

In the wild pilot usability assessment of a connected health system for stroke self management

Panagiotis Balatsoukas
King's College London
London, UK
panagiotis.balatsoukas@kcl.ac.uk

Nadin Kokciyan
University of Edinburgh
Edinburgh, UK
nadin.kokciyan@ed.ac.uk

Mark Ashworth
King's College London
London, UK
mark.ashworth@kcl.ac.uk

Isabel Sassoon
Brunel University London
London, UK
Isabel.Sassoon@kcl.ac.uk

Archie Drake
King's College London
London, UK
archie.drake@kcl.ac.uk

Vasa Curcin
King's College London
London, UK
vasa.curcin@kcl.ac.uk

Simon Parsons
University of Lincoln
Lincoln, UK
SParsons@kcl.ac.uk

Martin Chapman
King's College London
London, UK
martin.chapman@kcl.ac.uk

Sanjay Modgil
King's College London
London, UK
sanjay.modgil@kcl.ac.uk

Elizabeth Sklar
University of Lincoln
Lincoln, UK
ESklar@lincoln.ac.uk

Abstract— This paper reports on the findings of a pilot study for the formative “in the wild” assessment of the usability of CONSULT, a research-led connected health system for stroke self-management and prevention. CONSULT integrates data from commercial wellness sensors, electronic health records and clinical guidelines and enables users to monitor their vital signs to support self-monitoring and provision of tailored advice. The CONSULT system includes a dashboard and a chatbot. To assess the usability of our system, six volunteers were recruited to interact with CONSULT over a period of seven days. System logs confirmed that participants interacted with the CONSULT system throughout. CONSULT's ability to integrate data from different sensors was an aspect of this systems that all participants liked and kept them motivated to track their vital signs. The study also revealed several usability issues that designers of this type of systems should consider. Some of the most prevalent issues were: information overload, data misinterpretation, need for more anthropomorphic conversational capabilities for the chatbot; lack of visibility of the data transmission status. This paper concludes with reflections on the importance of these findings when assessing the usability of connected health systems, like CONSULT.

Keywords—human factors, usability, connected health

I. INTRODUCTION

The increasing incidence of primary and secondary stroke [1, 2] has led to the design of several preventive interventions that harness the power of modern technology [3]. More recently, the wide availability of wearable sensor devices has made possible the development of a new generation of data-driven decision support tools. Thanks to the ability of such tools to process large amounts of heterogeneous health data it becomes possible for patients to self-monitor and get personalised treatment plans, advice and assistance for better monitoring, self-management and stroke prevention [4,5, 6].

The CONSULT system [7] (Fig.1) was designed to serve this purpose through: (a) integration of data from sensor devices (wireless Blood Pressure cuff, smartwatch and wireless ECG patches), electronic health records and clinical guidelines; (b) information retrieval from publicly available and trusted sources, such as the UK's NHS website [8], and presentation of tailored health advice. All data and information are communicated to patients through our dedicated app.

Grounded by a participatory design approach [9] and the self-determination theory (SDT) of behaviour change [10], we designed two different versions of the CONSULT app with which a user can interact [11]. The first version makes use of a standard dashboard design to present information and support monitoring of vital signs. The CONSULT dashboard integrates measurements from three different sensors for blood pressure, heart rate and ECGs. In the second version, in addition to the standard dashboard design described above, the user can engage into a conversation with a chatbot in order to receive tailored information and advice. If information presented by the chatbot is not well understood by the user, it is possible for the user to ask the robot/agent for further explanations or clarifications.

II. AIM AND OBJECTIVES

The purpose of this pilot study was twofold: First, to test the feasibility and appropriateness of a research protocol developed to meet the needs of a forthcoming large-scale trial which aims to assess the usability and acceptability of our system. Second, to identify major usability and technical problems of our system, if any, that should be fixed before proceeding to the large-scale evaluation.

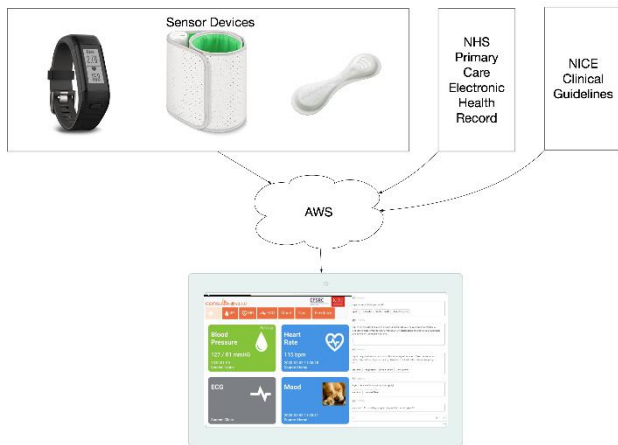


Fig. 1. Main components of the CONSULT system.

III. METHODS

A mixed methods within-subjects design in a naturalistic, in the wild [12], context was employed with six participants (all healthy volunteers). Participants interacted with two versions of the CONSULT app (with and without the chatbot) over a 7-day period, which is shorter than the period planned for the forthcoming large-scale trial. The study was split into three phases:

A. Phase 1

Day 1 of the 7-day period included induction and informed consent. Each participant was provided with an android tablet device, where the CONSULT application had been installed, as well as a set of 3 wireless wearable sensors – Vivosmart 4 by Garmin, a smartwatch used to track wellness and fitness activity (although heart rate only was of interest in this pilot); The wireless blood pressure (BP) monitor by Withings/NOKIA, similar in size and shape to a traditional BP cuff; The VitalPatch by MediBioSense, that recorded ECGs, a 11 grams patch with size 115x36x8mm and normally placed on the patient’s upper left chest. In the induction participants were shown how to use the app and the sensors. Instructions were provided that set a recommended minimum number of times per day that data should be recorded using the sensors: the Vivosmart 4 was recommended to be used by the participants as they would normally wear their own watch; The wireless BP cuff at least once daily; The VitalPatch was recommended to be worn for as long as it felt comfortable to avoid skin irritation. Data from the sensors was stored in a secure designated database, and associated virtual machine, provided by Amazon Web Services (AWS). Inductions took place at King’s College London and lasted approximately 2 hours with each participant attending individually. Participants were advised to familiarise themselves with the technology for the rest of day 1. Further support was available through a dedicated email address and phonenumber.

B. Phase 2 & 3

Both phases lasted three days each without a break between them. During phase 2 participants were assigned for three days one of the two versions of the CONSULT, then asked to shift to the other version for another three days (phase 3). Block randomisation was used to ensure an equal number of patients had used both versions. Given the nature of this specific study no washout period was deemed necessary. At the end of phase 2 (i.e. day 4) and phase 3 (day 7) a researcher interviewed each participant individually about their

experiences and perceptions on usability. Interviews were semi-structured and lasted approximately 30 to 60 minutes. Interviews took place at private venues agreed with each participant. Throughout the study log data of system usage (e.g. frequency and duration of usage, as well as features used) was recorded to infer the application’s level of usage and usability. Finally, at the end of phase 3 participants returned the tablet and the three sensors and were reimbursed (£50).

C. Data Analysis

Usage logs were analysed descriptively. Qualitative data was analysed inductively using a thematic analysis [13]. Heart rate and blood pressure data was displayed to the participants, through the CONSULT dashboard, both as instantaneous values (i.e. all individual measurements) and averages at hourly and daily intervals. ECG figures were displayed at daily intervals (e.g. last 24 hours). However, these physiological measurements were not subjected to any further analysis given the fact that the research protocol aimed at exploring usability and not effects on health outcomes.

D. Sample Characteristics and Recruitment

Six healthy volunteers were recruited from the student and staff population of King’s College London. People who had defibrillators or pacemakers were excluded from the research, no further exclusion criteria were applied. One participant was male and five were female. Their age ranged between 18 and 54 years old. They were all users of an android or iPhone device. Two of the participants had no previous experience of mobile health apps, while four were occasional users of fitness apps. This study was approved by the Research Ethics Committee of King’s College London.

IV. FINDINGS

The findings of the pilot study supported the feasibility of our connected health system. Participants were able to collect, upload and monitor heterogeneous physiological and self-reported data using the wearable sensor devices during the 7 days. All participants expressed positive feelings about the CONSULT app and the three wearable sensors, especially when it comes to the integration of heterogeneous health data in a single dashboard, the proactive behaviour of the chatbot (e.g. its capability to initiate a conversation in order to monitor mood or alert the patient for any abnormalities in BP), the comfort and ease of use of the sensor devices. However, the pilot also showed that participants’ experiences were hindered by various usability issues that were prevalent in the case of the app’s interface. The most common issues were: difficulty to interpret data (especially the ECG readings but also when trying to understand the meaning of their measurements across time and in combination with each other), the lack of a more natural conversation flow when interacting with the chatbot (e.g. close to the one that they would have with their GP), as well as the lack of visible notifications about whether or not data was transmitted successfully from the sensors to the app. The log analysis showed no differences in usage between the two conditions (with/without chatbot). The BP and HR plots were the most frequently accessed features, as well as the feedback tab that users could use to take notes.

Finally, this pilot confirmed that the research protocol used to assess the usability of our connected health system was appropriate. In-the-wild, naturalistic, approaches using a mix of quantitative and qualitative techniques (e.g.

phenomenography-driven interviews and data logs) can support the collection of rich longitudinal data to help assess human factors and experiences with this type of connected technologies.

V. CONCLUSIONS

This pilot study has limitations most of which are related to the sample's size and characteristics. For example, a larger sample size would have strengthened the generalisability of our findings. In addition, having stroke survivors in the sample would have helped us to better understand user experience in context. Finally, running a longer-term study (>7 days) would have helped us unpick some patterns in human behaviour when interacting with the two conditions. Following an iterative design research life-cycle, the next step for us would be to involve patients in a series of co-design activities in order to determine how to best address the issues that emerged and improve the usability of CONSULT before proceeding to a large-scale evaluation.

ACKNOWLEDGMENT

This research was funded by the UK Engineering Physical Sciences Research Council (EPSRC) under grant \#EP/P010105/1

REFERENCES

- [1] G. NHS. Available at: <https://www.nhs.uk/conditions/stroke/>
- [2] Burn, J., Dennis, M., Bamford, J., Sandercock, P., Wade, D. and Warlow, C., 1994. Long-term risk of recurrent stroke after a first-ever stroke. The Oxfordshire Community Stroke Project. *Stroke*, 25(2), pp.333-337
- [3] Parke, H.L., Epiphaniou, E., Pearce, G., Taylor, S.J., Sheikh, A., Griffiths, C.J., Greenhalgh, T. and Pinnock, H., 2015. Self-management support interventions for stroke survivors: a systematic meta-review. *PLoS One*, 10(7), p.e0131448
- [4] Kökciyan, N., Sassoön, I., Young, A.P., Chapman, M., Porat, T., Ashworth, M., Curcin, V., Modgil, S., Parsons, S. and Sklar, E., 2018. Towards an argumentation system for supporting patients in self-managing their chronic conditions. In *Proceedings of the AAAI Joint Workshop on Health Intelligence (W3PHIAI 2018)*
- [5] Contreras, I. and Vehi, J., 2018. Artificial Intelligence for Diabetes Management and Decision Support: Literature Review. *Journal of medical Internet research*, 20(5).
- [6] Thilarajah, S., Clark, R.A. and Williams, G., 2016. Wearable sensors and Mobile Health (mHealth) technologies to assess and promote physical activity in stroke: a narrative review. *Brain Impairment*, 17(1), pp.34-42.
- [7] Essers, K., Chapman, M., Kokciyan, K., Sassoön, I., Porat, T., Balatsoukas, P., Young, P., Ashworth, M., Curcin, V., Modgil, S., Parsons, S. and Sklar, E.I., 2018., The CONSULT System: Demonstration. In *Proceedings of the 6th International Conference on Human-Agent Interaction (HAI '18)*. Association for Computing Machinery, New York, NY, USA, 385–386. DOI:<https://doi.org/10.1145/3284432.3287170>
- [8] NHS. <https://www.nhs.uk/>.
- [9] Schuler, D. and Namioka, A. eds., 1993. *Participatory design: Principles and practices*. CRC Press.
- [10] Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (p. 416–436).
- [11] Balatsoukas, P., Porat, T., Sassoön, I., Essers, K., Kokciyan, N., Chapman, M., Drake, A., Modgil, S., Ashworth, M., Sklar, E. and Parsons, S., 2019, July. User involvement in the design of a data-driven self-management decision support tool for stroke survivors. In *IEEE EUROCON 2019-18th International Conference on Smart Technologies* (pp. 1-6). IEEE.
- [12] Waterson, S., Landay, J.A. and Matthews, T., 2002, April. In the lab and out in the wild: remote web usability testing for mobile devices. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems* (pp. 796-797).
- [13] Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), pp.77-101.