course.

To meet privacy and security requirements imposed by the department of defense from Burundi, several restrictions had to be made. All collected data had to be encrypted on the smart mobile device utilizing state-of-the-art encryption algorithms. Moreover, no wireless transfer of the collected data was allowed. To meet these requirements, the collected data may only be accessed through the application itself. In this case, the data is automatically encrypted and temporarily stored in the application's data folder, where it may be accessed via iTunes when re-connecting the device to the computer. When analyzing the data collected, the private key is needed to decrypt the file, e.g., to process it with statistical software afterwards.

During the course of the trial, Apple released several new versions of their mobile operating system iOS. In turn, this required significant adaptations of the source code of the described mobile application. As response to major iOS releases, even a partial re-implementation of the mobile application became necessary. Due to Apple's update policy, which does not allow older devices to be upgraded to the latest version of their operating system, as well as the fact that there was no budget to buy new tablets, at a certain stage of the project we stopped adapting the source code to newer iOS versions. In September 2013, we cloned and adaptated the original code to support another trial from clinical psychology in Uganda, where psychologists are investigating genetic factors in combination with PTSD [12].

IV. Lessons Learned

During the trial, several lessons could be learned and fundamental requirements for mobile data collection in the large scale be elicited. The most important requirements are captured in Table II. Furhtermore, selected ones will be discussed in detail in this section.

A. Selected Requirements

This section highlights requirements, which were of particular relevance when implementing the trial application.

In Burundi, where the trial was conducted, there exist two official languages – French and Kirundi. However, as most psychologists were German, and hence were unable to speak any of the two languages fluently, a third language (English) had to be supported by the mobile application. Furthermore, the subjects to be interviewed spoke various dialects depending on their home region. To cope with these language issues, human translators were required to translate the respective questions from French or English to Kirundi. In this context, problems regarding the understanding of the on-the-fly translated Kirundi version of a question could be observed, since the translators interpreted the questions slightly different. Hence, a variety of translations with a somewhat modified meaning originated for the same question. To effectively manage the different versions of the translations, the psychologists added a review process for translations. However, when finalizing the translation, all smart mobile devices had to be adapted to the new version of the question. As each interviewer had his own smart mobile device for collecting data, numerous devices had to be changed. We implemented a mechanism to distribute the translations to all devices to update them at the same time to the new version in order to prevent ambiguosities and heterogeneities during the interviews.

Another challenge to be tackled during this long-term trial was to cope with the very short release cycles of the mobile operating platform itself. When developing the application in March 2012, Apple just released iOS 5.1. When planning the next trial phase in September 2012, Apple had introduced iOS 6.0 a few days before. Most of the code could be easily migrated to the new target platform. However, several software components also had to be adapted. In particular, after the release of iOS 7.0, the user interface had to be completely rewritten since Apple had introduced a new application styleguide. Furthermore, iOS 8.0 introduced *Swift*, a completely new programming language providing more functionality as well as an easier to read syntax compared to ObjectiveC.

B. Discussion

This section discusses lessons learned when implementing the electronic questionnaire application.

Due to the limited amount of time available for developing the first variant of the mobile electronic questionnaire, the application was hard-coded. We used Apple's Storyboard technology to create the screens as well as the transitions between them. Although this was a convenient and fast approach in the first run, its limitations became quickly evident later. Particularly, each adaptation of the questionnaire and its structure (e.g., to match the different phases of the trial) had to be programmed by IT experts. In turn, smaller changes (e.g., regarding the wording of a question) could be managed by the psychologists on their own. Still, there were many issues, the psychologists were unable to perform due to lack of programming skills. This includes, for example, changes in the procedures for validating the entered data, or changes in the implementation of the questionnaire logic. What has been missing but is urgently needed in such a setting, is sophisticated support for *end-user programming*, allowing domain experts to easily create and configure their mobile electronic questionnaires on an abstract (i.e., domain-specific) level. That is, requiring no programming or other technical skills. We have presented first work towards this vision in [14].

State	Lessons Learned	Description
General Information	Multi-Language Support	The electronic questionnaire must be displayed in different languages. The subject filling in the questionnaire shall be allowed to choose among different languages.
	Multi-User Support	When working with the device, different users may interact with the electronic questionnaire. In turn, these users may have different roles (e.g., interviewer or subject).
	Different Questionnaire Modes	Electronic questionnaires may be used in different modes (e.g., interviewer or self-rating). The user interface and the provided functions may diverge between these modes.
	Ensuring Validity	The validity known from the original paper-based questionnaires must be kept; i.e., the electronic questionnaire need to be designed in the same style (e.g., keep the same structure and layout).
During Enactment	On-the-Fly Manipulation	Questions of already deployed electronic questionnaires might have to be adapted to fix spelling errors or to improve understandability. This should be feasible without need for programming skills.
	Adding Freetext Notes	During an interview, it often becomes necessary to add notes. The electronic questionnaire should allow documenting data in addition to the one captured by the regular answers.
	Integrating Sensors	To utilize the capabilities of smart mobile devices when collecting data, it should be possible to integrate sensors (e.g., heart rate). This would allow for more detailed analyses (cf. [13]).
	Self-Explaining User Interfaces	Standard control elements (e.g., buttons or sliders) should be used for accomplishing data input to ensure familiarity with the application. In addition, the user interface should be kept intuitive and easy to understand.
After Enactment	Anonymizing and Encrypting Collected Data	When collecting data, there often emerge legal issues, e.g., requiring the anonymization and encryption of the collected data. The application should provide automated support for respective functions.
	Exporting Collected Data	To allow for a fast evaluation, an export of the collected data to standard file formats (e.g., CSV) needs to be provided.
Adapting	Maintaining the Questionnaire	As the questionnaire evolves over time or new versions of the operating system might be released, the mobile application should be easy to maintain. Moreover, no programming skills should be required, i.e., domain experts should introduce changes on their own.

Table II
REQUIREMENTS AND LESSONS LEARNED FROM THE LARGE-SCALE SCENARIO IN BURUNDI

We realized electronic questionnaires on smart mobile devices for enabling data collection in other trials as well. For example, we created a mobile application for tracking tinnitus perception, which has been used by patients from all over the world [15]. As opposed to the Burundi questionnaires, the *Track Your Tinnitus*[1] application could be described as *crowd sensing* application. Depending on customizable settings, the mobile application notifies the patient several times a day, triggering a specific questionnaire for collecting data. Furthermore, additional meta information like the sound volume level of the environment (measured in dB) is collected and analyzed. In about 1 year, more than 600 registered users have filled in 8.000 of these daily questionnaires. However, as this application is specifically tailored for the described tinnitus use case, a more generic approach is needed to support crowd-sensing projects. [16] provides a technical view on a reference architecture, and presents also requirements for creating crowd-sensing platforms.

The amount of data collected in the Burundi trial is considerable. Overall, the 1.700 interviewed subjects (cf. Table I) answered 765.000 questions until today. By end of the trial, we expect 1.100.000 answers (i.e., data entries). When using paper-based questionnaires, this trial would not have been possible in such a short time period, considering the rather low number of interviewers. Assuming an estimated error rate of 3% when transcribing data collected with paper-based questionnaires [17] (e.g., interviewer forgot to check an answer or could not interpret the written answer when digitizing the data), this would result in about 30.000 incorrect data entries. The mobile application we implemented, notifies the interviewer, if he forgets to answer a question. Thus, the number of errors that occur during the process of data collection could be significantly reduced.

[1] https://www.trackyourtinnitus.org/

V. Summary and Outlook

This paper has demonstrated the feasibility of using smart mobile devices for collecting data in clinical psychology trials. The presented trial is characterized by its large number of subjects, its specific location, and its challenging environmental and infrastructural conditions. Due to the long runtime of the trial as well as the need to change the questionnaires from trial phase to trial phase, continuous adaptations of the implemented questionnaire were required. In addition, not only the questionnaires themselves, but also the mobile operation system evolved over time introducing further complexity to the project. Finally, we could elicit fundamental requirements for developing mobile data collection applications, especially when targeting difficult environment.

The project has revealed the need of providing a generic questionnaire framework to domain experts. Such a framework should allow them to create electronic questionnaires executable on a smart mobile device on their own at a high level of abstraction requiring no programming skills. Currently, we are working on such an *end-user programming* approach (see [14], [18] for first results). This approach allows mapping a questionnaire to a visual process model, which can be interpreted and run on any smart mobile device based on a mobile process engine [19]. We have designed an architecture for this approach, which allows for the easy deployment of the modeled questionnaires to smart mobile devices [20]. When developing such an end-user development framework, other issues need to be addressed, such as the *synchronization of questionnaires* between multiple smart mobile devices or a *domain-specific modeling language* for visually defining and testing the questionnaires. Finally, we want to integrate a sensor framework to the developed electronic questionnaire framework to be used to enrich the data collected with further information [13].

Acknowledgement

This work was funded by the *Deutsche Forschungsgemeinschaft* (DFG) in the context of the QuestionSys project[2].

References

[1] M. Ruf-Leuschner, R. Pryss, M. Liebrecht, J. Schobel, A. Spyridou, M. Reichert, and M. Schauer, "Preventing further trauma: KINDEX mum screen – assessing and reacting towards psychosocial risk factors in pregnant women with the help of smartphone technologies," in *XIII Congress Europ Society of Traumatic Stress Studies*, 2013.

[2] D. Isele, M. Ruf-Leuschner, R. Pryss, M. Schauer, M. Reichert, J. Schobel, A. Schindler, and T. Elbert, "Detecting adverse childhood experiences with a little help from tablet computers," in *XIII Congress Europ Society of Traumatic Stress Studies*, 2013.

[3] A. Crombach, C. Nandi, M. Bambonye, M. Liebrecht, R. Pryss, M. Reichert, T. Elbert, and R. Weierstall, "Screening for mental disorders in post-conflict regions using computer apps – a feasibility study from Burundi," in *XIII Congress Europ Society of Traumatic Stress Studies*, 2013.

[4] A. S. Mudano *et al.*, "Using tablet computers compared to interactive voice response to improve subject recruitment in osteoporosis pragmatic clinical trials: feasibility, satisfaction, and sample size," *Patient Preference and Adherence*, vol. 7, p. 517, 2013.

[5] J. Ainsworth *et al.*, "A comparison of two delivery modalities of a mobile phone-based assessment for serious mental illness: native smartphone application vs text-messaging only implementations," *Medical Internet Research*, vol. 15, no. 4, 2013.

[6] M. R. Friesen, C. Hamel, and R. D. McLeod, "A mHealth application for chronic wound care: Findings of a user trial," *International Journal of Environmental Research and Public Health*, vol. 10, no. 11, pp. 6199–6214, 2013.

[7] M. Kirwan, M. J. Duncan, C. Vandelanotte, and W. K. Mummery, "Using smartphone technology to monitor physical activity in the 10,000 steps program: a matched case-control trial," *Medical Internet Research*, vol. 14, no. 2, 2012.

[2]http://www.uni-ulm.de/en/in/dbis/research/projects/questionsys.html

[8] H. Q. Nguyen *et al.*, "Pilot study of a cell phone-based exercise persistence intervention post-rehabilitation for copd," *Int'l Journal Chronic Obstructive Pulmonary Disease*, vol. 4, p. 301, 2009.

[9] R. Pryss, N. Mundbrod, D. Langer, and M. Reichert, "Supporting medical ward rounds through mobile task and process management," *Information Systems and e-Business Management*, vol. 13, no. 1, pp. 107–146, February 2015.

[10] E. B. Foa, D. S. Riggs, C. V. Dancu, and B. O. Rothbaum, "Reliability and validity of a brief instrument for assessing post-traumatic stress disorder," *Traumatic Stress*, vol. 6, no. 4, pp. 459–473, 1993.

[11] R. Weierstall and T. Elbert, "The appetitive aggression scale development of an instrument for the assessment of human's attraction to violence." *Europ Psychotraumatol*, vol. 2, 2011.

[12] S. Wilker, A. Pfeiffer, S. Kolassa, T. Elbert, B. Lingenfelder, E. Ovuga, A. Papassotiropoulos, D. de Quervain, and I.-T. Kolassa, "The role of fkbp5 genotype in moderating long-term effectiveness of exposure-based psychotherapy for post-traumatic stress disorder," *Translational Psychiatry*, vol. 4, no. 6, 2014.

[13] J. Schobel, M. Schickler, R. Pryss, H. Nienhaus, and M. Reichert, "Using Vital Sensors in Mobile Healthcare Business Applications: Challenges, Examples, Lessons Learned," *9th Int'l Conf on Web Inf Sys and Technologies (WEBIST)*, 2013.

[14] J. Schobel, M. Schickler, R. Pryss, F. Maier, and M. Reichert, "Towards process-driven mobile data collection applications: Requirements, challenges, lessons learned," in *10th Int'l Conf on Web Inf Sys and Technologies (WEBIST 2014), Special Session on Business Apps*, April 2014.

[15] R. Pryss, M. Reichert, J. Herrmann, B. Langguth, and W. Schlee, "Mobile Crowd Sensing in Clinical and Psychological Trials – A Case Study," in *28th IEEE Int'l. Symp. on Computer-Based Medical Systems*, 2015.

[16] M. Schickler, J. Schobel, R. Pryss, and M. Reichert, "Mobile Crowd Sensing - A New Way of Collecting Data from Large Patient Communities?" in *14th Congress Europ Society of Traumatic Stress Studies*, 2015.

[17] R. Black, P. Woolman, and J. Kinsella, "Variation in the transcription of laboratory data in an intensive care unit," *Anaesthesia*, vol. 59, no. 8, pp. 767–769, 2004.

[18] J. Schobel, M. Ruf-Leuschner, R. Pryss, M. Reichert, M. Schickler, M. Schauer, R. Weierstall, D. Isele, C. Nandi, and T. Elbert, "A generic questionnaire framework supporting psychological studies with smartphone technologies," in *XIII Congress Europ Society of Traumatic Stress Studies*, 2013, pp. 69–69.

[19] P. Dadam and M. Reichert, "The ADEPT Project: A Decade of Research and Development for Robust and Flexible Process Support - Challenges and Achievements," *Computer Science - Research and Development*, vol. 23, no. 2, pp. 81–97, 2009.

[20] J. Schobel, M. Schickler, R. Pryss, M. Reichert, and T. Elbert, "A Domain-Specific Framework for Collecting Data in Trials with Smart Mobile Devices," in *14th Congress Europ Society of Traumatic Stress Studies*, 2015.