Illness behavior in patients with musculoskeletal disease
Bot, A.G.J.

Citation for published version (APA):
Bot, A. G. J. (2013). Illness behavior in patients with musculoskeletal disease

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER 16

Factors associated with incomplete DASH questionnaires

Arjan G.J. Bot, Steven Ferree, Valentin Neuhaus and David C. Ring

HAND (NY). 2013 Jan;8(1):71-6
Abstract

Background: Missing data are unavoidable in clinical research. Older age, female gender and fewer years of education are risk factors for missing items in a questionnaire. This study assessed the differences between patients with complete and incomplete DASH questionnaires in terms of demographics and psychological factors.

Methods: We analyzed a convenience sample of 1204 patients enrolled in 8 prospective studies. The DASH and the Pain Catastrophizing Scale (PCS) were completed by all patients. The Center for Epidemiologic Studies-Depression scale (CES-D), Patient Health Questionnaire (PHQ-9), Pain Anxiety Symptoms Scale (PASS), and an ordinal pain scale were completed by 745, 493, 545 and 391 patients respectively. Bivariate analysis and binary logistic regression were used to determine risk factors for incomplete (one or more unanswered question) or invalid (more than 3 unanswered questions) DASH questionnaires.

Results: Thirty-one percent of patients did not complete the DASH. Patients with an incomplete DASH were older, had fewer years of education and had higher levels of catastrophic thinking, depression and pain anxiety. Age and catastrophic thinking were retained in the best logistic regression models of predictors of both incomplete and invalid DASH questionnaires.

Conclusions: The observation that patients who complete disability questionnaires are different from patients who do not may affect the interpretation of clinical research. Questionnaires based on item response theory merit investigation as a means to avoid incomplete questionnaires.
Introduction

Missing data are unavoidable in clinical research\textsuperscript{12,26}. For example, in a large international study by Gandek and colleagues, up to 13\% of the participants did not complete the SF-36\textsuperscript{11}. Missing answers can compromise the validity of questionnaires and may contribute to inaccuracy or bias\textsuperscript{7,12,25}. Methods to deal with missing data\textsuperscript{5,6} are helpful, but complete data is ideal.

Prior studies have identified the following predictors of missing items in a questionnaire: age\textsuperscript{2,16,21,22,28}, female gender\textsuperscript{16,21,22,28} and fewer years of education\textsuperscript{16,22}. The aim of this study was to investigate the differences between patients with an incomplete DASH and patients with a complete questionnaire. Our null-hypothesis was that there are no differences in demographics, catastrophic thinking and depression between patients with complete and incomplete DASH. Secondary study questions addressed the hypotheses that there would be no differences between patients with a valid DASH and patients with an invalid DASH (more than 3 unanswered questions) and between completers and non-completers of the most frequent missing questions.

Materials and methods

Study design

A convenience sample of 1209 prospectively enrolled patients was collected from a database with subjects from 8 different prospective studies performed at our institution from 2006 to 2011. All studies were done under an IRB approved protocol and all patients provided informed consent to participate in the original study. We included studies where the DASH questionnaire\textsuperscript{13} and the Pain Catastrophizing Scale (PCS) were administered\textsuperscript{20,27}, and either the Center for Epidemiologic Studies Depression Scale (CES-D)\textsuperscript{23} or the Patient Health Questionnaire (PHQ-9)\textsuperscript{15}. In a portion of the cohort also the Pain Anxiety Symptoms Scale-20 (PASS-20)\textsuperscript{18} was administered and a pain scale. We excluded 4 patients because we did not have access to the primary data sheets; and one additional patient due to a missing DASH questionnaire, leaving a cohort of 1204 patients in the study. We defined an incomplete DASH as one or more questions left blank. When not all questions were completed, we considered the DASH incomplete. In line with the developers of the DASH, we defined an invalid DASH questionnaire when more than 3 questions were missing\textsuperscript{13}.

Patient characteristics (Table 1)

The 1204 patients included in the cohort had a mean age of 53 years and 69\% of the patients completed the DASH. Other patient characteristics can be found in table 1.

Education information was assessed in 848 patients (70\% of the total cohort).

Evaluation

Upper extremity specific disability was measured with the 30-item DASH questionnaire\textsuperscript{13}. A DASH score is valid for a maximum of 3 missing questions\textsuperscript{13}.

The PCS was used to measure catastrophic thinking\textsuperscript{20,27}. This questionnaire contains 13 items, which are scored on a 4 point Likert scale\textsuperscript{20,27}. Patients are asked to score how often a certain feeling occurred in the last week. Scores per question range from 1-4 points, where 1 corresponds with “not at all” and 4 with “all the time”. The outcome scores range from 13 to 52, with a higher score indicating more catastrophic thinking\textsuperscript{20,27}. There are three subscales in the PCS: rumination, magnification and helplessness\textsuperscript{20,27}. Thirteen patients had an invalid or missing PCS score, so the group mean was imputed. There were 56 missing questions in 50 patients, we imputed the mean of the patients’ other questions for those missing values.
The CES-D was used to measure the level of depressive thoughts. The CES-D was obtained from 745 patients in 4 different studies. This questionnaire has 20 items each answered on a 4 point Likert scale. The scores range from 0 to 60, with a higher score indicating more depressive symptoms. Seven hundred and forty-five patients were enrolled in a study where the CES-D was administered. Twenty-nine patients did not complete the questionnaire and we imputed the group mean. Forty-seven patients had one missing question in the CES-D and we imputed the mean of the patients’ other answers for those missing questions.

A group of 493 patients enrolled in 5 different studies where the PHQ-9 was evaluated. The PHQ-9 is a questionnaire that measures symptoms of depression. It consists of 9 items scored on a 4 point scale from 0 “not at all” to 3 “nearly every day.” The scores range from 0 to 27 and higher scores indicate more depressive symptoms. One patient did not complete this questionnaire and we imputed the group mean. There were 17 missing answers in 15 patients; we imputed the mean of the patients’ other questions for those missing questions.

A group of 545 patients in 3 different studies completed the PASS questionnaire in addition to a depression questionnaire. The PASS-20 is a 20 item questionnaire that assesses pain anxiety. The questionnaire has 4 subscales; cognitive anxiety, fear of pain, escape and avoidance and physiological anxiety. Patients are asked how often they are engaged in certain activities or thoughts, and the items are scored on a 6 point scale from 0 “never” to 5 “always.” Total scores of the PASS-20 range from 0-100, and scores of 0-25 on each of the subscales are possible. Higher scores indicate a higher level of pain anxiety. Thirty-two questions were missing in 30 patients, we inserted the mean of the patients’ other questions.

Pain was measured on an 11-point ordinal scale in 391 patients in 4 different studies. We imputed the mean pain score in 15 patients.
Statistical Evaluation

A power analysis indicated that 831 patients in the group with complete DASH and 373 with missing DASH questions provided 94% power to detect an effect-size of 0.22 in differences in the PCS, with a two tailed T-test.

We used a Pearson chi-square test to assess differences in the dichotomous and categorical variables (gender and diagnosis) between patients with a complete and incomplete DASH (yes/no). Independent samples T-test were executed to detect differences in age, education and the different questionnaires.

Pearson chi-square test and Fisher’s exact test were used to assess categorical variables. Unequal variance T-tests were used to detect differences in age, education and the different questionnaires. The most frequent missing question in the DASH was assessed and an unequal variance independent samples T-test was conducted to assess the differences in the PCS score and age between the completers and non-completers of the most frequent missing question. A Pearson chi-square test was used to assess differences in gender between completers and non-completers of the most frequent missing question in the DASH.

Due to adequate sample size and to account for possible confounding all variables that were available in all 1204 patients and had a p-value of less than 0.10 in the bivariate analysis were entered in a backwards, stepwise (likelihood) binary logistic regression analysis to assess the correlation of these variables with the outcome variable DASH complete (yes/no). We dummy coded the categorical variable diagnosis with “other diagnosis” as the reference category and investigated the interaction terms of age with all separate diagnoses.

To determine the predictors of an invalid DASH score, we inserted all variables with p<0.10 in bivariate analysis in a backwards, stepwise (likelihood) binary logistic regression analysis.

Results

Evaluation scores

Patients with an incomplete DASH were older, had fewer years of education, and had higher PCS, CES-D and PASS scores (Table 2).

Patients with an invalid DASH (more than 3 missing questions) were older, had fewer years of education and had higher PCS and CES-D scores (Table 2).

The most frequent missing question of the DASH was question 21 regarding difficulties with sexual activities. One hundred fifty patients (12%) did not complete this question. The non-completers of question 21 were significantly older than the completers; had a significantly higher PCS score; and were more often female. (Table 3)

The five most frequent not-completed questions were: question 21 (“Sexual activities”): 150 times (12% of the cohort); question 8 (“Garden or do yard work”): 54 times (4.5%); question 19 (“Recreational activities in which you take some force or impact through your arm, shoulder or hand (golf, hammering, tennis, etc.): 49 times (4.1%), question 27 (“Weakness in your arm, shoulder or hand”): 48 times (4.0%) and question 5 (“Push open a heavy door”: 40 times (3.3%).
Table 2 Outcome differences in patients with complete/incomplete and valid/invalid †DASH

<table>
<thead>
<tr>
<th></th>
<th>Total cohort</th>
<th>†DASH complete</th>
<th>†DASH incomplete</th>
<th>p</th>
<th>t</th>
<th>Valid †DASH</th>
<th>Invalid †DASH</th>
<th>p</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1204</td>
<td>831</td>
<td>373</td>
<td>1152</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>6.8</td>
<td>6.2</td>
<td>8.2</td>
<td>6.7</td>
<td>9.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>8.4</td>
<td>8.0</td>
<td>9.1</td>
<td>8.2</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>13-52</td>
<td>0-39</td>
<td>0-39</td>
<td>&lt;0.001</td>
<td>3.6</td>
<td>0-39</td>
<td>0-37</td>
<td>0.049</td>
<td>2.0</td>
</tr>
</tbody>
</table>

‡PCS: Pain Catastrophizing Scale

•CES-D: Center for Epidemiological Studies Depression Scale

♦PHQ-9: Patient Health Scale-9

°PASS: Pain Anxiety Symptoms Scale

The best logistic regression model for predictors of an incomplete DASH questionnaire included age ($\beta=0.035$, $p<0.001$, OR 1.04, CI 1.03-1.05) and PCS ($\beta=0.034$, $p<0.001$, OR 1.04 CI 1.02-1.05, constant $\beta_0 = -3.5$) (Model fit; Hosmer and Lemeshow test: Chi-square=8.4, $p=0.40$, Nagelkerke R-square=0.10). (Table 4). These findings mean that with each year of increased age or 1 point increase of the PCS, the odds to have missing questions in the DASH increases with 4% (more or less comparable to a relative risk of 1.04 or 4% more chance to be a non-completer).

The best logistic regression model for predictors of an invalid DASH questionnaire included age ($\beta=0.036$, $p<0.001$, OR 1.04, CI 1.02-1.06) and PCS ($\beta=0.038$, $p=0.007$, OR 1.04 CI 1.01-1.07, constant $\beta_0 = -5.9$). (Model fit; Hosmer and Lemeshow test: Chi-square=4.6, $p=0.80$, Nagelkerke R-square=0.055). (Table 5). So the odds to be a non-completer increase with 4% per year or point on the PCS.

**Discussion**

Both age and catastrophic thinking were associated with significantly higher probability to have an incomplete or invalid DASH questionnaire.

Our findings that patients with incomplete DASH score were older and had less years of education is in agreement with findings in literature\(^2,16,21,22,28\).

The finding that the majority of patients only missed a few questions (e.g. only 52 patients had more than 3 missing questions on the DASH) is in line with the findings of Liu and colleagues\(^16\). The authors also found a similar percentage of missing questions\(^16\). The percentage of missing questions per study and questionnaire vary. For instance Burns and colleagues found 0.5-95% of items missing on the MMSE\(^2\). Other studies found 13% and 41%
missing items on the SF-36 \(^{11,19}\), 11\% missing items on the Expanded Prostate Cancer Index Composite-26 \(^4\), 24\% on the MSKCC Bowel Function instrument \(^{29}\) and 10\% on the SF-12 \(^{21}\).

Patients with incomplete questionnaires had higher scores on measures of depression and pain anxiety, but did not have higher levels of pain.

The question “difficulties with sexual activities” was the most frequent missing question in the DASH and patients who did not complete this question were older, female and had more catastrophic thinking. The finding that older female patients are at risk for not completing the question about sexual activities was also reported in the study by Fayers and colleagues \(^{7}\). The authors reported that 19\% of patients left this question blank in the Rotterdam Symptom Checklist \(^7\). In our study, 150 patients (12\%) did not complete this question, which was lower than the 29\% Beaton and colleagues found in their study \(^1\).

The majority of the most frequent unanswered questions are things that not everyone participates in: e.g. sex, gardening, and forceful recreational activities. This reinforces one of the advantages of outcome measures based on item response theory (IRT). This approach uses Computerized Adaptive Tests (CAT), with which the next question depends on the answer of the previous question \(^3,24\). The questions are ranked so that once a patient reports difficulty with a specific task, difficulty with all more difficult tasks is assumed. IRT based questionnaires can arrive at a patient’s disability level with fewer questions \(^9,10,14,24\). Questions that are not relevant to the patient can be bypassed \(^14\).

### Table 3: Most frequent missing question analysis of the †DASH

<table>
<thead>
<tr>
<th>Question 21</th>
<th>Completers of question 21</th>
<th>Non completers of question 21</th>
<th>p</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Mean ± SD</td>
<td>Range</td>
<td>Mean ± SD</td>
<td>Range</td>
</tr>
<tr>
<td>‡PCS</td>
<td>6.5 ± 8.1</td>
<td>13-52</td>
<td>9.0 ± 9.6</td>
<td>13-50</td>
</tr>
<tr>
<td>Sex</td>
<td>Males (%)</td>
<td>44</td>
<td>31</td>
<td>0.003</td>
</tr>
</tbody>
</table>

†DASH: Disabilities of the Arm, Shoulder and Hand
§Question 21: Sexual activities
‡PCS: Pain Catastrophizing Scale

### Table 4: Logistic regression for predictors of an incomplete †DASH questionnaire

<table>
<thead>
<tr>
<th>Predictors of an incomplete †DASH total cohort</th>
<th>β</th>
<th>p</th>
<th>Odds ratio</th>
<th>95% CI for Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.035</td>
<td>&lt;0.001</td>
<td>1.04</td>
<td>1.03 - 1.05</td>
</tr>
<tr>
<td>‡PCS</td>
<td>0.034</td>
<td>&lt;0.001</td>
<td>1.04</td>
<td>1.02 - 1.05</td>
</tr>
<tr>
<td>Age*fracture</td>
<td>0.005</td>
<td>0.076</td>
<td>1.01</td>
<td>1.00 - 1.01</td>
</tr>
<tr>
<td>Constant</td>
<td>β₀ = -3.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

†DASH: Disabilities of the Arm, Shoulder and Hand
‡PCS = Pain Catastrophizing Scale
N= number of patients in the cohort
Variables included in the regression: age, sex, PCS, diagnoses (with dummy coding), age by all diagnoses
Table 5: Logistic regression for predictors of an invalid †DASH questionnaire

<table>
<thead>
<tr>
<th>Predictors of a invalid †DASH total cohort</th>
<th>n=1204</th>
<th>β</th>
<th>p</th>
<th>Odds ratio</th>
<th>95% CI for Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.036</td>
<td>&lt;0.001</td>
<td>1.04</td>
<td>1.02</td>
<td>1.06</td>
</tr>
<tr>
<td>‡PCS</td>
<td>0.038</td>
<td>0.007</td>
<td>1.04</td>
<td>1.01</td>
<td>1.07</td>
</tr>
<tr>
<td>Constant</td>
<td>β₀=-5.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

†DASH: Disabilities of the Arm, Shoulder and Hand
‡PCS = Pain Catastrophizing Scale
β= number of patients in the cohort
Variables included in the regression: age, sex, PCS, diagnoses (with dummy coding), age by all diagnoses

Using IRT for the QuickDASH was described by Lyren and Atroshi, where a minimum number of 6 questions was required; in contrast to a minimum of 10 questions in a traditional administration of the questionnaire ¹⁷. This technique has also been used for the PHQ-9 ⁸ and the CTS-6 questionnaire ¹⁷.

This study should be viewed in the light of several shortcomings. Some comparisons might be underpowered. For instance only 52 patients had invalid questionnaires. Our conclusions concerning education, depression (based on CES-D and PHQ-9), pain anxiety and pain are based on subpopulations of 862, 745, 493, 545 and 391 patients respectively, so are less strong than the conclusions for age and catastrophic thinking. That is also the reason why we did not include those variables in the logistic regression model.

In conclusion, this study found that catastrophic thinking and age were risk factors for incomplete or invalid DASH scores. The finding that patient characteristics affect how questionnaires are completed is important for future research. The fact that 31% of the patients did not complete the DASH score may be a reason to consider moving to measures based on Item Response Theory, where the same information is obtained with fewer questions, subjects can bypass questions that are not relevant to them, and incomplete questionnaires are not possible.
References


