

A comparison of patients' responses about their disability with and without attribution to their affected area[☆]

Robert G. Marx^{a,b}, Sheilah Hogg-Johnson^b, Pam Hudak^b, Dorcas Beaton^b, Susanne Shields^b,
Claire Bombardier^{b,c}, James G. Wright^{a,*}

^aDepartment of Surgery and Public Health Sciences, Clinical Epidemiology and Health Care Research Program, University of Toronto,
The Hospital for Sick Children, 555 University Avenue, Toronto, ON, M5G 1X8 Canada

^bInstitute for Work and Health, 250 Bloor Street E, 7th Floor, Suite 702, Toronto, ON M4W 1E6 Canada

^cDepartment of Medicine, Health Administration, University of Toronto, Institute for Work and Health,
250 Bloor Street E, 7th Floor, Suite 702, Toronto, ON M4W 1E6 Canada

Received 15 July 1998, received in revised form 2 October 2000, accepted 4 October 2000

Abstract

The purpose of this study was to determine whether individual items in a disability questionnaire were answered differently depending on whether or not the questions were attributed to the upper limb (i.e., “do you have difficulty eating due to your arm or hand problem?” or “do you have difficulty eating?”, respectively). The a priori hypothesis was that the same or more disability would be detected by non-attributed items. Four hundred sixty-seven patients with upper extremity disorders completed the SF-36 general health survey, which does not attribute health problems to affected areas. Patients also completed six additional questions, modified from the SF-36, regarding work (four questions) and social function with friends and family (two questions), which attributed their disability to their affected upper extremity. Of 467 patients, 419–431 (89–92%) responded to both versions of the questions. Although we demonstrated a significant order effect (Generalized Estimating Equation; $P=.003$), comparison of the responses to the six questions showed that for five of the six questions (Generalized Estimating Equation; $P\leq.001$) patients reported more disability when the questions were worded with attribution to the upper extremity. Even considering the order effect, patients demonstrated a counterintuitive result by reporting more disability when questions were attributed to their affected area. Thus, both the wording of questions and order of questions can significantly affect patients' responses about their disability and raises questions about the validity of patients' reports of their disability. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Disability evaluation; Questionnaires; Health status; Reproducibility of results

1. Introduction

Questionnaires used to measure health may be generic or specific [1,2]. *Generic* instruments comprehensively assess patients' overall health and can therefore be used to compare the status of patients with different diseases [1,3]. *Specific* instruments pertain either to pathologic entities, called disease- or condition-specific scales, or anatomic regions, called region-specific scales [1,3,4]. Generic instruments have the advantage of allowing comparisons among patients with different conditions. Disease-specific scales have the

advantage of focusing on those aspects of the condition that are most relevant to patients and clinicians [2,3,5]. Specific instruments usually provide different assessment of “health” when compared with generic measures [1]. This difference has generally been assumed to be due to the content of disease-specific questions and their focus on disease-specific problems, which are not reflected in the more global generic measure [1,3,6,7]. An alternative explanation, which has not been previously considered, is that the difference may be explained by the wording or “framing” of the questions.

Specific measures of health-related quality of life, in addition to focusing on the aspects of the condition that are important to patients (and clinicians), usually explicitly or implicitly attribute the complaints to the diseased area [8]. For example, patients are asked how much difficulty they have dressing themselves *due to their shoulder problem*.

[☆] Presented at the Canadian Orthopaedic Association Meeting, Hamilton, Ontario, 1997.

* Corresponding author. The Hospital for Sick Children, 555 University Avenue, S-107, Toronto, ON M5G 1X8 Canada. Tel.: (416) 813-6433; fax: (416) 813-6414.

(Items in generic measures can also be attributed to their diseased area [9]; e.g., how is your life in general due to your shoulder problem?) For patients with only one problem or affected area, patients would be expected to express the same level of disability whether or not the questions were attributed to their disease or affected area. For those patients with multiple or comorbid diseases, one might expect less disability to be detected by items that are specifically attributed to the diseased area, as this would reflect only a portion of their overall disability. For example, an item asking about disability due to a shoulder problem should not pick up a problem at work *due to knee arthritis* (or other medical or physical disorders). Thus, it might be expected that patients without comorbid disease would report the same degree of health problems whether or not questions were attributed to specific regions or diseases. Furthermore, we would also anticipate the response to the two ways of asking questions (attributed and nonattributed) to be different in patients with coexistent physical or medical problems (comorbidity).

The decision to use attributed and nonattributed items in a disability questionnaire has several important implications. First, the wording of items (with or without attribution) may affect the focus of the scale (overall disability or disability due to the affected area) and, thereby, the phenomenon being measured. Second, different wording may affect the performance of items and consequently the decision to eliminate or select items in the item reduction phase of instrument development. Finally, the wording of the items, as previously discussed, may affect the evaluation of disability in the final scale (nonattributed items are more likely to be affected by comorbid disease). The purpose of this study was: (i) to determine whether individual items were answered differently by patients depending on whether or not they were attributed to their upper limb problem, and (ii) to determine how the presence of comorbidity affected patients' responses.

2. Methods

The American Academy of Orthopaedic Surgeons and the Institute for Work and Health (a nonprofit research organization that attempts to improve the understanding of factors that determine workplace health, injury, and disability) collaborated to develop a measure of upper extremity symptoms and disabilities, called the DASH (Disabilities of the Arm, Shoulder, and Hand). The DASH is an evaluative and discriminative, region-specific instrument [10].

Items for the DASH were obtained from previous scales, clinicians, and focus groups [10]. After eliminating redundant and irrelevant items, the scale was reduced based on clinicians' ratings to 70 items, prior to final item reduction to a 30-item scale. Of the 70 items considered for inclusion in the final version of the DASH, four relating to work and two pertaining to social function were modified questions from the SF-36 [11]. These six questions were modified

such that patients' disabilities were attributed to their upper extremity condition (these were referred to as "attributed" items). Patients also completed the SF-36 [11], which included the same six (four work and two social function) questions that were not attributed to their upper limb. These were referred to as "nonattributed" items. The wording of the questions is provided in the Appendix. The two different formats (attributed and nonattributed) of these six questions were the focus of this investigation.

2.1. Part I

Patients were recruited from 20 centers across Canada, Australia, and the United States. All patients referred for upper limb complaints were eligible. Diagnosis, gender, hand dominance, employment status, age, worker's compensation status, and side affected were documented. As previously discussed, patients completed the 70-item DASH questionnaire (which included the six attributed questions) and the SF-36 (which included the six nonattributed questions). All attributed questions were asked ahead of the nonattributed questions, as the DASH questionnaires was ahead of the SF-36 in the package distributed to the patients.

The two versions (attributed and nonattributed) of the two social activity items had identical response options (see Appendix). The nonattributed versions of the four work-related questions had binary response options, as in the original SF-36. The attributed versions of the four work-related questions had five response options (see Appendix). In order to compare the four attributed and nonattributed work-related questions directly, the five category ordinal responses were reduced to binary responses. The "not at all" responses were converted to "no" and the remainder of the responses ("slightly" to "extremely") were converted to "yes." Responses 2 through 5 were grouped together because they express the presence of disability.

Comorbidity was assessed using a self-report scale (Jeff Katz, personal communication) that asked patients if they had heart disease, high blood pressure, lung disease, diabetes mellitus, ulcer or stomach disease, kidney disease, liver disease, anemia or other blood disease, cancer, depression, osteoarthritis or degenerative arthritis, back pain, or rheumatoid arthritis. Patients were asked if they "currently had the problem and if they were receiving treatment for it." We hypothesized that patients without comorbidity would express similar disability (with or without attribution) and patients with comorbidity would express greater disability on the generic (nonattributed) items compared with the specific (attributed) items because their responses should be based on their associated problems. Thus, the analyses were performed comparing the responses of patients both with and without comorbidity.

The statistical significance of the disagreement was tested with McNemar's chi-square [12]. This statistical test is used to test the direction of disagreement for matched pairs. This test is generally used for 2-by-2 tables; however, we also used it for the 5-by-5 tables by comparing the num-

ber of individuals for each direction of disagreement. The results were also compared for patients with and without comorbidity.

Part I of this study had two methodologic limitations: (i) the attributed questions were always asked ahead of the nonattributed questions, and (ii) the number of response options for four of the six questions were not the same. In Part II of the study, the attributed and nonattributed versions were given in random order to 63 patients. Furthermore, the four work questions for both the attributed and nonattributed versions were given binary response options so that both versions for all six questions had the same number of response options. Patients were consecutively assessed at a single Canadian centre that had participated in Part I. The sample size was based on a delta of 0.15 (the delta was based on the smallest difference observed in Part I) for the difference in the off-diagonal percentages with an alpha = .05 and beta = .90 [13].

3. Results

3.1. Part I

Of the 407 patients who were enrolled in the study, 88–92% of the respondents answered both the attributed and nonattributed versions of each of the six items. Patients' sex, hand dominance, employment status, age, worker's compensation status, and side affected are listed in Table 1.

Table 1

Characteristics	Distribution (n=407)
Sex	
Male	166 (41%)
Female	206 (51%)
Missing	35 (8%)
Dominant hand	
Right	365 (90%)
Left	37 (9%)
Missing	5 (1%)
Employed	
Yes	198 (49%)
No	161 (40%)
Missing	48 (11%)
Why not working?	
Upper limb problem	59 (15%)
Other health problem	6 (2%)
Unemployed	10 (2%)
Retired	49 (12%)
Other	21 (5%)
Missing	16 (3%)
Average age in years	45.0 (S.D. = 16.7)
Worker's compensation	
yes	46 (11%)
no	361 (89%)
Side affected	
Right	175 (43%)
Left	125 (31%)
Both	103 (25%)
Missing	4 (1%)

Table 2

Attributed and nonattributed responses to item 1: "Extent social activities"

	NA1	NA2	NA3	NA4	NA5	Total
A1	101	17	5	0	0	123
A2	49	35	10	1	0	95
A3	10	19	21	9	0	59
A4	7	12	16	13	0	48
A5	5	2	8	12	7	34
Total	172	85	60	35	7	359

A: attributed response; NA: nonattributed response.

McNemar's Test Statistic: $\chi^2 = 57.2$, $P \leq .001$.

Percent agreement = 49.3%.

The diagnoses of the patients included rotator cuff tendinopathy (18%), carpal tunnel syndrome (9%), osteoarthritis (7%), rheumatoid arthritis (6%), lateral epicondylitis (5%), thumb carpo-metacarpal arthritis (4%), Colle's fracture (2%), DeQuervain's tenosynovitis (1%), humerus fracture (1%), nonspecific (6%), unknown (10%) and other (32%). Of the 407 patients, 155 had back pain (39%), 85 osteoarthritis (21%), 72 hypertension (18%), 47 rheumatoid arthritis (12%), 40 ulcer disease (10%), 27 reported heart disease (7%), 16 anemia or other blood diseases (4%), 15 depression (4%), 13 diabetes mellitus (3%), 7 cancer (2%), and 6 renal disease (2%).

Tables 2 and 3 demonstrate patients' responses to the two formats (attributed and nonattributed) of the social function items (both administered with 5-point response scales). Tables 4 through 7 demonstrate the patients' responses to the four work-related items. For five of the six questions, patients were more likely to report disability on the attributed items compared with the nonattributed items (McNemar's chi-square, $P \leq .001$).

The analyses were performed separately for patients with and without comorbidity. The percentage agreement for the items, for patients without comorbidity, was 48, 65, 75, 80, 73, and 74 for items 1 through 6, respectively. The percentage agreement for the items, for patients with comorbidity was 49, 56, 75, 82, 75 and 83 for items 1 through 6, respectively, showing similar levels of agreement with the "no comorbidity" group. The percentage of patients (without comorbidity) who reported being disabled on the attributed version of the questions (vs. the nonattributed version) was

Table 3

Attributed and nonattributed responses to item 2 "Frequency social activities"

	NA1	NA2	NA3	NA4	NA5	Total
A1	150	23	13	1	1	188
A2	21	30	12	6	2	71
A3	8	18	31	4	0	61
A4	1	4	14	7	5	31
A5	3	2	1	2	6	14
Total	183	77	71	20	14	365

A: attributed response; NA: nonattributed response.

McNemar's Test Statistic: $\chi^2 = 0.26$, $P = .36$.

Percent agreement = 61.4%.

Table 4
Attributed and nonattributed responses to item 3: “Cut down on work or other activities”

	No (NA)	Yes (NA)	Total (NA)
No (A)	106	16	122
Yes (A)	76	171	247
Total (A)	182	187	369

A: attributed response; NA: nonattributed response.
McNemar’s Test Statistic: $\chi^2 = 37.8, P \leq .001$.
Percent agreement = 75.1%.

40 vs. 13, 18 vs. 17, 19 vs. 3, 19 vs. 0, 26 vs. 1, and 25 vs. 1 for items 1 through 6, respectively. The percentage of patients (with comorbidity) who reported being disabled on the attributed version of the questions (vs. the nonattributed version) was 41 vs. 10, 23 vs. 21, 22 vs. 6, 15 vs. 3, 23 vs. 2, and 16 vs. 1 for items 1 through 6, respectively.

3.2. Part II

Of 68 patients in Part II of the study who were approached to participate, 63 (93%) completed the questions. The characteristics of the patients in Part II (listed in Table 1) were not significantly different from the patients who participated in Part I.

Table 8 shows the summarized responses to the six questions for the 63 patients stratified by whether the attributed or nonattributed questions (determined by random assignment) were asked first or second. For all six questions, no matter what the order, more patients rated themselves as disabled on the attributed compared with the nonattributed questions ($P \leq .009$). The order the patients completed the questions did affect the magnitude of the discordance ($P = .003$; Generalized Estimating Equation model with a logit link). The discordance was greater when the attributed questions were asked first.

Because of the order effect, we reanalyzed our results as a parallel groups design comparing the percentage of patients who rated themselves as disabled when responding to the attributed question (when asked first) compared with the percentage who rated themselves as disabled when responding to the nonattributed question (when asked first). For all six questions more patients counterintuitively rated themselves as disabled when responding to the attributed ques-

Table 5
Attributed and nonattributed responses to item 4: “Accomplish less than you would like”

	No (NA)	Yes (NA)	Total (NA)
No (A)	63	5	68
Yes (A)	66	239	305
Total (A)	129	244	373

A: attributed response; NA: nonattributed response.
McNemar’s Test Statistic: $\chi^2 = 50.7, P \leq .001$.
Percent agreement = 81.0%.

Table 6
Attributed and nonattributed responses to item 5: “Limited in the kind of work or other activities”

	No (NA)	Yes (NA)	Total (NA)
No (A)	52	5	57
Yes (A)	91	223	314
Total (A)	143	228	371

A: attributed response; NA: nonattributed response.
McNemar’s Test Statistic: $\chi^2 = 75.3, P \leq .001$.
Percent agreement = 74.1%.

tion compared with nonattributed question (0.73 vs. 0.56, $P = .16$; 0.80 vs. 0.69, $P = .31$; 0.83 vs. 0.82, $P = .87$; 0.86 vs. 0.75, $P = .27$; 0.80 vs. 0.61, $P = .09$; 0.77 vs. 0.56, $P = .09$, respectively). Although the parallel groups analyses were underpowered, the Generalized Estimating Equation model, while adjusting for the order effect, confirmed that disability was more often reported on the attributed compared with nonattributed version ($P \leq .001$).

4. Discussion

Attribution of patients’ disability to their area of complaint was expected to detect less disability than questions that did not attribute the disability to any cause. Patients, however, were more likely to report disability on the attributed items than the nonattributed items for one of the two social disability items and for all four work-related items. Only when patients were asked about the frequency of their disability for social activity did they report similar amounts of disability (attributed or not). However, we did detect a significant order effect. When patients were asked the nonattributed questions first, the difference between the two versions of questions was less. If patients had responded to questions in a “logical” way, then the same number of or more patients should rate themselves as disabled when responding to nonattributed questions compared with attributed questions. For both orders (order = 0 and order = 1, as shown in Table 8) more patients rated themselves as “disabled” on the attributed questions compared with nonattributed questions. Although the differences between the two versions was less when the nonattributed questions appeared first (order = 1), even a finding of similar disability

Table 7
Attributed and nonattributed responses to item 6: “Difficulty performing the work or other activities”

	No (NA)	Yes (NA)	Total (NA)
No (A)	48	3	51
Yes (A)	76	230	306
Total (A)	124	233	357

A: attributed response; NA: nonattributed response.
McNemar’s Test Statistic: $\chi^2 = 65.6, P \leq .001$.
Percent agreement = 77.9%.

Table 8
Comparison of attributed and nonattributed questions when given in random order

Question	Number of response categories	Order = 0: Attributed questions appeared first on questionnaire			Order = 1: Nonattributed questions appeared first on questionnaire		
		Agree	More disabled on attributed questions	More disabled on nonattributed questions	Agree	More disabled on attributed questions	More disabled on nonattributed questions
1. To what extent... interfered with your normal social activities...	5	9/30 (30%)	19/30 [19/21] (63.3%) [90.5%]	2/30 (6.7%)	16/33 (48.5%)	10/33 [10/17] (30.3%) [58.8%]	7/33 (21.2%)
2. ...how much of the time has your...interfered with your normal social activities	5	12/30 (40.0%)	16/30 [16/18] (53.3%) [88.9%]	2/30 (6.7%)	14/32 (43.8%)	12/32 [12/18] (37.5%) [66.7%]	6/32 (18.8%)
3. Cut down on the amount of time you spent on work or other activities	2	25/30 (83.3%)	5/30 [5/5] (16.7%) [100%]	0/30 (0%)	25/32 (78.1%)	6/32 [6/7] (18.8%) [85.7%]	1/32 (3.1%)
4. Accomplished less than you would like	2	23/30 (76.7%)	7/30 [7/7] (23.3%) [100%]	0/30 (0%)	28/32 (86.5%)	3/32 [3/4] (9.4%) [75%]	1/32 (3.1%)
5. Were limited in the kind of work or other activities	2	21/30 (70.0%)	9/30 [9/9] (30.0%) [100%]	0/30 (0%)	28/33 (84.8%)	2/33 [2/5] (6.1%) [40%]	3/33 (9.1%)
6. Had difficulty performing the work or other activities	2	22/29 (75.9%)	7/29 [7/7] (24.1%) [100%]	0/29 (0%)	31/32 (96.9%)	1/32 [1/1] (3.1%) [100%]	0/32 (0.0%)

Parentheses show percentages with all responses in the denominator. Brackets use just the discordant pairs as the denominator (i.e., all cases where the same response was given on the generic and specific versions of the question are removed from consideration).

for the two versions would have been surprising. Furthermore, the Generalized Estimating Equation, while adjusting for the statistically significant order effect, confirmed that patients were statistically more likely to report disability on the attributed questions. Thus, despite the order effect, not only did this study fail to confirm the hypothesis that attributed items would detect the same or less disability, but the results were counterintuitive because more patients rated themselves disabled on the attributed questions than on questions asked in a nonattributed fashion. This counterintuitive finding questions the validity of patients' self-reports of their disability and condition.

Previous studies have shown that patients interpret risks and make decisions differently based on how risks are presented [14]. Presentation of the risks associated with different choices has been called "framing" and has been found to impact on both medical and nonmedical decision-making [15]. For example, when patients are presented with treatment options associated with explicit risks of dying, rather than living, the treatment is viewed much less favourably [16]. In this study we also found a significant order effect whereby the difference between the two versions was greatest when the attributed questions were asked first. Thus, framing and order both appear to have important effects on how patients respond to questionnaires.

In view of the paradoxical results of this study, patients may be overstating their disability when responding to attributed items or understating their disability when responding to nonattributed items. Patients with upper extremity disorders may understate their disability when asked non-specific (nonattributed) questions as opposed to specific (attributed) questions for several reasons: (i) they may not think to comment on how conditions other than their upper

extremity problem impair their function, (ii) they may discount their disability due to other conditions because they are focusing on their upper extremity problem (the reason for seeing the doctor), and (iii) they may see health as separate or different from their upper extremity problem. A second explanation is that when people with upper limb pathology are asked if they have trouble due to their arm problem, they may overstate their disability because they are focusing on the reasons that prevents them from working or the reasons that cause them to visit their doctor. Finally, the discrepancy between patients, responses to the attributed and nonattributed items may be due to a combination of these two factors.

Patients without comorbidity had a greater discrepancy between their responses to the attributed and nonattributed versions for most questions. This finding was anticipated because patients with other health problems have more reasons to have disability with their work and social activities in addition to their upper extremity disorder. However, overall, more patients still reported disability on the attributed versions for five of the six questions. Only one prior study has investigated the effect of attribution (of questions) on the response to health questionnaires. This work, however, compared the scores of generic and condition-specific instruments, but not the responses to the items themselves [8].

This research, in view of the counter intuitive results and order effects, raises the question of which method of asking patients about their disability provides accurate results. Previous research on framing of patient decisions cannot fully address this issue because in decision-making there is no "right" answer. Because people seem to respond with more "appropriate" decisions when provided more explicit information [17], perhaps the attributed versions may be provid-

ing more accurate disability assessments. However, further studies need to compare patients' ratings of their disability with an external standard, such as objective or observational measures of function. This research may also be relevant to the observation that disease-specific measures are more responsive than generic measures [1]. The increased disability detected by attributed questions (typical of many disease-specific questionnaires) may partially account for the improved responsiveness of disease-specific measures compared with generic measures.

Our results demonstrate that the wording and order of the questions may significantly affect the performance of the items. Prior to testing items, the investigator must word the questions to most clearly reflect the concept they are attempting to measure. The language used is important in order to arrive at a scale that is measuring the phenomenon of interest. As well, the wording of items may affect the way patients perceive their disability and may affect the ultimate measure of this concept.

Part I of this study had several potential limitations. First, four of the six items had different response options for the attributed and nonattributed versions. In order to compare the two versions, the five category ordinal scale (used for the attributed items) was converted to a binary response category scale. Although the use of different response option formats might have affected the magnitude of the differences between the attributed and nonattributed questions, it was unlikely to completely reverse the response pattern. The second limitation was that the attributed questions were asked before the nonattributed questions in Part I. Part II of this study addressed these two limitations by administering the two versions of the questions in random order and creating questions with the same number of response options. Although the order in which the attributed and nonattributed questions were answered affected the size of the differences, for all questions patients were more likely to report more disability on the attributed versions. Finally, this study examined the effect of attribution to the upper limb on disability with respect to social functioning and work in a population of patients with limb complaints. These results may not be generalizable to other conditions and anatomic regions.

In conclusion, the results demonstrate that patients do not answer questionnaires the way clinicians may expect. The wording and order of questions is important and can substantially affect a patient's response. Further studies are required to determine which type of wording most accurately reflects patients' disability and whether or not patients are able to accurately attribute their disability to a specific disease or condition.

Acknowledgments

Dr. Marx is supported by the Arthritis Society, the Institute for Work and Health and the Surgeon Scientist Program at the University of Toronto. Dr. Wright is the Robert B.

Salter Chair in Surgical Research and is supported as a Scientist by the Medical Research Council of Canada.

References

- [1] Bombardier C, Melfi CA, Paul J, Green R, Hawker G, Wright J, Coyte P. Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement. *Med Care* 1995; 33:AS131–AS144.
- [2] Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;27:S217–S232.
- [3] Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622–9.
- [4] Bergner M. Health status measures: an overview and guide for selection. *Ann Rev Public Health* 1987;8:191.
- [5] Wiklund I, Karlberg J. Evaluation of quality of life in clinical trials. Selecting quality-of-life measures. *Controlled Clin Trials* 1991;12: 204S–216S.
- [6] MacKenzie CR, Charlson ME, DiGiosa D, Kelley K. Can the sickness impact profile measure change: an example of scale assessment. *J Chron Dis* 1986;39:429–38.
- [7] MacKenzie CR, Charlson ME, DiGiosa D, Kelley K. A patient-specific measure of change in maximal function. *Arch Intern Med* 1986; 146:1325–9.
- [8] Kantz ME, Harris WJ, Levitsky K, Ware JE, Ross-Davies A. Methods for assessing condition-specific and generic functional status outcomes after total knee replacement. *Med Care* 1992;30:MS240–MS252.
- [9] Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
- [10] Hudak PL, Amadio PC, Bombardier CB, Beaton D, Cole D, Davis A, Hawker G, Katz JN, Makela M, Marx RG, Punnett L, Wright JG. Development of an upper extremity outcome measure: the DASH (Disabilities of the Arm, Shoulder and Hand). *Am J Indust Med* 1996;29: 602–8.
- [11] Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 health survey manual and interpretation guide. Boston: The Health Institute, 1993.
- [12] Rosner B. *Fundamental of biostatistics*. Toronto: Duxbury Press, 1995.
- [13] Connor R. Sample size for testing differences in proportions for the paired-sample design. *Biometrics* 43:207–11.
- [14] Redelmeier DA, Rozin P, Kahneman D. Understanding patients' decisions. Cognitive and emotional perspectives. *J Am Med Assoc* 1993;270:72–6.
- [15] Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science* 1981;211:453–8.
- [16] McNeil BJ, Pauker SG, Sox HCJ, Tversky A. On the elicitation of preferences for alternative therapies. *N Engl J Med* 1982;306:1259–62.
- [17] Redelmeier DA, Koehler DJ, Liberman V, Tversky A. Probability judgement in medicine: discounting unspecified possibilities. *Med Decision Making* 1995;15:227–30.

Appendix

A.1 Attributed items

A1. During the past week, to what extent has your arm, shoulder or hand problem interfered with your normal social activities with family, neighbors or groups?

1. Not at all

2. Slightly
3. Moderately
4. Quite a bit
5. Extremely

A2. During the past week, how much of the time has your arm, shoulder or hand problem interfered with your normal social activities with family, neighbors or groups?

1. None of the time
2. A little of the time
3. Some of the time
4. Most of the time
5. All of the time

Preamble for items A3–6: During the past week, have you had any of the following problems with your work or other regular daily activities as a result of your arm, shoulder or hand problem:

A3. Cut down on the amount of time you spent on work or other activities.

1. Have not cut down the amount of time at all
2. Have cut down on the amount of time slightly
3. Have cut down on the amount of time moderately
4. Have cut down on the amount of time a lot
5. Have been unable to spend any time on work or other activities

A4. Accomplished less than you would like.

1. Able to accomplish as much as I would like.
2. Accomplishing slightly less than I would like
3. Accomplishing moderately less than I would like
4. Accomplishing a lot less than I would like
5. Unable to accomplish anything I would like

A5. Were limited in the kind of work or other activity

1. Not limited in the kinds of work or other activities at all
2. Slightly limited in the kinds of work or activities
3. Moderately limited in the kinds of work or activities
4. Very limited in the kinds of work or activities
5. Unable to accomplish any kind of work or other activities

A6. Had difficulty performing the work or other activities (for example, it took extra effort)

1. No difficulty
2. Mild difficulty
3. Moderate difficulty
4. Severe difficulty
5. Unable

A.2. *Nonattributed items*

NA1 (SF-36 question #6). During the past week, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors or groups?

1. Not at all
2. Slightly
3. Moderately
4. Quite a bit
5. Extremely

NA2 (SF-36 question #10). During the past week, how much of the time has your physical health or emotional problems interfered with your normal social activities (like visiting with friends, relatives, etc.)?

1. None of the time
2. A little of the time
3. Some of the time
4. Most of the time
5. All of the time

NA3-6. During the past week, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

Item	Yes	No
NA3 (SF-36 question #4a). Cut down on the amount of time you spend on work or other activities.	1	2
NA4 (SF-36 question #4b). Accomplished less than you would like.	1	2
NA5 (SF-36 question #4c). Were limited in the kind of work or other activities.	1	2
NA6 (SF-36 question #4d). Had difficulty performing the work or other activities (for example, it took extra effort).	1	2