

The Reliability of Using Tablet Technology for Screening the Health of Older Adults

Lex VAN VELSEN^{ab,1}, Sanne FRAZER^a, Aurèle N'DJA^c, Nadir AMMOUR^c, Susanna DEL SIGNORE^d, Gianluca ZIA^e & Hermie HERMENS^{ab}

^a*Roessingh Research and Development, Telemedicine cluster, The Netherlands,*

^b*University of Twente, Biomedical Signals and Systems group, The Netherlands,*

^c*Sanofi R&D, France,* ^d*BLUECOMPANION LTD, United Kingdom,*

^e*Caretek s.r.l., Italy*

Abstract. In this study, we assessed the reliability of using a tablet application for collecting health data among older adults, in comparison to using paper surveys for this goal. Test-retest reliability between the two modalities, usability, user experience factors, and older adults' preference were determined. The results show perfect agreement between tablet and paper for the SARC-F and high agreement for the SF-36 physical scale and EQ-5D. Usability and user experience factors were perceived the same for both modalities. The majority of the participants preferred the tablet for health screening purposes, mainly because of its ease of use. This study shows that using tablets for health screenings among older adults does not affect test reliability, and that older adults prefer the tablet to paper for completing these tests.

Keywords. Population health screening, test reliability, tablet computers

1. Introduction

The rapid adoption of tablet technology among older adults has tremendously expanded our options for offering digital health interventions. Tablet-based applications can sort positive effects on their management of chronic conditions, maintenance of physical and cognitive health, and social wellbeing [1]. Another area in which the use of tablets may be valuable is public health screening. Deploying large-scale health screenings digitally, instead of on paper, is likely to yield cost-savings, will minimize missing and illegible data, and allows us to inform the older adult of the screening results immediately.

High reliability of health screening surveys, completed on a tablet, is of course a prerequisite for their use. Hess and colleagues [2] assessed the difficulty which a large sample of (relatively young, highly-educated) primary care patients had with completing a health screening survey on a tablet. They found that the majority of their participants had no difficulty with using a tablet here, but also state that special attention needs to be paid to potentially vulnerable groups. A sample of older adults was involved in the comparison between test results of three surveys provided via a tablet and on paper [3]. The study provided the preliminary evidence that the modality in which a screening is offered does not affect its results.

¹ Lex van Velsen: Senior researcher, Roessingh Research and Development, P.O. box 310, 7500 AH Enschede, The Netherlands; e-mail: l.vanvelsen@rrd.nl.

In this study, we assessed the reliability of using a tablet application for collecting physical health data among older adults, compared to using paper surveys. We determined test-retest reliability between the two modalities, older adults' preferences, and the user experience of working with each modality. The results of this study allow us to make well-informed policy decisions about the way administer health screenings among older adults.

2. Methods

2.1. Health screening surveys

In recent years, it has become widely recognized that population-wide screenings among older adults are an important strategy to identify frailty. New technologies have been found to be potentially very valuable for improving the quality of these screenings [4]. Therefore, we selected three screening surveys from the frailty domain:

1. SARC-F: A survey with 5 multiple choice questions (3 answering options) that identifies sarcopenia [5];
2. SF-36: A survey to determine an individual's health status from which we utilized the physical functioning scale, consisting of 10 multiple choice questions (3 answering options) [6];
3. EQ-5D: A survey to assess an individual's health status, consisting of 5 multiple choice questions (3 answering options) and one 100-point visual analogue scale on which the respondent must rate his/her health [7].

Tests were provided on paper and via a tablet-based clinimetrics service, ReQuest [8].

2.2. Participant recruitment

Participants were recruited in the Netherlands via convenience sampling and visited at home. Participants had to be 60 years or older, needed to speak Dutch fluently, and should have been willing to provide informed consent. We targeted for inclusion of robust, pre-frail and frail participants (as determined by a physical therapist) to have representative sample for older adults that start running the risk of becoming, or are frail.

2.3. Data collection and analysis

In each session, a researcher first assessed demographics and administered the Mini-Mental State Examination (MMSE). Then the participants completed the questionnaires on paper and on a tablet (a Samsung Galaxy Tab 4, with a 10.1" screen) in a randomized order. After completing the screenings on each modality, its usability was assessed via the System Usability Scale (SUS) [9], as well as its perceived usefulness (scale by [10]), and the participants' perceptions of enjoyment (scale by [11]), and control (scale by [12]). Each scale used a 5-point rating scale. Each session ended by asking the participant for his/her preference for one of the two modalities. Reliability was assessed by calculating the agreement between paper and tablet results via a Spearman correlation, intraclass correlation coefficient (ICC), and Cohen's kappa.

3. Results

3.1. Participants

Nine men (37.5%) and 15 women (62.5%) participated in the study ($n = 24$). The average age was 71.63 years (range: 62 to 87 years). Fifteen persons (62.5%) had to deal with one or more physical conditions or complaints (e.g., difficulty with walking, cardiac problems), nine persons (37.5%) did not. Nineteen participants (79.2%) used the Internet via a smartphone ($n = 16$), tablet PC ($n = 14$), or laptop PC ($n = 13$). All participants scored in the highest category of the MMSE (range: 25 to 30 points), meaning that their cognitive abilities should not affect their survey replies. During the test, ten participants (41.7%) requested the use of a tablet stylus or were given one by the test administrator when a participant had difficulty with pressing the screen buttons.

3.2. SARC-F

Table 1 shows high agreement between the answers given for the single items on the paper and tablet version of the SARC-F. ICC and Kappa scores for the single items are moderate to good. All participants were classified exactly the same by the instrument, regardless of its modality. As a result, all reliability measures were excellent.

Table 1. Correspondence between SARC-F scores on paper and tablet.

Item	Agreement	Spearman Correlation	ICC	Cohen's kappa
SARC-F 1: Strength	22/24 (91.67%)	.80**	.80	.86
SARC-F 2: Walking	21/24 (87.50%)	.87**	.77	.61
SARC-F 3: Chair rise	22/24 (91.67%)	.70**	.71	.70
SARC-F 4: Stairs	19/24 (79.17%)	.72**	.60	.64
SARC-F 5: Falls	23/24 (95.83%)	.90**	.90	.90
SARC-F total score	18/24 (75.00%)	.96**	.87	.70
SARC-F classification	24/24 (100.0%)	1.00**	1.00	1.00

* $p < .001$ ** $p < .01$

Table 2. Correspondence between SF-36 scores on paper and tablet.

Item	Agreement	Spearman Correlation	ICC	Cohen's kappa
SF-36 1: Vigorous activities	19/24 (79.17%)	.78*	.79	.66
SF-36 2: Moderate activities	20/24 (83.33%)	.82*	.84	.73
SF-36 3: Groceries	20/24 (83.33%)	.84*	.82	.72
SF-36 4: Climbing > 1 Stairs	21/24 (87.50%)	.91*	.90	.81
SF-36 5: Climbing 1 stairs	18/24 (75.00%)	.67*	.63	.54
SF-36 6: Bending	15/24 (62.50%)	.44**	.45	.42
SF-36 7: Walking > mile	21/24 (87.50%)	.93*	.92	.79
SF-36 8: Walking > blocks	22/24 (91.67%)	.88*	.93	.83
SF-36 9: Walking 1 block	21/24 (87.50%)	.85*	.72	.67
SF-36 10: Bathing or dressing	21/24 (87.50%)	-.06 (n.s.)	-.06	-.06
SF-36 total score	7/24 (29.17%)	.93***	.90	.23
SF-36 classification	21/24 (87.50%)	.71*	.72	.71

* $p < .01$ ** $p < .05$ *** $p < .001$

3.3. SF-36: Limitations of activities

Table 2 shows that several items of the SF-36 resulted in high agreement and good reliability scores. Two items (5 and 6) showed low agreement and low reliability scores. Item 10 displayed low reliability scores, despite a high agreement. This is because there was little variance in the data (participants only used two out of three answering options). Hence, reliability metrics could not be calculated correctly. Most participants (21/24) were classified the same by the instrument on both modalities. Two participants scored 'normal' on paper, while they scored as having 'functional decline' on the tablet; one person scored as having 'functional decline' on the tablet, while being classified as 'normal' on paper. Reliability measures of the classification were moderate to good.

3.4. EQ-5D

Table 3 shows that agreement and reliability scores for three items (1, 3 and 4) are moderate. Item 2 showed high agreement and low reliability scores, due to low variance in participants' answers. Item 6 displayed a relatively low agreement and reliability scores, but utilized a 100-point VAS scale. Therefore, we think that an agreement of 11 out of 24 is very high. As the EQ-5D does not classify people in nominal categories, but results in an index (partly derived from a VAS score), agreement indices are relatively low. However, the Spearman correlation shows good agreement.

Table 3. Correspondence between EQ-5D scores on paper and tablet.

Item	Agreement	Spearman correlation	ICC	Cohen's kappa
EQ-5D 1: Mobility	20/24 (83.33%)	.64*	.65	.64
EQ-5D 2: Self-care	22/24 (91.67%)	-.04 (n.s.)	-.04	-.02
EQ-5D 3: Usual activities	18/24 (75.00%)	.55*	.64	.52
EQ-5D 4: Pain/discomfort	19/24 (79.17%)	.66*	.64	.60
EQ-5D 5: Anxiety/depression	24/24 (100.0%)	1.00	1.00	1.00
EQ-5D 6: VAS score	11/24 (45.83%)	.37 (n.s.)	.08	.41
EQ-5D index	11/24 (45.83%)	.75**	.53	.36

* $p < .01$ ** $p < .001$

3.5. Modality preference and user experience

The tablet and paper screenings yielded equal scores on usability ($m=76.25(sd\ 22.35)$ vs. $m=81.78(sd\ 14.23)$; $t(23)=1.13$, $p=.27$), pleasure ($m=3.27(sd\ .40)$ vs. $m=3.16(sd\ .33)$; $t(23)=1.44$, $p=.16$), perceptions of control ($m=3.75(sd\ 1.18)$ vs. $m=3.65(sd\ .94)$; $t(23)=.34$, $p=.74$), and usefulness ($m=3.78(sd\ 1.01)$ vs. $m=3.43(sd\ 1.35)$; $t(23)=-.97$, $p=.34$). The usability of both modalities was appreciated well. Pleasure was regarded to be neutral, control and usefulness to be moderately high.

Seventeen persons (70.8%) preferred the tablet, seven persons (19.2%) rather used paper. Reasons for preferring the tablet (given more than once) were ease of use ($n = 13$), it is environment-friendly ($n = 3$), it provides a better overview ($n = 3$), it is easier to correct a mistake ($n = 3$), it is more efficient ($n = 3$), and it has a more relaxing experience ($n = 2$). Reasons for preferring paper (stated more than once) were its trustworthiness ($n = 4$), out of habit ($n = 3$), computer anxiety ($n = 2$), or ease of use ($n = 2$).

4. Discussion and conclusion

In this study, we assessed the reliability of using tablet computers for screening the health of older adults, by comparing the results of the same screening provided via a tablet and on paper. Test results were in perfect or good agreement; only a very small subset of participants received different health classifications. Future research should point out whether this needs further attention. The participants judged the usability and user experience factors the same. A majority of the older adults preferred the use of a tablet for this task, mostly because they found it easier to use. Of course, the findings should be interpreted in the light of their limitations: Our sample size was relatively small and we had to use technology use as a proxy for digital literacy.

Our findings are in line with previous studies that focused on the same topic [2,3]. It shows that switching from paper to tablet computers for administering health-screening surveys among older adults does not affect the reliability of the test outcomes and can count on high acceptance.

Acknowledgement

This study was undertaken within the context of the FOCUS project (EU joint action: 664367 / FOCUS).

References

- [1] D. Dasgupta, B. Chaudhry, E. Koh, and N. Chawla, A survey of tablet applications for promoting successful aging in older adults, *IEEE Access* **4** (2016), 9005-9017.
- [2] R. Hess, A. Santucci, K. McTigue, G. Fischer, and W. Kapoor, Patient Difficulty Using Tablet Computers to Screen in Primary Care, *J Gen Intern Med* **23** (2008), 476-480.
- [3] J. Fanning and E. McAuley, A Comparison of Tablet Computer and Paper-Based Questionnaires in Healthy Aging Research, *JMIR Res Prot* **3** (2014), e38.
- [4] L. van Velsen, et al., A Community-Based, Technology-Supported Health Service for Detecting and Preventing Frailty among Older Adults: A Participatory Design Development Process, *J Aging Res* **2015** (2015), article no. 216084.
- [5] T.K. Malmstrom, D.K. Miller, E.M. Simonsick, L. Ferrucci, and J.E. Morley, SARC-F: a symptom score to predict persons with sarcopenia at risk for poor functional outcomes, *J Cachexia Sarcopenia Muscle* **7** (2016), 28-36.
- [6] J.E. Ware and C.D. Sherbourne, The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection, *Med Care* **30** (1992), 473-483.
- [7] R. Brooks and E. Group, EuroQol: the current state of play, *Health Policy* **37** (1996), 53-72.
- [8] J.-W. Van 't Klooster and M. Vollenbroek-Hutten, Design and evaluation of a flexible web-based screening service for clinimetrics, *Int J Comput Sci Inform Sys* **11** (2016), 47-59.
- [9] A. Bangor, P.T. Kortum, and J.T. Miller, An Empirical Evaluation of the System Usability Scale, *Int J Human-Comput Interact* **24** (2008), 574-594.
- [10] A.H. Segars and V. Grover, Re-Examining Perceived Ease of Use and Usefulness: A Confirmatory Factor Analysis, *MIS Q* **17** (1993), 517-525.
- [11] H. Van der Heijden, User acceptance of hedonic information systems, *MIS Q.* **28** (2004), 695-704.
- [12] L. van Velsen, T. van der Geest, L. van de Wijngaert, S. van den Berg, and M. Steehouder, Personalization has a Price, Controllability is the Currency: Predictors for the Intention to use Personalized eGovernment Websites, *J Organ Comput Electronic Commerce* **25** (2015), 76-97.