Clin Orthop Relat Res (2013) 471:4037–4044
DOI 10.1007/s11999-013-3213-2

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

CLINICAL RESEARCH

Abbreviated Psychologic Questionnaires Are Valid in Patients With Hand Conditions

Arjan G. J. Bot MD, Stéphanie J. E. Becker MD, C. Niek van Dijk MD, PhD, David Ring MD, PhD, Ana-Maria Vranceanu PhD

Received: 15 February 2013 / Accepted: 25 July 2013 / Published online: 3 August 2013

Abstract

Background The Pain Catastrophizing Scale (PCS) and Short Health Anxiety Inventory (SHAI) can help hand surgeons identify opportunities for psychologic support, but they are time consuming. If easier-to-use tools were available and valid, they might be widely adopted.

Questions/purposes We tested the validity of shorter versions of the PCS and SHAI, the PCS-4 and the SHAI-5, by assessing: (1) the difference in mean scaled scores of the short and long questionnaires; (2) floor and ceiling effects between the short and long questionnaires; (3) correlation between the short questionnaires and the outcome measures (an indication of construct validity); and (4) variability in disability and pain, between the short and long questionnaires.

Methods One hundred sixty-four new or followup adult patients in one hand surgery clinic completed the SHAI-18, SHAI-5, PCS-13, PCS-4, Patient Health Questionnaire (PHQ)-9, PHQ-2, DASH, and QuickDASH questionnaires, and an ordinal pain scale, as part of a prospective cross-sectional study. Mean scores for the short and long questionnaires were compared with paired t-tests. Floor and ceiling effects were calculated. Pearson's correlation was used to assess the correlation between the short and long questionnaires and with outcome measures. Regression analyses were performed to find predictors of pain and disability.

Results There were small, but significant differences between the mean scores for the DASH and QuickDASH (QuickDASH higher), SHAI-18 and SHAI-5 (SHAI-18 higher), and PCS-13 and PCS-4 (PCS-4 higher), but not the PHQ-9 and PHQ-2. Floor effects ranged between 0% and 65% and ceiling effects between 0% and 3%. There were greater floor effects for the PHQ-2 than for the PHQ-9, but floor and

The institution of one of the authors (CND) has received, during the study period, funding from Boehringer Ingelheim (Alkmaar, the Netherlands), Biomet (Dordrecht, the Netherlands), Stryker (Waardenburg, the Netherlands), and Amgen BV (Breda, the Netherlands).
One of the authors (CND) certifies that he or she, or a member of his or her immediate family, has received or may receive payments or benefits, during the study period, an amount of USD 10,000 to 100,000, from Smith & Nephew BV (Hoofddorp, the Netherlands).
One of the authors (DR) certifies that he or she, or a member of his or her immediate family, has received or may receive payments or benefits, during the study period, an amount of USD 10,000 to 100,000, from Skeletal Dynamics; less than USD 10,000 from Wright Medical Technology, Inc (Arlington, TN, USA) and Biomet, Inc (Warsaw, IN, USA); and less than USD 10,000 from AO North America (Paoli, PA, USA) and AO International (Dübendorf, Switzerland).
One of the authors (AGJB) certifies that he has received, during the study period, funding from Prins Bernhard Cultuurfonds/Banning-de Jong Fonds (Amsterdam, the Netherlands), VSBfonds (Utrecht, the Netherlands), and Anna Fonds (Oegstgeest, the Netherlands).
One of the authors (SJEB) certifies that he has received, during the study period, funding from Anna Fonds, Genootschap Noorthey (Bussum, the Netherlands), and Vreedefonds (Voorburg, the Netherlands).
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research* editors and board members are on file with the publication and can be viewed on request.
Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.
The study was performed at the Orthopaedic Hand and Upper Extremity Service, Massachusetts General Hospital, Boston, MA, USA.

A. G. J. Bot, S. J. E. Becker, D. Ring (✉)
Department of Orthopaedic Surgery, Harvard Medical School, Boston, MA, USA
e-mail: dring@partners.org

A. G. J. Bot, S. J. E. Becker, D. Ring
Orthopaedic Hand and Upper Extremity Service, Massachusetts General Hospital, 55 Fruit Street, Yawkey Center, Suite 2100, Boston, MA 02114, USA

ceiling effects were otherwise comparable for the other short and long questionnaires. All questionnaires showed convergent and divergent validity and criterion validity was shown in multivariable analyses.

Conclusions Content validity, construct convergent validity, and criterion validity were established for the short versions of the PCS and SHAI. Using shorter forms creates small differences in mean values that we believe are unlikely to affect study results and are more efficient and advantageous because of the decreased responder burden.

Level of Evidence Level III, diagnostic study. See Instructions for Authors for a complete description of levels of evidence.

Introduction

The DASH questionnaire, Pain Catastrophizing Scale (PCS), Patient Health Questionnaire (PHQ), and Short Health Anxiety Inventory (SHAI) are reliable and valid questionnaires that have proved useful in hand surgery research [1, 19–21, 27, 28, 32]. These measures also might help hand surgeons identify opportunities for psychologic support in daily practice, but one should limit the burden on the patient [23]. Asking patients to complete the 30-item DASH, 9-item PHQ-9, 13-item PCS, or 18-item SHAI will take most patients approximately 20 minutes.

Shorter forms of the original questionnaire have been developed for the DASH (the QuickDASH) and the PHQ-9 (the PHQ-2) [2, 9, 12, 14]. A previous unpublished study by our group involved the development of the shorter versions of the PCS and SHAI, the PCS-4 and SHAI-5. In the current study, we sought to validate these abbreviated measures.

Specifically, we assessed: (1) differences in mean scaled scores of the short and long questionnaires; (2) floor and ceiling effects between the short and long questionnaires; (3) correlation between the short questionnaires and the outcome measures (an indication of construct validity); and (4) variability in disability and pain between the short and long questionnaires.

C. N. van Dijk
Department of Orthopaedic Surgery, Academic Medical Center, Amsterdam, The Netherlands

A.-M. Vranceanu
Department of Psychology, Harvard Medical School, Boston, MA, USA

A.-M. Vranceanu
Department of Behavioral Medicine, Benson Henry Mind Body Institute, Massachusetts General Hospital, Boston, MA, USA

Patients and Methods

For this study, we used the prospectively enrolled cross-sectional cohort of 164 patients used to develop the shortened version of the questionnaires using the methodology as described by McCracken and Dhingra [16]. This methodology selects questions with the highest value of inter-item correlation in each of the subscales of the questionnaire (three for the PCS and 2 for the SHAI-18) [16]. When a chosen question would lead to decreased variance of the subscale or reduced scale variability, the item with the second highest correlation is chosen [16]. This method resulted in a four-item PCS (PCS-4) and a five-question SHAI (SHAI-5). Our institutional review board approved the study and all patients provided informed consent. We approached patients visiting our orthopaedic hand and upper extremity clinic regardless of their diagnosis. Adult, new or followup, English-speaking patients were enrolled between December 2009 and June 2011. A total of 236 consecutive patients were asked to participate in our study, 66 declined based on the time necessary to participate in the study (approximately 20 minutes). From the cohort of 170 patients, one patient withdrew while completing the questionnaires and five patients were excluded because they had invalid DASH questionnaires, leaving 164 patients in the study.

The cohort consisted of 54% women, had a mean age of 51 years, and had various diagnoses (fracture, carpal tunnel syndrome, osteoarthritis, trigger finger, nonspecific pain, and other diagnoses) (Table 1). Forty-one percent of the patients were new and 59% were returning patients. Fifty-four patients had previous surgery.

Arm-specific disability was evaluated with the DASH questionnaire [10]. The DASH contains 30 questions, which are answered on 5-point Likert scales, and is scaled from 0 to 100 points, where a higher score indicates greater disability [10]. The QuickDASH [2] is a validated short form of the DASH, containing 11 of the original 30 questions [9].

Pain was scored on an ordinal scale from 0 to 10, where 0 was no pain and 10 the worst pain the patients ever experienced. Three patients had a missing pain score. We used the method of regression prediction plus error imputation [8, 25] for these three missing questionnaires.

To assess symptoms of depression, the validated PHQ-9 [11] was used. The original long questionnaire contains nine questions, answered on a 4-point Likert scale, where 0 represents "not at all" and 4 "nearly every day" [29]. The first two questions of the PHQ-9 were used [12, 14] to form the PHQ-2 of the original questionnaire. The PHQ-2 is a validated screening tool for depression [12, 14]. To compare the differences between the two scores, the mean scores for the PHQ-2 and PHQ-9 were calculated. Five patients had one missing question on the PHQ-9 so we

Table 1. Patient demographics (n = 164)

Parameter	Value
Age (years)*	51 ± 15 (18–99)
Months since first pain experience*	18 ± 39 (0–240)
Number of prior surgeries*	1.3 ± 0.7 (1–5)
Patient self-rating of health (points)*	8.4 ±1.5 (3–10)
Education (years)*	15 ± 2.6 (5–20)
Sex	
Male	28 (34%)
Female	54 (66%)
Visit type	
First	67 (41%)
Followup	97 (59%)
Prior surgery	
Yes	54 (33%)
No	110 (67%)
Other pain conditions	
Yes	59 (36%)
No	105 (64%)
Smoking	
Yes	19 (12%)
No	145 (88%)
Marital status	
Single	49 (30%)
Living with partner	5 (3.0%)
Married	74 (45%)
Separated/divorced	26 (16%)
Widowed	10 (6.1%)
Working status	
Full-time	80 (49%)
Part-time	23 (14%)
Homemaker	6 (3.7%)
Retired	29 (18%)
Unemployed	23 (14%)
Workers Compensation	3 (1.8%)

* Values are expressed as mean ± SD, with range in parentheses.

imputed the mean of the patients' other questions for these missing values [7].

The 13-item PCS (PCS-13) was used to measure misinterpretation or overinterpretation of nociception (catastrophic thinking). This questionnaire contains 13 questions, answered on a 4-point Likert scale, where 0 represents "not at all" and 4 "all the time"; thus, total score can range from 0 to 52 [26]. A higher score indicates a more amplified negative orientation toward pain [26]. For this study, we used the PCS-4, which contains Questions 3, 6, 8, and 11 of the original scale (Table 2). Five patients had one missing question on the PCS-13, and we used mean imputation of the average of the other questions to deal with these missing values [7].

Table 2. Questions in the PCS-4 and SHAI-5

Question	Description
PCS-4	
3	It's terrible and I think it's never going to get any better
6	I become afraid that the pain may get worse
8	I anxiously want the pain to go away
11	I keep thinking about how badly I want the pain to stop
SHAI-5	
2	I notice aches and pains…
3	….aware of bodily sensation or changes
12	I ….think I have a serious illness
15	If I had a serious illness I would…
17	A serious illness would ruin…..aspects of my life

PCS-4 = four-item Pain Catastrophizing Scale; SHAI-5 = five-item Short Health Anxiety Index.

The 18-item SHAI (SHAI-18) was used to assess health anxiety (the sense that one has a serious health problem) [22]. Each item has a score ranging from 0 to 3 [22]. We selected Questions 2, 3, 12, 15, and 17 to form the SHAI-5 (Table 2). There were nine missing questions for five patients. We substituted the missing answers with the patient's mean of the other answers in the questionnaire [7].

We extracted the PCS-4 and SHAI-5 from the same surveys [16]. We chose not to burden the patients with the same questions again. The SHAI and PCS scores are usually the sum of the answers of the individual questions. We calculated the average per question as a way to scale the long- and short-form scores.

All questionnaires were administered in pen and paper format.

An a priori power analysis with a two-tailed paired t-test for our primary study question determined that 128 patients were needed to provide 80% power to detect a 0.25 SD difference between the short and long questionnaires of the PCS, with an alpha of 0.05.

The average score per question was calculated for each of the questionnaires. The scaled means of the short and long questionnaires were compared with paired t-tests.

To test content validity, we assessed floor and ceiling effects for all questionnaires. We calculated the percentage of patients having the lowest possible score (floor effect) for each questionnaire and the percentage of patients having the highest possible score (ceiling effect). When there is a high floor or ceiling effect, the questionnaire is less capable of detecting differences for the mildest and most severe ranges.

To investigate the construct convergent validity, Pearson's correlation coefficients were calculated for correlations between the short and long questionnaires and for correlations between these questionnaires and the following response variables: pain and disability (QuickDASH and DASH). Separate analyses were performed for the long and short versions of the questionnaires. A valid questionnaire should be able to detect low and high correlations (divergent and convergent validity). We considered a correlation low when r was between 0.10 to 0.29, moderate when r was between 0.30 and 0.49, and high when r was between 0.50 and 1.0 [5, 17].

In bivariate analysis of factors associated with pain (ordinal pain scale) and disability (DASH and QuickDASH), Pearson's test was used for continuous variables, Student's t-test for dichotomous variables, and one-way ANOVA for categorical variables. Factors with p values less than 0.10 in bivariate analysis were included in stepwise backward multivariable analyses (the factors that contribute to the best performing model are automatically selected in an iterative statistical process). When we entered categorical variables into the multivariable analysis, we dummy coded the separate categories (eg, nonspecific arm pain versus all other diagnoses). We then inserted all dummy categories to determine the most important categories. To determine whether the short questionnaire correlated with the same dependent variables as the long questionnaire (criterion validity), we compared the best models in the multivariable analysis for the dependent variables. If short questionnaires have adequate criterion validity, the predictors retained in the best stepwise multivariable model should be the same as for the original questionnaires.

Results

There were small differences in mean scaled scores for some of the shorter questionnaires: mean DASH and QuickDASH scores (29 ± 19 points versus 33 ± 21; $p < 0.001$); SHAI-18 and SHAI-5 scores (0.60 ± 0.38 versus 0.79 ± 0.48; $p < 0.001$); and PCS-13 and PCS-4 scores (0.41 ± 0.53 versus 0.50 ± 0.68; $p < 0.001$). There was no difference in the mean PHQ-9 and PHQ-2 scores (0.39 ± 0.49 versus 0.39 ± 0.68; $p = 0.97$).

There was a substantially higher floor effect for the PHQ-2 compared with the PHQ-9 (65% had the lowest possible score versus 29%), but all other floor and ceiling effects were comparable. The floor effects were 3% for QuickDASH versus 2.4% for DASH, 42% for PCS-4 versus 32% for PCS-13, and 4.9% for SHAI-5 versus 0% for SHAI-18. There was no ceiling effect for the QuickDASH, DASH, PHQ-9, SHAI-18, SHAI-5, and PCS. There were ceiling effects of 3% for the PHQ-2 and 0.6% for the PCS-4.

Convergent and divergent validity was shown for the long and the short questionnaires as follows. The long and short questionnaires correlated well with one other (all $p < 0.001$): DASH with QuickDASH ($r = 0.97$), PHQ-9 with PHQ-2 ($r = 0.84$), PCS-13 with PCS-4 ($r = 0.96$), and SHAI-18 with SHAI-5 ($r = 0.87$). All short and long questionnaires correlated significantly with the DASH (range, $p < 0.001$ to $p = 0.004$), QuickDASH (range, $p < 0.001$ to $p = 0.005$), and all questionnaires except for the PHQ-2 correlated with pain (range, $p < 0.001$ to $p = 0.014$) (Table 3). The SHAI-5 showed a moderate correlation with the PHQ-9 ($r = 0.43$; $p < 0.001$) and the correlation with pain intensity was low ($r = 0.20$; $p = 0.009$). The PCS-4 had a moderate correlation with the DASH ($r = -0.46$; $p < 0.001$) and a high correlation with the PHQ-9 ($r = 0.57$; $p < 0.001$).

Multivariable analyses identified comparable factors associated with the short and long forms of the questionnaires supporting criterion validity. Depression (as measured with the PHQ-2 or PHQ-9) and pain catastrophizing (PCS-13 or PCS-4) were associated with disability. The PCS (PCS-13 and PCS-4) was associated with pain intensity. The best multivariable model (the model that selected the variables, which together explained most of the variability in the outcome measure) for the DASH using the long questionnaires as explanatory variables included the PCS-13, PHQ-9, and age ($R^2 = 0.28$; $p < 0.001$) (Table 4). The best model using the short questionnaires included the PCS-4, PHQ-2, and age ($R^2 = 0.32$; $p < 0.001$). The best model for the QuickDASH using the long questionnaires included the PHQ-9, PCS-13, and differences in work status ($R^2 = 0.30$; $p < 0.001$). We then conducted another regression with the same predictors, but instead of using the long questionnaires, we inserted the shortened questionnaires. The best model using the short questionnaires included the PHQ-2, PCS-4, and differences in work status ($R^2 = 0.31$; $p < 0.001$). The best multivariable model for pain using the long questionnaires included the PCS-13, months since first pain experience, less education, no other pain conditions, and differences in diagnosis ($R^2 = 0.33$; $p < 0.001$) (Table 5). The best model for pain using the short questionnaires included the PCS-4, months since first pain experience, less education, no other pain conditions, and differences in diagnosis ($R^2 = 0.32$; $p < 0.001$).

Discussion

Previous studies established the value of the DASH, PCS, PHQ, and SHAI questionnaires in hand surgery [1, 19–21, 27, 28, 32]. Short forms exist for the DASH [2] and PHQ [12, 14], but not for the PCS and SHAI. Our study showed (1) small differences in scaled means of the short and long questionnaires; (2) comparable floor and ceiling effects of both versions of the questionnaires; (3) the short questionnaires had sufficient convergent and divergent validity (construct validity); and (4) the short and long questionnaires identified similar factors associated with disability and pain (criterion validity).

Table 3. Bivariate analysis of variables with the DASH, QuickDASH, and pain (n = 164)*

Variable	DASH		QuickDASH		Pain	
	R, T, or F value	p value	R, T, or F value	p value	R, T, or F value	p value
Pearson's correlation (r)						
PHQ-9	0.44	< 0.001	0.45	< 0.001	0.20	0.010
PHQ-2	0.47	< 0.001	0.48	< 0.001	0.12	0.12
PCS-13	0.50	< 0.001	0.51	< 0.001	0.46	< 0.001
PCS-4	0.46	< 0.001	0.47	< 0.001	0.46	< 0.001
SHAI-18	0.26	0.001	0.25	0.001	0.19	0.014
SHAI-5	0.23	0.004	0.22	0.005	0.21	0.007
Age	0.13	0.099	0.12	0.12	0.039	0.62
Years of education	−0.13	0.091	−0.16	0.043	−0.25	0.001
Months since first pain experience	0.18	0.024	0.17	0.035	0.22	0.005
T-test (T)						
Sex	1.6	0.10	1.2	0.23	1.5	0.14
Previous surgeries	0.51	0.61	−0.70	0.49	1.5	0.15
Other pain conditions	−2.3	0.021	−2.5	0.015	−3.7	< 0.001
Smoking	1.2	0.24	1.5	0.13	0.26	0.79
One-way ANOVA (F)						
Surgeon	1.5	0.21	1.7	0.16	1.3	0.27
Diagnosis	0.73	0.61	0.36	0.88	4.6	0.001
Marital status	0.56	0.69	0.43	0.79	1.9	0.12
Work status	4.5	0.001	4.6	0.001	1.9	0.099

* Bivariate analysis = Pearson's correlation, t-test, and one-way ANOVA were conducted for each of the individual predictors to find the association with the outcome measures: DASH, QuickDASH, and pain; PHQ-9 and -2 = nine- and two-item Patient Health Questionnaire; PCS-13 and -4 = 13- and four-item Pain Catastrophizing Scale; SHAI-18 and -5 = 18- and five-item Short Health Anxiety Index.

The study should be considered in light of its shortcomings. First, there were some missing answers on the questionnaires all substituted using mean imputation, but given the small number of missing questions for the large number of study subjects, this should have limited influence. Mean imputation is commonly used to address missing data [7]. This method is best for a small number of missing questions in a questionnaire. More sophisticated methods such as regression prediction are preferred when addressing complete missing data points such as entire questionnaires [8, 25]. We used the simpler mean imputation for missing items in a questionnaire because we had few missing data points and we used regression prediction plus error imputation for missing pain scores. Second, this study used the same cohort of patients as our previous unpublished study used in the creation of the shorter questionnaires. Ideally the creation and validation of questionnaires should be done in two different cohorts, as was done by our research group in a different study [31]. We did this for this study as well, but the data were lost by a visiting researcher. The other study by our group contained two cohorts, one of 65 patients to construct the questionnaire and 90 patients to validate [31]. For shortening of a questionnaire, however, an earlier study described using one cohort [16]. However, the cohort was large and sufficiently powered, and this study had a unique hypothesis and aim and comprised the same number of patients as the study of Vranceanu et al. [31]. However, it is not simply a matter of power to determine if there were enough patients for both study purposes; preferably a small study should have been used to create the questionnaire and a larger sample to validate. We now have only one large cohort with which we created and validated the questionnaires, which is a limitation of the current study. Third, a large cohort of patients was approached to participate but 28% of patients declined. Most patients declined based on the time investment necessary to complete the study. We saw this in a previous study [4], therefore this study is not different from most others.

The average QuickDASH score was slightly but significantly higher than the average complete DASH score, which is in agreement with a previous study [9]. In our opinion and in agreement with the literature [24], this small difference probably is not clinically important. The finding that the variability in DASH was predicted by pain catastrophizing and age is in agreement with the study of Novak et al. [18], and the correlation between disability and depression and pain catastrophizing also has been described [13, 15, 18, 30].

Table 4. Multivariable analysis of the short and long questionnaires and demographics on disability

Model*	Variables included in model	Adjusted $R^{2\dagger}$	p value	Part $R^{2\ddagger}$	p value
Including long questionnaires§					
DASH	Overall	0.28	< 0.001		
	PHQ-9			0.028	0.013
	PCS-13			0.078	< 0.001
	Age			0.018	0.043
QuickDASH	Overall	0.30	< 0.001		
	PHQ-9			0.019	0.038
	PCS-13			0.065	< 0.001
	Differences in work status				
	Retired versus full-time			0.015	0.066
	Workers Compensation versus full-time			0.013	0.077
	Unemployed versus full- time			0.018	0.043
Including short questionnaires§					
DASH	Overall	0.32	< 0.001		
	PHQ-2			0.092	< 0.001
	PCS-4			0.082	< 0.001
	Age			0.026	0.014
QuickDASH	Overall	0.31	< 0.001		
	PHQ-2			0.094	< 0.001
	PCS-4			0.076	< 0.001
	Differences in work status				
	Retired versus full-time			0.012	0.096

* The predictors found in the bivariate analysis were inserted in a backward, stepwise regression; this yields a model with the predictors that determine most of the variability of the variables; †percentage of the overall variability in the dependent variable explained or accounted for by the independent variables in the model; ‡part R^2 = the individual contribution of each variable to the adjusted R^2; §the regressions were conducted with all significant predictors from the bivariate analysis; when the questionnaires were entered, the first regression inserted the full questionnaires (eg, PHQ-9, PCS-13, SHAI-18); the second regression (including short questionnaires) inserted the same predictors, but instead of using the full questionnaires (eg, PHQ-9, PCS-13, SHAI-18), the PHQ-2, PCS-4 and SHAI-5 were inserted; the comparison of the models with the short and long questionnaires gives us an indication of the criterion validity; PHQ-9 and -2 = nine- and two-item Patient Health Questionnaire; PCS-13 and -4 = 13- and four-item Pain Catastrophizing Scale.

The percentage of floor and ceiling effects can help the clinician to determine whether the questionnaire is useful to discriminate the characteristics of interest in practice. The floor effects of the short (42%) and long (32%) versions of the PCS were different, but not far apart. The observed floor effect of the PHQ-2 (65%) was high and was higher than that (29%) for the long version. The PHQ-2 threshold score for major depression is 3. Sixty-five percent of the patients who did not meet the threshold had a score of 0 in our study and in a group of patients without depression in the studies of Kroenke et al. [12] and Lowe et al. [14]. The high floor effect of the PHQ-2 means that it is not useful for discriminating lower levels of depressive symptoms. The floor effect of the short form of the PCS is 10% greater than the long version. These floor effects seem acceptable for research because the correlations are still representative and acceptable for clinical care where only substantially maladaptive mindset merits a different treatment approach.

The PCS-4 and SHAI-5 showed convergent and divergent validity, establishing construct validity. The finding that pain did not correlate with depression in this study is interesting since pain is a good predictor of disability [3, 6], and disability correlates with depression [17] in patients with hand and arm conditions. We do not have an explanation for this finding and it might be spurious.

The multivariable models showed that DASH, QuickDASH, and pain were predicted by the same questionnaires in short and long formats. This indicates criterion validity. This is an argument for the use of the short questionnaires in orthopaedic practice. The questionnaires tested in this study were predictors of disability and pain, consistent with earlier studies [3, 17, 29, 32].

We found small but significant differences in the mean scores of the original questionnaires and abbreviated forms. Except for the PHQ, the floor and ceiling effects were comparable for all short and long questionnaires, and with respect to our third study question, we showed construct validity because short and long questionnaires had similar correlation with the outcome measures. Criterion validity was established by the observation that the short and long

Table 5. Multivariable analysis of the short and long questionnaires and demographics on pain intensity

Model*	Variables included in model	Adjusted $R^{2\dagger}$	p value	Part $R^{2\ddagger}$	p value
Including long questionnaires[§]					
Pain	Overall	0.33	< 0.001		
	PCS-13			0.14	< 0.001
	Months since first pain experience			0.019	0.033
	Years of education			0.025	0.014
	Differences in diagnosis				
	No other pain conditions			0.011	0.097
	Osteoarthritis compared with nonspecific arm pain			0.054	< 0.001
Including short questionnaires[§]					
Pain	Overall	0.32	< 0.001		
	PCS-4			0.13	< 0.001
	Months since first pain experience			0.018	0.037
	Years of education			0.023	0.019
	No other pain conditions			0.013	0.076
	Differences in diagnosis				
	Osteoarthritis compared with nonspecific arm pain			0.045	0.001

* The predictors found in the bivariate analysis were inserted in a backward, stepwise regression; this yields a model with the predictors that determine most of the variability of the variables; †percentage of the overall variability in the dependent variable explained or accounted for by the independent variables in the model; ‡part R^2 = the individual contribution of each variable to the adjusted R^2; §the regressions were conducted with all significant predictors from the bivariate analysis; when the questionnaires were entered, the first regression inserted the full questionnaires (eg, PHQ-9, PCS-13, SHAI-18); the second regression (including short questionnaires) inserted the same predictors, but instead of using the full questionnaires (eg, PHQ-9, PCS-13, SHAI-18), the PHQ-2, PCS-4 and SHAI-5 were inserted; the comparison of the models with the short and long questionnaires gives us an indication of the criterion validity; PHQ-9 and -2 = nine- and two-item Patient Health Questionnaire; PCS-13 and -4 = 13- and four-item Pain Catastrophizing Scale.

forms were included in the best multivariable models for various response variables and the models explained a similar percentage of the variance, although there were some differences in the variables included in the models. These shorter forms of questionnaires may be less burdensome for patients, researchers, and for clinicians who use them to track patient-centered outcomes. We propose using the shorter forms in an orthopaedic hand clinic.

References

1. Barsky AJ, Borus JF. Somatization and medicalization in the era of managed care. *JAMA.* 1995;274:1931–1934.
2. Beaton DE, Wright JG, Katz JN; Upper Extremity Collaborative Group. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am.* 2005;87:1038–1046.
3. Bot AG, Doornberg JN, Lindenhovius AL, Ring D, Goslings JC, van Dijk CN. Long-term outcomes of fractures of both bones of the forearm. *J Bone Joint Surg Am.* 2011;93:527–532.
4. Bot AG, Menendez ME, Neuhaus V, Mudgal CS, Ring D. The comparison of paper- and web-based questionnaires in patients with hand and upper extremity illness *Hand (NY).* 2013;8:210–214.
5. Cohen J, Cohen P, West S, Aiken L. *Applied Multiple Regression-Correlation Analysis for the Behavioral Sciences*. 3rd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 2003.
6. Doornberg JN, Ring D, Fabian LM, Malhotra L, Zurakowski D, Jupiter JB. Pain dominates measurements of elbow function and health status. *J Bone Joint Surg Am.* 2005;87:1725–1731.
7. Faris PD, Ghali WA, Brant R, Norris CM, Galbraith PD, Knudtson ML; APPROACH Investigators. Alberta Provincial Program for Outcome Assessment in Coronary Heart Disease. Multiple imputation versus data enhancement for dealing with missing data in observational health care outcome analyses. *J Clin Epidemiol.* 2002;55:184–191.
8. Graham JW. Missing data analysis: making it work in the real world. *Annu Rev Psychol.* 2009;60:549–576.
9. Gummesson C, Ward MM, Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. *BMC Musculoskelet Disord.* 2006;7:44.
10. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med.* 1996;29:602–608.
11. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16: 606–613.
12. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care.* 2003;41:1284–1292.
13. Lindenhovius A, Henket M, Gilligan BP, Lozano-Calderon S, Jupiter JB, Ring D. Injection of dexamethasone versus placebo for lateral elbow pain: a prospective, double-blind, randomized clinical trial. *J Hand Surg Am.* 2008;33:909–919.
14. Lowe B, Kroenke K, Grafe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J Psychosom Res.* 2005;58:163–171.
15. Lozano-Calderon SA, Souer JS, Jupiter JB, Ring D. Psychological differences between patients that elect operative or nonoperative

treatment for trapeziometacarpal joint arthrosis. *Hand (NY).* 2008;3:271–275.
16. McCracken LM, Dhingra L. A short version of the Pain Anxiety Symptoms Scale (PASS-20): preliminary development and validity. *Pain Res Manag.* 2002;7:45–50.
17. Niekel MC, Lindenhovius AL, Watson JB, Vranceanu AM, Ring D. Correlation of DASH and QuickDASH with measures of psychological distress. *J Hand Surg Am.* 2009;34:1499–1505.
18. Novak CB, Anastakis DJ, Beaton DE, Mackinnon SE, Katz J. Biomedical and psychosocial factors associated with disability after peripheral nerve injury. *J Bone Joint Surg Am.* 2011;93: 929–936.
19. Ring D, Guss D, Malhotra L, Jupiter JB. Idiopathic arm pain. *J Bone Joint Surg Am.* 2004;86:1387–1391.
20. Ring D, Kadzielski J, Fabian L, Zurakowski D, Malhotra LR, Jupiter JB. Self-reported upper extremity health status correlates with depression. *J Bone Joint Surg Am.* 2006;88:1983–1988.
21. Ring D, Kadzielski J, Malhotra L, Lee SG, Jupiter JB. Psychological factors associated with idiopathic arm pain. *J Bone Joint Surg Am.* 2005;87:374–380.
22. Salkovskis PM, Rimes KA, Warwick HM, Clark DM. The Health Anxiety Inventory: development and validation of scales for the measurement of health anxiety and hypochondriasis. *Psychol Med.* 2002;32:843–853.
23. Smith MV, Calfee RP, Baumgarten KM, Brophy RH, Wright RW. Upper extremity-specific measures of disability and outcomes in orthopaedic surgery. *J Bone Joint Surg Am.* 2012;94:277–285.
24. Sorensen AA, Howard D, Tan WH, Ketchersid J, Calfee RP. Minimal clinically important differences of 3 patient-rated outcomes instruments. *J Hand Surg Am.* 2013;38:641–649.
25. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ.* 2009;338:b2393.
26. Sullivan MJ, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. . *Psychological Assessment.* 1995;7: 525–532.
27. Turner JA, Aaron LA. Pain-related catastrophizing: what is it? *Clin J Pain.* 2001;17:65–71.
28. Turner JA, Jensen MP, Warms CA, Cardenas DD. Catastrophizing is associated with pain intensity, psychological distress, and pain-related disability among individuals with chronic pain after spinal cord injury. *Pain.* 2002;98:127–134.
29. Vranceanu AM, Jupiter JB, Mudgal CS, Ring D. Predictors of pain intensity and disability after minor hand surgery. *J Hand Surg Am.* 2010;35:956–960.
30. Vranceanu AM, Kadzielski J, Hwang R, Ring D. A patient-specific version of the Disabilities of the Arm, Shoulder, and Hand Questionnaire. *J Hand Surg Am.* 2010;35:824–826.
31. Vranceanu AM, Safren S, Cowan J, Ring D. The development of the negative pain thoughts questionnaire. *Pain Pract.* 2008;8:337–341.
32. Vranceanu AM, Safren S, Zhao M, Cowan J, Ring D. Disability and psychologic distress in patients with nonspecific and specific arm pain. *Clin Orthop Relat Res.* 2008;466:2820–2826.