

1 **Measurement of equivalence between the web and paper versions of the Japanese**
2 **Orthopaedic Association Hip Disease Evaluation Questionnaire (JHEQ)**

3 **Abstract**

4 *Objectives:* Digitized patient-reported outcome may be beneficial for physicians and patients. The
5 Japanese Orthopaedic Association Hip Disease Evaluation Questionnaire (JHEQ) can only be completed
6 with paper and pencil (pJHEQ). We newly developed a web version of the JHEQ (wJHEQ). This study
7 aimed to determine whether the scores obtained with the wJHEQ are equivalent to those from the pJHEQ,
8 how much the wJHEQ would decrease missing answers, and which JHEQ the participants preferred to
9 use.

10 *Methods:* To measure equivalence between the pJHEQ and wJHEQ, we evaluated the mean score
11 difference for each subscale (pain, movement, mental) and Visual Analogue Scale (VAS) (satisfaction,
12 right hip pain, left hip pain) and then assessed the intraclass correlation coefficients (ICC) between the
13 two scores. ICC values ≥ 0.75 were defined as excellent agreement. We used Bland-Altman analysis to
14 assess level of agreement between the values of the two questionnaires. We compared the number of
15 incomplete forms and amount of missing data between the two questionnaires. We investigated ease of
16 use by asking the participants which format was easier to use.

17 *Results:* This study comprised 113 patients (mean age 58.1 years, 81% female) with hip disease. Mean
18 score differences for each subscale between the wJHEQ and pJHEQ were not significantly different. The
19 values of ICC for each subscale and each VAS were all > 0.75 . All 113 participants completed the wJHEQ
20 questionnaire, whereas nine patients did not complete the pJHEQ form. There was a significant statistical
21 difference between the completion rate of the wJHEQ and that of the pJHEQ ($p=0.0017$). Fifty-seven
22 participants (55%) preferred the wJHEQ, whereas thirty-three participants (32%) preferred the pJHEQ.

23 *Conclusions:* The wJHEQ was found to be equivalent to the original pJHEQ. The wJHEQ significantly
24 decreased the numbers of missing answers and incomplete forms. The participants felt ease of use was
25 nearly equivalent. The wJHEQ might help facilitate more complete assessments in clinical trials and
26 research.

27 **1. Introduction**

28 Patient-reported outcomes (PROs) are increasingly used in clinical trials and other research, including
29 in the orthopedics community, as tools to objectively assess the patients' subjective data, which indicates
30 their pain and treatment satisfaction [1].

31 The Japanese Orthopaedic Association Hip Disease Evaluation Questionnaire (JHEQ) was developed
32 as one of the PROs for patients with diseases of the hip [2]. The JHEQ consists of three subscales: pain
33 (28 points), movement (28 points), and mental (28 points). Higher scores of the JHEQ indicate a better
34 outcome. In addition, the JHEQ has a characteristic feature: it includes three visual analogue scales (VAS)
35 to measure patients' satisfaction and their left and right hip pain.

36 Currently, several studies have recommended that electronic PROs should be introduced because
37 electronic systems allow efficient standardized assessments, decreased response burden, improved ease of
38 use, and fewer missing data compared with paper-based PROs measurements [3-5]. Until now, the data
39 for the JHEQ could only be entered manually with paper and a pencil (pJHEQ). We originally developed
40 an electronic version of the JHEQ in 2015. However, in our previous study we could not demonstrate
41 reliability and validity of the electronic version because of the use of a slider bar system to input VAS [6].
42 Thus, we have newly developed a web version of the JHEQ (wJHEQ) that uses a touch panel system to
43 input VAS instead of a slider bar.

44 This study aimed to determine whether the scores obtained via the wJHEQ are equivalent to those
45 obtained from the pJHEQ, how much the wJHEQ would decrease the number of missing answers, and
46 participant preference for the wJHEQ or pJHEQ.

47

48 **2. Material and Methods**

49 This study was approved by our institution's research ethics boards, and all participants provided
50 written informed consent.

51

52 *2.1 wJHEQ*

53 We developed the wJHEQ to reproduce as faithfully as possible the pJHEQ. We improved the data
54 input system for VAS in the wJHEQ by adopting a touch panel system instead of the slider bar system
55 reported in our previous study [6]. We uploaded the wJHEQ on the Internet service at

56 <http://www.jheq.sakura.ne.jp>. The wJHEQ has several characteristic features: 1) it provides an automatic
57 alert if the participant fails to answer a question; 2) it transfers the recorded data to the examiner's email
58 address as a text file via the Internet; 3) the results of the wJHEQ are displayed immediately at the end of
59 the questionnaire; and 4) the personal data that is inputted into the browser is erased at the end of every
60 use to protect the patient's personal information.

61

62 *2.2 Participants*

63 From December 2017 to February 2018, eligible patients were recruited at their regularly scheduled
64 follow-up visit at our outpatient clinic. The inclusion criteria were at least a 3-month history of
65 symptomatic osteoarthritis of the hip or a history of any surgery for the hip such as total hip arthroplasty
66 and osteotomy, and the ability to comprehend the Japanese language. Exclusion criteria were age under
67 20 years (one participant), inability to use a computer for any reason (one participant because of eye
68 impairment), and no desire to participate in this study (five participants because they had no time to
69 complete both the pJHEQ and wJHEQ). Participants were invited to complete both the pJHEQ and the
70 wJHEQ. They received either the pJHEQ or the wJHEQ first, the order of which was randomly assigned
71 by a computer software program according to age, sex, and disease [7]. There was a 15-minute interval
72 between the two assessments, as described in a previous study by Bischoff-Ferrari et al. [8]. All patients
73 had completed the original pJHEQ at their first visit to our outpatient clinic.

74

75 *2.3 Data analysis*

76 We defined the values of the pJHEQ as being the gold standard because the reliability and validity of
77 the pJHEQ have been confirmed previously [9]. To measure equivalence between the pJHEQ and the
78 wJHEQ, we evaluated the mean score difference for each subscale and the three VASs. We used the
79 paired *t*-test to compare the wJHEQ to the pJHEQ. We then assessed the intraclass correlation
80 coefficients (ICC) between the two scores. We considered ICC values ≥ 0.75 to indicate excellent
81 agreement and a value < 0.75 to indicate poor to moderate agreement [5]. In addition, we used
82 Bland-Altman analysis [10] to assess the level of agreement between the values of the pJHEQ and those
83 of the wJHEQ. The smallest detectable difference is defined as $1.96 * (\text{standard deviation of the$

84 difference between the scores), and also the limits of agreement, according to prior studies [11,12]. The
85 agreement between the scores was also illustrated by Bland-Altman plots [10].

86 The wJHEQ is equipped with an automatic alert when the participant does not respond to the
87 questionnaire. Furthermore, we prohibited advancement to the next question until the current question
88 was answered. We therefore expected use of the wJHEQ to decrease the amount of missing data and the
89 numbers of incomplete forms. Thus, we compared the number of incomplete forms and the amount of
90 missing data between the pJHEQ and the wJHEQ with Fisher's exact test.

91 To investigate ease of use, the participants were asked the following question: "Which format was
92 easier for you to use: the computer format, paper format or both equally?" The participants could only
93 choose only one answer according to a previous study [8]. In addition, we tried to clarify the effect of
94 differences in the participants' background on preference. We grouped the participants into three groups
95 according to their preferred input method and compared their backgrounds. We used analysis of variance
96 (ANOVA) for age followed by the Holm test as a post hoc test and Fisher's exact test for gender and
97 disease.

98 The level of statistical significance was set at $p < 0.05$. All statistical analyses were performed with
99 EZR (Saitama Medical Center, Jichi Medical University) [13].

100

101 **3. Results**

102 In total, 113 patients participated in our study. Table 1 shows the characteristic of the study population.
103 The age range was 26 to 80 years, with a mean of 58.1 years. Approximately 81% of the participants were
104 women. Most of the participants had osteoarthritis of the hip. Three participants had a bone tumor, of
105 whom two had pigmented villonodular synovitis of the hip, and one had a chondroblastoma of the
106 femoral head. The demographics of the two groups according to their preference for the wJHEQ or
107 pJHEQ are shown in Supplemental Table 1.

108 All 113 participants completed the wJHEQ questionnaire, whereas nine patients, one man and eight
109 women, did not completely fill out the pJHEQ form. On the pJHEQ, four participants missed one item,
110 one missed two items, one missed five items, one missed six items, one missed eight items, and one
111 missed 12 items. There was a significant difference between the completion rate of the wJHEQ and that

112 of the pJHEQ ($p=0.0017$). The demographics of the participants who could not complete the pJHEQ are
113 shown in Supplemental Table 2.

114 After excluding the data from these nine participants, the data of the remaining 104 participants were
115 evaluated for further study. First, we assessed the measurement of the equivalence between the wJHEQ
116 and the pJHEQ. The means and standard deviations of the JHEQ subscale scores (pain, movement, and
117 mental) and the value of VASs (satisfaction, right hip pain, and left hip pain) are presented in Table 2.
118 The difference in each mean between the wJHEQ and pJHEQ was not significantly different ($P=0.40$,
119 0.15 , 0.52 , 0.62 , 0.65 , and 0.61 , respectively). The differences in each VAS varied from 1.6 mm to 2.0
120 mm (satisfaction: 2.0, right hip pain: 1.6, and left hip pain: 1.8). The ICC values of each subscale and
121 each VAS were all >0.75 : pain: 0.95, movement: 0.97, mental: 0.96, satisfaction: 0.96, right hip pain:
122 0.97, and left hip pain: 0.97. The agreement between scores obtained by the wJHEQ and pJHEQ is
123 illustrated by Bland-Altman plots for the VASs of patient satisfaction, right hip pain, and left hip pain in
124 Fig. 1. There was agreement between the VAS values of the pJHEQ and those of the wJHEQ in most
125 cases (98 cases [94.2%] for satisfaction, 97 cases [93.2%] for right hip pain, and 98 cases [94.2%] for left
126 hip pain). There were no significant differences between the wJHEQ and pJHEQ scores in the two groups
127 (Supplemental Tables 3 and 4).

128 Finally, in terms of ease of use of the two questionnaires, slightly more than half of the participants,
129 57 (55%), preferred the wJHEQ, whereas 33 (32%) preferred the pJHEQ, and 14 (13%) preferred both
130 equally well. The group of participants preferring the pJHEQ was significantly older than that preferring
131 the wJHEQ or both ($P=0.003$ and 0.017 , respectively) (Fig. 2). There was also no significant relation of
132 gender or disease with their preferred choice ($P=0.23$ and 0.079 , respectively) (Tables 3 and 4).

133

134 **4. Discussion**

135 This study showed that both the subscale scores and VAS values of the wJHEQ were similar to those
136 of the pJHEQ, which is the original paper-based JHEQ. Use of the wJHEQ could significantly reduce the
137 amount of incomplete data. In addition, the participants found the wJHEQ to be an easy-to-use tool
138 similar to the pJHEQ.

139 Our results suggested that the wJHEQ was equivalent to the pJHEQ. A recent meta-analysis showed
140 that the electronic data collection for PROs was adequately validated in several studies. [14]. However,

141 the electronic version of the JHEQ developed in our previous study did not provide reliable values
142 because of the slider bar system used to capture the VAS data. In the previous study, we concluded that
143 the target of the slider bar was hidden by the examinee's finger when inputting the data. The lack of
144 visual feedback resulted in unreliable VAS values [6]. We therefore newly established a touch panel
145 system to capture this data, and with its incorporation, we were able to show equivalence between the
146 wJHEQ and the pJHEQ.

147 The VAS values for satisfaction and hip joint pain showed differences from 1.6 mm to 2.0 mm
148 between the wJHEQ and the pJHEQ. In several prior studies, the minimum clinically important difference
149 for pain was defined to vary from 9 to 14 [15,16]. This result suggested that the difference in our VAS
150 values did not have clinical importance.

151 In our study, 57 participants (55%) said it was easier to input data via a computer, 14 participants
152 (13%) thought both methods were equally easy to use, and 33 participants (32%) preferred the pJHEQ. In
153 other previous studies, 65% of participants preferred the digital format [8], and 92% of the subjects either
154 stated that both computer formats and paper formats were equally easy to use or that the computer format
155 was easier to use [10]. Our system established in this research resulted in similar ease of use with the
156 pJHEQ. The group of participants preferring the pJHEQ was significantly older than that of the
157 participants preferring the wJHEQ or both ($P=0.003$ and 0.017 , respectively) (Fig. 2). A previous study
158 comparing young and older participants found that older adults feel less confident than their younger
159 counterparts in their own computer knowledge [17], suggesting that the paper version might be preferable
160 for elderly participants.

161 This study has several limitations. First, our results may be susceptible to selection bias because of
162 participant selection. An assistant initially explained how to use the computer version of the test to the
163 participants, and if they had any questions, the assistant offered appropriate help. As the applicants in this
164 study were limited to those who wanted to participate, persons not confident in the use of computers
165 chose not to participate. This resulted in most participants being familiar with computer use. In addition,
166 this study was conducted in only one institution, which could also result in selection bias. Second, the
167 wJHEQ forces the participants to answer all questions. If there are questions that the participants could
168 not answer, our system may result in incorrect answers. The automatic alert may also mentally stress the
169 patients because of forcing them to complete the form. To avoid this problem, other researchers designed

170 a system to have a skip question button or “no answer” button [8]. It is thus necessary to determine how
171 to manage the “no answer” issue in the wJHEQ. Third, we enforced a 15-minute interval between the test
172 and the retest as described in a previous study by Bischoff-Ferrari et al. [8]. This short interval might
173 affect the results showing high correlation and statistically insignificant differences in our study. Fourth,
174 we did not interview the participants as to their previous experience using any electronic devices. A
175 cross-sectional study revealed that previous experience with computer use was related to a reduction in
176 anxiety when using a computer [18]. This may also have affected the results of this study. Fifth, no
177 patients had rheumatoid arthritis (RA) in our study because the patients with RA were followed by
178 another group in our institution. It is possible that hand impairment could affect the ability to input data
179 on a tablet computer. Future research should also include patients with hand impairment.

180 In conclusion, the wJHEQ showed equivalency with the pJHEQ, the original manual entry version
181 requiring paper and pencil. The wJHEQ could significantly decrease the numbers of missing answers and
182 incomplete forms. The participants felt that both questionnaires were about equally easy to use, however
183 the paper version might be preferable for elderly participants.

184 The wJHEQ may facilitate more complete assessment in clinical trials and other research.

185

186 **Conflicts of interest**

187 There are no conflicts of interest.

188

189 **References**

- 190 [1] Ayers DC, Zheng H, Franklin PD. Integrating patient-reported outcomes into orthopaedic clinical
191 practice: Proof of concept from FORCE-TJR. *Clin Orthop Relat Res.* 2013;471:3419–25.
- 192 [2] Matsumoto T, Kaneuji A, Hiejima Y, Sugiyama H, Akiyama H, Atsumi T, et al. Japanese
193 Orthopaedic Association Hip Disease Evaluation Questionnaire (JHEQ): a patient-based
194 evaluation tool for the hip-joint disease. The Subcommittee on Hip Disease Evaluation of the
195 Clinical Outcome Committee of the Japanese Orthopaedic Association. *J Orthop Sci.* 2012;17:25–
196 38.
- 197 [3] Jensen RE, Snyder CF, Abernethy AP, Basch E, Potosky AL, Roberts AC, et al. Review of
198 electronic patient-reported outcomes systems used in cancer clinical care. *J Oncol Pract.* 2014;
199 10:e215–22.
- 200 [4] Shervin N, Dorrwachter J, Bragdon CR, Shervin D, Zurakowski D, Malchau H. Comparison of
201 paper and computer-based questionnaire modes for measuring health outcomes in patients
202 undergoing total hip arthroplasty. *J Bone Joint Surg Am.* 2011; 93:285–93.
- 203 [5] Marsh JD, Bryant DM, Macdonald SJ, Naudie DD. Patients respond similarly to paper and
204 electronic versions of the WOMAC and SF-12 following total joint arthroplasty. *J Arthroplasty.*
205 2014, 29:670–3.
- 206 [6] Takegami Y, Seki T, Kaneuji A, Nakao A, Hasegawa Y, Ishiguro N. Validity of a tablet computer
207 version of the Japanese Orthopaedic Association hip disease evaluation questionnaire: a pilot
208 study. *Nagoya J Med Sci.* 2016;78:237–44.
- 209 [7] Saghaei M, Saghaei S. Implementation of an open-source customizable minimization program for
210 allocation of patients to parallel groups in clinical trials. *J Biomed Eng.* 2011; 04:734–9.
- 211 [8] Bischoff-Ferrari HA, Vondechend M, Bellamy N, Theiler R. Validation and patient acceptance of
212 a computer touchscreen version of the WOMAC 3.1 osteoarthritis index. *Ann Rheum Dis.*
213 2005;64:80–4.
- 214 [9] Seki T, Ishiguro N, Hasegawa Y, Ikeuchi K, Hiejima Y. Reliability and validity of the Japanese
215 Orthopaedic Association hip disease evaluation questionnaire (JHEQ) for patients with hip disease.
216 *J Orthop Sci.*2013;18:782–7.

- 217 [10] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of
218 clinical measurement. *Lancet*. 1986;327:307–10.
- 219 [11] Kvien TK, Mowinckel P, Heiberg T, Dammann KL, Dale Ø, Aanerud GJ, et al. Performance of
220 health status measures with a pen based personal digital assistant. *Ann Rheum Dis*. 2005;64:1480–
221 4.
- 222 [12] Reinhart K, Kuhn HJ, Hartog C, Bredle DL. Continuous central venous and pulmonary artery
223 oxygen saturation monitoring in the critically ill. *Intensive Care Med*. 2004;30:1572–8.
- 224 [13] Kanda Y. Investigation of the freely available easy-to-use software 'EZR' for medical statistics.
225 *Bone Marrow Transplant*. 2013;48:452e8.
- 226 [14] Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil
227 administration of patient-reported outcome measures: A meta-analytic review. *Value Health*.
228 2008;11:322–33.
- 229 [15] Kelly AM. Does the clinically significant difference in visual analog scale pain scores vary with
230 gender, age, or cause of pain? *Acad Emerg Med*. 1998;5:1086–90.
- 231 [16] Bodian CA, Freedman G, Hossain S, Eisenkraft JB, Beilin Y. The visual analog scale for pain:
232 clinical significance in postoperative patients. *Anesthesiology*. 2001;95:1356–61.
- 233 [17] Marquié JC, Jourdan-Boddaert L, Huet N. Do older adults underestimate their actual computer
234 knowledge? *Behav Inf Technol*. 2010;21:273–80.
- 235 [18] Igbaria M, Chakrabarti A. Computer anxiety and attitudes towards microcomputer use. *Behav Inf*
236 *Technol*. 1990;9(3):229–41.
- 237

238 **Figure legend**

239 **Fig. 1.** Scatter plot and Bland-Altman plot of the comparison of each VAS between the wJHEQ and
240 pJHEQ. **A** Patient satisfaction, **B** VAS of right hip pain, and **C** VAS of left hip pain.

241

242 **Fig. 2.** Age difference of the participants according to preference for the input method. Values are the
243 mean and standard deviation. *P value < 0.05 for post hoc analysis.