Digital Personalized Health and Medicine
L.B. Pape-Haugaard et al. (Eds.)

doi:10.3233/SHTI200301

Development of a Mobile Digital Manikin to Measure Pain Location and Intensity

Sabine N VAN DER VEER[a,1] Anna L BEUKENHORST[b], S Mustafa ALI[b], Ben JAMES[c], Pedro SILVA[c], John McBETH[b], William G DIXON[b]
[a] *Centre for Health Informatics, University of Manchester, Manchester, UK*
[b] *Centre for Epidemiology Versus Arthritis, University of Manchester, Manchester, UK*
[c] *uMotif, London, UK*

Abstract. Painful conditions are prevalent and substantially contribute to disability worldwide. Digital manikins are body-shaped drawings to facilitate self-reporting of pain. Some of them have been validated, but without allowing for recording of location-specific pain intensity and for use on a smartphone. This paper describes the initial development of a digital pain manikin to support self-reporting of pain location and location-specific intensity using people's own mobile device. Subsequently, we conducted reliability and usability tests with eight researchers and seven patient representatives. Test-retest reliability depended on the manikin's level of detail, but was generally high with most intraclass correlation coefficients ≥0.70 and all similarity coefficients ≥0.50. Participants found the manikin easy to use, but suggested clearer orientation (front/back, certain body locations) and would value additional feedback and diary functions. We will address these issues in the next version of the manikin before conducting a validation study.

Keywords. Pain measurement; Patient-generated health data; mHealth

1. Introduction

Pain is one of the leading causes of disability worldwide [1]. It is a symptom of all musculoskeletal diseases, and prominent in many other conditions, such as cancer. Pain substantially affects people's quality of life and imposes a significant burden on society due to treatment costs and absence from work [2]. To optimally manage pain and understand treatment response, we need insight into how it changes over time. This requires validated instruments to capture key aspects of pain, which often vary per location, such as pain intensity.

Digital manikins are human-shaped drawings where people can shade areas to self-report the location of their pain on a computer. In contrast to their paper counterparts, digital manikins support automated (rather than manual) calculations of e.g. the extent of the affected area. Previous studies presented validated, digital pain manikins [3–5], but none of them facilitated the capture of location-specific pain intensity (i.e. different intensities for different locations within one pain report). They also required a dedicated device with a pre-defined screen size, limiting opportunities for people to use their personal devices, which in turn hampers deploying these manikins for large

[1] Corresponding Author; E-mail: sabine.vanderveer@manchester.ac.uk

mHealth studies. Manikins incorporated in pain self-management apps that are currently available in app stores (e.g. https://www.catchmypain.com/) are unsuitable as research tools because they have not been validated as measurement instruments. We, therefore, developed a digital pain manikin to enable people to self-report the location and location-specific intensity of their pain using any (mobile) device. This paper reports its initial development and usability testing.

2. Methods

2.1. Initial development

We developed a prototype, for Android IOS in the first instance, together with an industry partner (https://www.umotif.com/). The prototype design was guided by the manikin's intended purpose, published digital pain manikin studies [3–5], and input from clinicians and researchers with pain measurement expertise. The manikin was two dimensional, black, and gender-neutral with anchors for chin/back of the head and groin/buttocks; users could select a zoomed-in view on the lower arm and hand. There was a slider to indicate pain intensity before shading an area; score and colour ranged from 1 (grey; no pain) to 10 (red; severe pain). Users could erase incorrectly shared areas by selecting an intensity of 1. There was a visible grid with squares overlaying the manikin; 'active' squares were those that had been shaded with intensity >1. A mechanism for exporting a table with the coordinates and intensity scores of all active squares facilitated a pain score calculation that was not influenced by screen size.

2.2. Selecting the optimal grid

As we expected the grid's granularity to affect both user experience and the accuracy of pain drawings, we evaluated the usability and reliability of three grids with different granularities: coarse (40x80 squares); medium (80x160); and fine (100x200); we tested full body and zoomed-in arm manikins separately. We aimed to identify the grid that best balanced ease of use with drawing accuracy.

We recruited eight researchers from our professional network; five had experience with measuring pain. A sample size of eight was sufficient to detect an intraclass correlation coefficient (ICC) of 0.70 or more [6], found by previous studies [3–5]. After receiving verbal instructions and completing one practice drawing, participants completed a first test set of six drawings (i.e. one for each grid-manikin combination) on a Samsung S6 smartphone by copying an example manikin from paper (Figure 1); we randomised the order in which grids were tested. We then interviewed participants for 30 minutes to explore their views on the ease of use of each grid before asking them to complete a second set of the same six drawings (i.e. retest).

For each grid-manikin combination we evaluated the test-retest reliability of pain extent by calculating (i) the size (i.e. mean of difference) and range (i.e. limit of agreement as $\pm 1.96\times$standard deviation) of within-participant differences between test and retest [7] for the proportion of active squares within the manikin perimeter, and (ii) the ICC [8]. For the test-retest reliability of pain location we calculated the Jaccard similarity coefficient (per [3]), which is the number of squares that were active in both test and retest divided by the number of squares that were active in either test or retest; values range from 0 to 1 (no to complete overlap in pain location). We also calculated mean completion time during the retest and synthesised interview notes per grid.

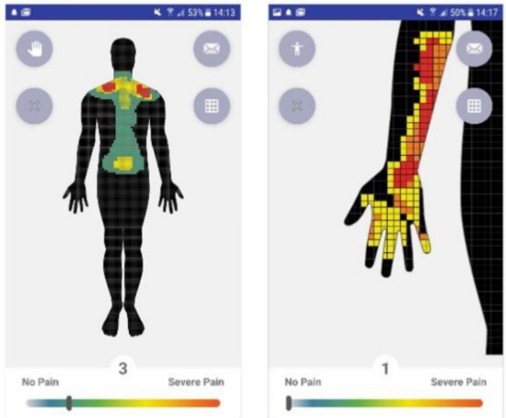

Figure 1. Screenshots of the medium grid body manikin (left panel) and arm manikin (right panel), as completed on a Samsung S6 smartphone by one of the participants

2.3. Exploring usability

We explored the general usability of the digital pain manikin, including the manikin's ease of use; its look and feel; and suggestions for additional functionalities to improve its usefulness for pain management more widely. We collected qualitative data through the researcher interviews (see 2.2), complemented with a 3-hour workshop with seven members of a patient and public involvement (PPI) panel with personal experience of chronic pain. The workshop included individual 30-minute sessions where participants were interviewed and performed a series of tasks with the manikin using the medium grid. They then shared their experiences in a group discussion. We imported written interview and workshop notes into NVivo and analysed them thematically.

3. Results

3.1. Grid selection

Table 1 shows that for pain extent, the coarse grid had the lowest test-retest reliability for both manikins with the highest mean of difference and the lowest ICC. The medium grid performed best for the body manikin, with the lowest mean of difference and the highest ICC. The fine grid showed the best results for the arm manikin. For pain location, the test-retest overlap was comparable between grids for the body manikin. For the arm manikin, the coarse grid showed most overlap, but also the highest proportion of active squares in test and retest. Within grids, completion time varied widely between participants but generally increased with granularity.

Compared to the fine and medium grids, participants found the coarse grid easier to use but also felt it produced the least accurate pain drawings. Others reported a higher risk of unintentionally activating squares, and having to dab (rather than draw) to shade areas. In contrast, they considered the fine grid the most time-consuming, but also the most accurate method that felt most like drawing. Some participants questioned whether real pain reports would require such a fine level of detail. For the medium grid, most participants reported similar experiences as for the fine grid, although some explicitly stated that the medium grid better balanced ease of use with drawing accuracy.

Table 1. Test-retest reliability of pain extent and location, and completion time per grid-manikin type combination

	Coarse grid (40x80)		Medium grid (80x160)		Fine grid (100x200)	
Manikin type →	***Body***	***Arm***	***Body***	***Arm***	***Body***	***Arm***
Extent						
Number of available squares[a)]	1,866	133	6,922	483	10,519	734
Mean (SD) proportion of active squares[b)]						
Test	0.17 (0.09)	0.39 (0.09)	0.16 (0.07)	0.34 (0.06)	0.14 (0.08)	0.34 (0.10)
Retest	0.16 (0.06)	0.44 (0.07)	0.13 (0.06)	0.36 (0.05)	0.14 (0.05)	0.37 (0.07)
Mean (1.96×SD) of difference[c)]	4.39 (5.13)	5.83 (12.62)	2.31 (3.90)	4.48 (6.92)	2.60 (6.72)	3.90 (7.98)
ICC (95% CI)	0.77 (0.27 – 0.95)	0.48 (-0.23 – 0.87)	0.90 (0.63 – 0.98)	0.51 (-0.20 – 0.87)	0.81 (0.36 – 0.96)	0.80 (0.33 – 0.96)
Location						
Mean (SD) Jaccard coefficient	0.50 (0.15)	0.71 (0.12)	0.51 (0.16)	0.59 (0.12)	0.54 (0.15)	0.59 (0.13)
Completion time (retest)						
Mean (SD) number of seconds	64.0 (57.4)	47.5 (21.7)	72.0 (52.2)	68.2 (40.7)	82.1 (33.7)	70.1 (42.2)

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient; SD, standard deviation

[a)] Total number of squares in the grid that fall within the manikin perimeter

[b)] Proportion of available squares shaded with pain intensity >1

[c)] Difference and limit of agreement between test—retest [7] in the proportion of active squares

3.2. Usability of the manikin

All participants found completing a manikin report easy and expected basic user instructions and some practice to be sufficient for most people to use it unsupervised. Some participants made sweeping movements to activate squares, while others dabbed, especially PPI members. Dabbing was more time consuming, which may explain why some PPI members would prefer a coarser grid. Many participants ‘pinched’ in an attempt to zoom in on certain body locations, but were equally happy to use the predefined zoomed-in views. Some participants struggled to activate squares as intended, especially on the border of the manikin perimeter. The method for deactivating squares was considered acceptable, although some suggested an eraser or ‘undo’ function.

Most participants saw the manikin as male, and PPI members additionally characterised it as “elongated” and “a frog man”. Additionally, many had trouble distinguishing front from back and therefore suggested adding labels (front/back, left/right), as well as more and clearer anchors (e.g. chest, lower back). They generally liked the grid being visible, and most considered the colour scheme for pain intensity appropriate. However, some found green colours for scores of 4-5 counterintuitive and many would prefer an intensity score of ‘0’ (rather than ‘1’) to reflect ‘no pain’. No one expressed concerns with the manikin being black.

Participants suggested several additional functionalities for improving the manikin’s usefulness for pain management, including: option to record additional pain aspects (e.g. location-specific pain type (throbbing, stabbing, etc.)), pain interference with daily life); feedback of pain reports over time (e.g. heat maps); diary to record free text information alongside a manikin report (e.g. description of the pain experience); and an option to share manikin reports with your doctor to inform (tele)consultations.

4. Discussion and conclusion

We developed and tested a digital pain manikin that facilitates measuring self-reported pain location and location-specific intensity on any device. Study participants found the manikin easy to use, and it generally showed moderate to very high reliability.

A limitation of our study is that we asked participants to copy a detailed manikin example, which might be more complex than reporting a personal pain experience. They also had to do this 2×6 times instead of the usual 2×1. The increased task complexity and repetitiveness may explain why our ICCs for pain extent did not exceed the 0.90 reported by other studies [3–5]. The reliability of pain location seemed less affected by this, with Jaccard similarity coefficients in some other studies not exceeding 0.50 [3,5].

Overall, the fine grid came out as most reliable across manikins, but the usability tests suggested that it might be too difficult and time consuming to use. The next version of the manikin will, therefore, use the medium grid, as well as addressing most of the issues and suggestions this study identified, such as: providing written user instructions; improving the look and feel of the manikin; facilitating the capture of additional pain aspects, and including diary and feedback functions. Through continued collaborations with patients, clinicians, researchers and industry, we will use the new manikin version to collect digital pain reports in a range of painful conditions to (i) reassess the manikin's reliability for capturing people's personal pain experience, and (ii) develop and validate algorithms for calculating quantitative manikin scores that combine pain location, extent and intensity. Ultimately, we expect this contribute to our digital pain manikin becoming a new standard for digital pain reports in large mHealth studies, while also being suitable for supporting clinical care and self-management.

Acknowledgments This project was supported by the Centre for Epidemiology Versus Arthritis (ref 20380).

References

[1] S.L. James, D. Abate, K.H. Abate, et al., Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017, *Lancet*. (2018). doi:10.1016/S0140-6736(18)32279-7.

[2] Institute of Medicine, Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research, The National Academies Press, Washington, DC, 2011. doi:10.17226/13172.

[3] M. Barbero, F. Moresi, D. Leoni, R. Gatti, M. Egloff, and D. Falla, Test-retest reliability of pain extent and pain location using a novel method for pain drawing analysis, *Eur. J. Pain (United Kingdom)*. (2015). doi:10.1002/ejp.636.

[4] D. Leoni, D. Falla, C. Heitz, G. Capra, R. Clijsen, M. Egloff, C. Cescon, J.P. Baeyens, and M. Barbero, Test–retest Reliability in Reporting the Pain Induced by a Pain Provocation Test: Further Validation of a Novel Approach for Pain Drawing Acquisition and Analysis, *Pain Pract.* (2017). doi:10.1111/papr.12429.

[5] M.A. Bujang, and N. Baharum, A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review, *Arch. Orofac. Sci.* (2017).

[6] J.M. Bland, and D.G. Altman, Measuring agreement in method comparison studies, *Stat. Methods Med. Res.* (1999). doi:10.1177/096228029900800204.

[7] T.A. Neubert, M. Dusch, M. Karst, and F. Beissner, Designing a tablet-based software app for mapping bodily symptoms: Usability evaluation and reproducibility analysis, *JMIR MHealth UHealth*. (2018). doi:10.2196/mhealth.8409.

[8] S. Chinn, The assessment of methods of measurement, *Stat. Med.* **9** (1990) 351–362.