Feedback to Renal Transplant Patients in a Self-management Support System

Wenxin Wang[a] Willem-Paul Brinkman[a] Ton J.M. Rövekamp[b] Paul van der Boog[c] Laurence Alpay[b] Mark A. Neerincx[a,d]

[a]Delft University of Technology, Mekelweg 4, 2628 CD Delft, the Netherlands. [b]TNO Lifestyle, Leiden, the Netherlands. [c]Leiden University Medical Center, Leiden, the Netherlands. [d]TNO Human Factors, the Netherlands.
{w.wang-1, w.p.brinkman, m.a.neerincx}@tudelft.nl {ton.rovekamp, laurence.alpay}@tno.nl
p.j.m.van_der_boog@lumc.nl

ABSTRACT

Motivation – A key function of a future self-management support system (SMSS) for renal transplant patients is to provide feedback about their health status. This study investigates patients' understanding, preference, and trust of such feedback.

Research approach – Three types of feedback form, namely simplicity, empathy, and empowerment, were designed and tested with 12 non-patients. The task completeness and the participants' preference were compared.

Findings – The users did trust the empowerment feedback more than other feedback. Furthermore, the feedback types seemed to influence users' ability of reporting their previous days' health status.

Research limitations/Implications – This research worked out three feedback types and provided insight into their effectiveness and preference. However, the number of participants was small, and they were non-patients, highly educated and relatively young.

Originality/Value – This research investigated different feedback types for self-management support systems in the healthcare domain.

Take away message – The different way of presenting the same information might influence users' trust and understanding of their health status.

Keywords

Self-management support system, feedback, usability, trust, renal transplant patient

INTRODUCTION

Chronic kidney disease is regarded as a major public health problem (National Kidney Foundation, 2002). In the Netherlands, 14,794 people were under treatment for the end-stage renal disease (ESRD) at the end of 2009, and 827 renal transplantations, the ideal treatment of ESRD, were performed in 2009. However, rejection and medication side effects can occur after transplantation (Crespo et al., 2001; Veenstra, Best, Hornberger, Sullivan, & Hricik, 1999). Therefore, renal transplant patients are treated as chronically ill.

These patients, as other chronic patients, are required to adapt their behavior actively (Bodenheimer, Lorig, Holman, & Grumbach, 2002; Lorig & Holman, 2003). Therefore, self-management, the process of managing symptoms, treatment, physical and psychosocial consequences, has been proposed (Barlow, Wright, Sheasby, Turner, & Hainsworth, 2002).

To help implement self-management, self-management support systems (SMSS) have been proposed (Lorig, Ritter, Laurent, & Plant, 2006). A SMSS can help empower the patients having control of their care process and daily activities, and therefore increasing their autonomy (Lorig, et al., 2006). This paper reports on the work carried out in the context of the ADMIRE project (Assessment of a Disease management system with Medical devices in Renal disease). In ADMIRE, a SMSS for renal transplant patients is proposed to guide them conducting daily self-management, such as self-measuring, getting feedback, and reacting. By doing so, it is expected that the patients can know their health status better, be more alert, and visit hospital less.

One of the core functions of this SMSS is providing feedback of patients' renal function. After patients measure themselves, they can input the measurements (e.g., creatinine level, temperature, pulse, and weight) into the SMSS, which should then give them feedback and suggestions. However, because complications and side effects can occur, patients might have reservation towards trusting and accepting the feedback. Therefore, this study focuses on investigating people's understanding, preference, and trust of such feedback.

To let patients accept a SMSS, its usability and safety seem important (Davis, 1989; Venkatesh, Morris, Gordon, & Davis, 2003) because 1) the patients should be able to use it easily and 2) patients should be able to understand it and take action correctly. Therefore, before testing with renal patients, prototypes were tested with non-patients as the first step in this study. The clarity of the information, and people's preferences and opinions on different designs were investigated. Furthermore, the study also aimed at understanding how people interpret and interact with the designs, in order to develop a transparent, easy to use, and trustworthy system.

BACKGROUND

Among all measurements, creatinine level was regarded most critical to detect rejection. It is unlikely that there

is a rejection if the creatinine level is stable or decreases. On the other hand, when it increases obviously, there may be rejection (Chapman, 2010). Therefore in ADMIRE project, the SMSS suggests patients to measure once more if their creatinine increases obviously, and contact the hospital if it increases too much. In this study, therefore, whether users could understand the creatinine status and whether they would follow the instructions was tested.

Considering the user needs obtained in previous research and existing design principles, several design suggestions can be made for presenting feedback in a SMSS. In this study three different main design ideas were compared: simplicity, empathy and empowerment. In the simplicity design, the "Minimal Manual (MM)" guidelines (Carroll, 1987) were applied. MM focuses on the real task and adequate text. Thus in this design a patient's task was simply illustrated by traffic lights. Green stands for "all right", orange for "measure once more", and red for "call the hospital". There are two variations: showing today's status with traffic lights, and showing today's status with traffic lights combined with colour-coded previous measurements (Figure 1.a).

Including emotional communication in computer system has been suggested to help in decision making and learning by increasing users' empathy experience (Barkhuysen, Krahmer, & Swerts, 2005; Looije, Cnossen, & Neerincx, 2006; Picard, 1995). The second design applied affective computing ideas by showing empathy from an avatar. Its various facial expressions and gestures convey a patient's renal function and express empathy. Such empathy might let patients understand easily by simulating human communication.

Empowerment design encourages people to gain the skills and knowledge that help them overcome obstacles (Carlos, 2010). By presenting information in a proper way, technology can increase people's understanding and performance (Fogg, 2003). The third type, therefore, attempted to empowerment patients by presenting information that help them understand their status. This information concerns mainly the dynamic thresholds that tell how the healthy or not results come out. The presentation of the thresholds has four variations: 1) colour-coded lines, 2) colour-coded bars, 3) colour-coded background of definite thresholds, and 4) colour-coded background of blurred thresholds.

All these designs had creatinine level graphs and explanatory texts. In addition, in each design, there were three situations that patient 1) is all right, 2) needs to pay attention, and 3) needs to contact the hospital. Examples of the three feedback types are in Figure 1.

The research question of this study was whether the feedback types (simplicity, empathy and empowerment) have an influence on users' task completeness, liking, perceived ease of use (PEOU), and trust of the system.

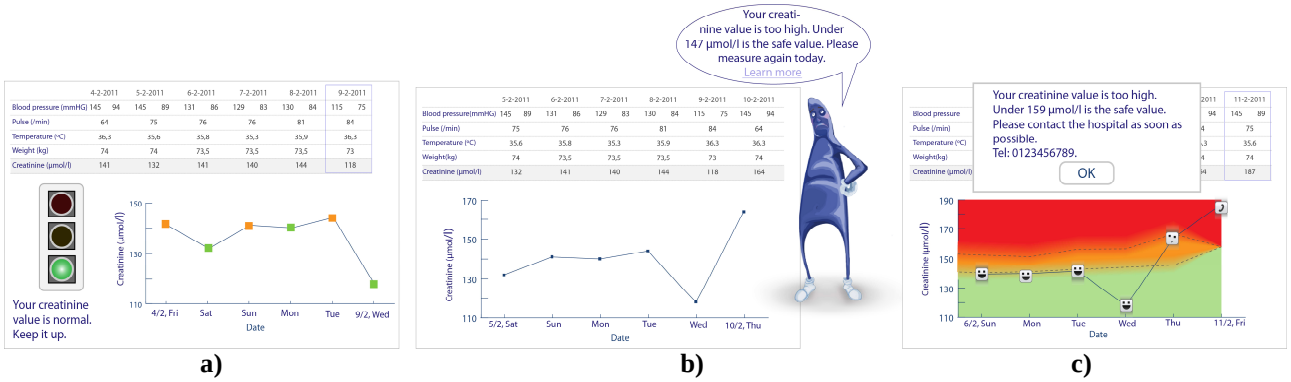
METHOD

The experiment had a within-subjects design. All three kinds of feedback were shown to each participant as paper-prototypes. To avoid learning effect, the order of feedback and the tasks were counterbalanced. This resulted in six sequences. The designs of traffic lights and dynamic threshold had more than one version. They, as well as the three health statuses, were shown to each participant randomly. The experiment was approved by the university ethics committee.

Participants were recruited from Delft University of Technology community. The 12 participants, 9 male and 3 female, were 22 to 38 years old (M = 29, SD = 4) Dutch-speaking non-patients. They all had a bachelor's or higher educational level in science or engineering.

At the beginning participants received an introduction about the study aim, and had to sign a consent form. Afterward, they completed a basic information questionnaire. Then they were asked to imagine that they were renal transplant patients, and were explained the relation between creatinine level and renal function.

They were asked to think aloud during the main part, which consisted of three sessions. In each session one of the three feedback types was presented. Participants were asked to complete some predefined tasks and questions, such as to find out current and/or previous health status. They were asked to, besides think aloud, tell how they interpret the information. After each session, they filled out a preference questionnaire. At the end of the experiment, participants were debriefed about the experiment. The entire experiment took around one hour.



a) **b)** **c)**

Figure 1. Screenshots of the prototypes. **a)** Simplicity design showing today's status and previous status. **b)** Empathy design. **c)** Empowerment design with colour-coded background and blurred thresholds.

Before the interaction with the prototypes, participants completed a questionnaire of their basic information such as gender, age, and educational level. After each session, all the prototypes of the same feedback type (i.e. simplicity, empathy, or empowerment) were shown to the participants, and they completed a specifically designed preference questionnaire. Here they were asked to rate on 7-point Likert scales how much they liked it, trusted it, and how easy it was to use,. The participants' interpretation of the prototype and understanding of the information were recorded using think aloud protocol. Besides subjective data, behavioral data of the task completion was collected.

RESULTS

Task Completeness

The performance of each task was classified as wrong, not precise, or right. Although there were more than one variations in two feedback types, the task completeness were compared between feedback types, instead of prototypes, because the participants interacted randomly with one variation. Friedman tests showed a border line significant difference of task completeness of finding out current health status ($\chi^2(2) = 6.00$, $p. = 0.05$), and a significant difference of finding out previous health status ($\chi^2(2) = 9.10$, $p. = 0.011$). It can be seen in Table 1 that with the empathy feedback less tasks were completed than with other two feedback types.

Table 1. Task Performance Results

Task	Feedback type	Wrong (No.)	Not precise (No.)	Right (No.)
Finding current/ previous status	Simplicity	0/1	0/1	10/8
	Empathy	0/5	4/3	6/2
	Empower-ment	0/0	0/4	10/6

Preference

Within Simplicity and within Empowerment Feedback

Between two variations of the simplicity prototypes, Wilcoxon signed ranks pair wise comparison tests found significant higher liking ($Z = -2.51$, $p. = 0.012$), PEOU ($Z = -2.72$, $p. = 0.007$), and trust ($Z = -2.46$, $p. = 0.014$) of the one with colour-coded dots ($Mdn_{liking} = 5$, $Mdn_{PEOU} = 5$, $Mdn_{trust} = 5$) than the one without such dots ($Mdn_{liking} = 3.5$, $Mdn_{PEOU} = 4$, $Mdn_{trust} = 4$).

For the four empowerment designs, Friedman tests revealed significant differences of liking, PEOU, and trust ($\chi^2(3) = 32.48$; 32.89; 24.72, respectively, and all $p. < 0.001$). Wilcoxon signed ranks tests were conducted, and significant differences were found in all liking, PEOU, and trust for all the pairs, except between colour-coded background with blurred and with definite thresholds, and the trust between colour-coded lines and colour-coded bars. The medians are shown in Figure 2. It can be seen that the colour-coded background with blurred and with definite thresholds designs were best.

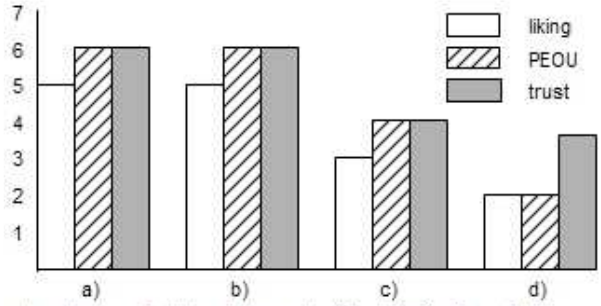


Figure 2. Median preference of empowerment designs.

Between Feedback types

Since there were significant preference differences between the variations of both simplicity feedback and empowerment feedback, the one with the highest median of each type was selected for the comparison between feedback types. Thus, the designs compared here were simplicity feedback of colour-coded dots, empathy feedback, and empowerment feedback of background with blurred thresholds.

Friedman tests showed no significant differences in the liking rating ($\chi^2(2) = 4.54$, $p. = 0.10$) or PEOU rating ($\chi^2(2) = 1.24$, $p. = 0.54$), but showed a significant difference in the trust rating ($\chi^2(2) = 6.44$, $p. = 0.040$). In Wilcoxon signed ranks tests, trust was only found significantly higher ($Z = -2.25$, $p. = 0.024$) for empowerment feedback than on empathy (Figure 3).

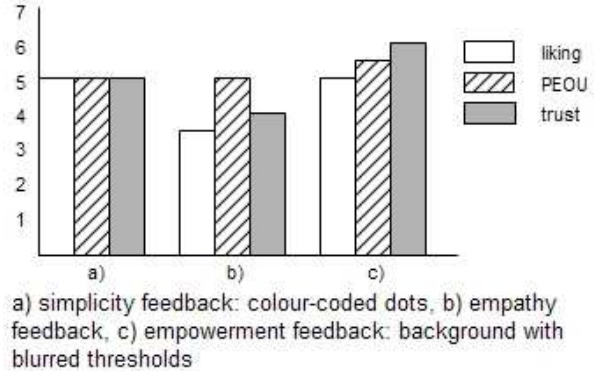


Figure 3. Median preference of feedback types.

Participants' Comments

In the aloud thinking and debriefing, participants commented on each design. Some noticeable ones were: 1)Some preferred to have the thresholds, so that they could know how far they were from it, 2) Since the thresholds were dynamic, the same measured value can be sometimes safe while sometimes not. If the thresholds were not shown, participants may be confused and not trust it, 3) The avatar's expressions and gestures were confusing, and sometimes were conflicting with the text. This might explain why the participants did not trusted it, and 4) Some participants did not understand the dynamic thresholds. They wondered why it was dynamic.

CONCLUSION

The significant difference of task completeness was only found from reporting previous status, which seems logical, because the previous status was explicitly shown in the simplicity and empowerment feedback, but not in the empathy one.

Significant differences of users' preference (liking, PEOU, and trust) were found within the same types of feedback (i.e., simplicity and empowerment). It implies that even having the same information, the different presenting way can influence users' preference. The significant differences between different feedback types is only found in trust.

DISCUSSION

Although several significant differences were found in the users' preference, especially within the same feedback type, the difference between feedback types was only found significant in trust. Besides, there was only one design for the empathy feedback; it could be that this design itself was not being trusted, instead of the empathy type (e.g., the figure of avatar was considered ugly or wired, or the facial expressions were not clear enough). There was no significant difference found between simplicity and empowerment designs, although a few participants commented that with thresholds they could know more, and some also mentioned that the simplicity design was quite direct.

Nevertheless, the number of the participants was relatively small. In addition, they were non-patients and were younger than most Dutch ESRD patients (M_{age} = 58.5, *SD* = 16.3) (ERA-EDTA Registry, 2011). Usability is however different for seniors, and their preference might therefore also be different (Coyne & Nielsen, 2002). In the future, an experiment with a high fidelity prototype and patients as participants will be conducted to investigate the responses of real users.

ACKNOWLEDGEMENTS

As part of ADMIRE project, this work is funded by the Netherlands Organisation for Health Research and Development (ZonMw).

REFERENCES

Barkhuysen, P., Krahmer, E., & Swerts, M. (2005). Problem detection in human–machine interactions based on facial expressions of users. *Speech Communication, 45*(3), 343-359.

Barlow, J., Wright, C., Sheasby, J., Turner, A., & Hainsworth, J. (2002). Self-management approaches for people with chronic conditions: a review. *Patient Education and Counseling, 48*(2), 177-187.

Bodenheimer, T., Lorig, K., Holman, H., & Grumbach, K. (2002). Patient Self-management of Chronic Disease in Primary Care. *JAMA: The Journal of the American Medical Association, 288*(19), 2469-2475.

Carlos, J. P. (2010). *Empowerment Takes More Than a Minute*: ReadHowYouWant.com, Limited.

Carroll, J. M. (1987). Minimalist design for active users. In R. M. Baecker (Ed.), *Human-computer interaction* (pp. 621-626): Morgan Kaufmann Publishers Inc.

Chapman, J. R. (2010). The KDIGO Clinical Practice Guidelines for the Care of Kidney Transplant Recipients. *Transplantation, 89*(6), 644-645.

Coyne, K., & Nielsen, J. (2002). Web Usability for Senior Citizens: 46 Design Guidelines Based on Usability Studies with People Age 65 and Older.

Crespo, M., Pascual, M., Tolkoff-Rubin, N., Mauiyyedi, S., Collins, A. B., Fitzpatrick, D., . . . Saidman, S. L. (2001). Acute humoral rejection in renal allograft recipients: I. Incidence, serology and clinical characteristics. [Article; Proceedings Paper]. *Transplantation, 71*(5), 652-658.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319-340.

ERA-EDTA Registry. (2011). ERA-EDTA Registry Annual Report 2009. Amsterdam, The Netherlands: Academic Medical Center, Department of Medical Informatics.

Fogg, B. J. (2003). *Persuasive technology: using computers to change what we think and do*: Morgan Kaufmann Publishers.

Looije, R., Cnossen, F., & Neerincx, M. A. (2006, 6-8 Sept. 2006). *Incorporating guidelines for health assistance into a socially intelligent robot.* Paper presented at the Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on.

Lorig, K. R., & Holman, H. R. (2003). Self-management education: History, definition, outcomes, and mechanisms. *Annals of Behavioral Medicine, 26*(1), 1-7.

Lorig, K. R., Ritter, P. L., Laurent, D. D., & Plant, K. (2006). Internet-Based Chronic Disease Self-Management: A Randomized Trial. *Medical Care, 44*(11), 964-971.

National Kidney Foundation. (2002). KDOQI Clinical Practice Guidelines for Chronic Kidney Disease: Evaluation, Classification, and Stratification. *American journal of kidney diseases : the official journal of the National Kidney Foundation, 39*(2), S1-266.

Picard, R. W. (1995). *Affective computing* (Rev. Nov. 26, 1995. ed.). Cambridge, Mass.: Perceptual Computing Section, Media Laboratory, Massachusetts Institute of Technology.

Veenstra, D. L., Best, J. H., Hornberger, J., Sullivan, S. D., & Hricik, D. E. (1999). Incidence and long-term cost of steroid-related side effects after renal transplantation. *American Journal of Kidney Diseases, 33*(5), 829-839.

Venkatesh, V., Morris, M. G., Gordon, B. D., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly, 27*(3), 425-478.