Values for function in rheumatoid arthritis: patients, professionals, and public

S Hewlett, A P Smith, J R Kirwan

Abstract

Background—Although many disability questionnaires measure fact very efficiently, they do not allow for consideration of the relevance of that disability to the patient. Data suggest that professionals misinterpret the relevance of disability for the patient and thus, also, the outcome of treatment.

Objectives—Firstly, to examine agreement on levels of importance for the items on a validated disability scale (Health Assessment Questionnaire (HAQ) and Modified HAQ (MHAQ)), within groups of patients with rheumatoid arthritis, health professionals, and controls. Secondly, to see if functional items important to patients are included in the HAQ, and whether the HAQ items are important to patients.

Methods—25 patients with RA, 25 rheumatology health professionals, and 25 healthy controls were asked to rate the importance of the HAQ (20 items) and MHAQ (eight domains). Before seeing the HAQ, patients were asked to generate items of function important to them.

Results—Only a slight-fair agreement within each group was found for the level of importance of the HAQ and MHAQ, and also within any combination of the groups (k values <0.38). Most of the functional items valued by patients were contained on the HAQ (70%), and no HAQ items were consistently rated as unimportant.

Conclusion—Patients, professionals, and healthy controls do not agree on the importance of disabilities. These data support the need to assess the personal impact of disability, as well as disability itself. Individual importance of disability weighted by level of disability is proposed as a model for calculating the personal impact of disability. A new tool to assess the personal impact of disability is being developed.

Over the past 20 years outcome measurement in arthritis has undergone a paradigm shift from process measures, such as bony erosions and C reactive protein, to patient centred outcome measures, such as function or physical and mental wellbeing. Disability in rheumatoid arthritis (RA) is often high at presentation, improves during the first year, and then slowly increases, and the measurement of disability (or activity limitation as opposed to impairment at tissue level) is now a well validated science, commonplace in clinical practice and research. The Health Assessment Questionnaire (HAQ) is a self reporting tool that measures disability in arthritis and correlates well with observed performance. It has been widely reviewed, its use is widespread, and it is recommended as a core measure in trials.

Many disability questionnaires measure fact efficiently using approaches such as, “I have much difficulty climbing stairs”, and are not intended to measure the impact of these facts on the individual patient. However, disabilities such as climbing stairs, for example, may be of little importance to someone living in a bungalow but of major concern to someone with an upstairs bathroom. The differing impact of disability on individual patients may be one explanation for the relatively modest associations reported between level of disability and dissatisfaction with disability, for the discrepancy between clinicians and patients on disability levels, and for the relatively poor association between actual (calculated) change in disability and patient perception of change. Large changes in functions of little personal relevance and small changes in functions of great personal relevance may account for these results. Activities that are rated by patients with RA as difficult and important are associated with reduced psychological wellbeing, and stopping 10% or more of such activities is a predictor of depression.

A measure of disability impact, based on personal importance and used in addition to a measure of disability, would therefore be useful, allowing patients and clinicians to place a disability score within an individual context.

Attempts have been made to incorporate relevance to patients in outcome measures by using mean importance weightings derived from professionals or from the general population. Mean weightings assume agreement between patients’ views and the group from whom the mean is calculated, but a growing body of evidence suggests that patients’ views about the importance of aspects of illness differ from those of clinicians and controls. Studies have shown that doctors’ ratings of their patients’ quality of life and their most important health domains differ significantly from their patients’ views. Differences of opinion about the value of self management activities or functions have been found between patients and professionals and between patients and controls, and people who use wheelchairs rate the importance of abilities that they lack significantly lower than do controls. Whereas some studies show no significant difference between ratings given to quality of life situations by patients with RA
Table 1 Demographic data (n=25 in each group)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patients with RA</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>58.6</td>
<td>16.6</td>
<td>20–82</td>
</tr>
<tr>
<td>Disease duration (years)</td>
<td>16.7</td>
<td>10.8</td>
<td>1–42</td>
</tr>
<tr>
<td>Disability (HAQ)</td>
<td>2.16</td>
<td>0.36</td>
<td>1.5–2.75</td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>53.4</td>
<td>9.5</td>
<td>36–75</td>
</tr>
<tr>
<td>Health professionals</td>
<td>7.64</td>
<td>4.7</td>
<td>1–17</td>
</tr>
<tr>
<td>Rheumatology experience (years)</td>
<td>5.5</td>
<td>3.6</td>
<td>0–13</td>
</tr>
</tbody>
</table>

Table 2 Frequency of importance scores for 20 Health Assessment Questionnaire activities of daily living (HAQ ADLs) (n=25 in each group)

<table>
<thead>
<tr>
<th>ADL</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dress</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>20</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>Shampoo hair</td>
<td>0</td>
<td>12</td>
<td>11</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Rise (chair)</td>
<td>0</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>3</td>
<td>4</td>
<td>11</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>Rise (bed)</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>21</td>
<td>1</td>
<td>3</td>
<td>16</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>Cut meat</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>Lift cup</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>17</td>
<td>2</td>
<td>2</td>
<td>9</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>Open carton</td>
<td>5</td>
<td>11</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Walk (flat)</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>14</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>Stairs</td>
<td>0</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>Wash body</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>14</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>Bath</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>On/off toilet</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Reach up</td>
<td>11</td>
<td>10</td>
<td>4</td>
<td>0</td>
<td>9</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Reach down</td>
<td>1</td>
<td>6</td>
<td>15</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>16</td>
<td>6</td>
<td>7</td>
<td>11</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Open car door</td>
<td>6</td>
<td>13</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Open jars</td>
<td>0</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Turn taps</td>
<td>1</td>
<td>2</td>
<td>12</td>
<td>12</td>
<td>1</td>
<td>12</td>
<td>2</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Shops</td>
<td>1</td>
<td>9</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>14</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>In/out car</td>
<td>0</td>
<td>6</td>
<td>14</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Chores</td>
<td>2</td>
<td>10</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>8</td>
<td>12</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>32</td>
<td>128</td>
<td>156</td>
<td>184</td>
<td>49</td>
<td>67</td>
<td>143</td>
<td>241</td>
<td>21</td>
<td>33</td>
<td>98</td>
<td>348</td>
</tr>
</tbody>
</table>

Proportion of group (%) 6.4 25.6 31.2 36.8 9.8 13.4 28.6 48.2 4.2 6.6 19.6 69.6

0 = not at all important; 1 = a little important; 2 = quite important; 3 = very important.

Impact of disability it must include functions important to patients and exclude unimportant functions. The most common disability measure (HAQ) appropriately uses items selected for their ability to measure function, and therefore the second aim of the study is to examine whether items of importance generated by patients are included in the HAQ, and whether any HAQ items are unimportant to patients.

Methods
Thirty one rheumatology health professionals from a variety of professions, locations, and backgrounds (clinical and academic) were invited to complete the questionnaire, with five local professionals completing them as structured interviews. They were asked to rate the importance of the 20 items on the HAQ and the eight domains on the Modified HAQ in response to the question: “Some questionnaires use healthcare professionals’ values to weight patient centred scales. Could you tick a column for each of the HAQ questions to show the value or importance level you think they hold for the patient?” (not at all important, a little, quite, very important, 0–3). The HAQ comprises 20 activity of daily living questions (ADLs), grouped into eight functional categories. Level of difficulty over the previous week is recorded (no difficulty, some difficulty, much difficulty, or unable to do, 0–3) and aids or assistance required. The eight category scores are summed and averaged, yielding a disability score of 0–3, where 3 is extreme disability. The Modified HAQ (MHAQ) of Pincus et al uses eight of the ADLs to represent the 20 item version. To try to capture more of the HAQ functions while maintaining the brevity of the MHAQ questionnaire, the wording of some of the eight ADLs selected by Pincus was expanded slightly (into domains) to incorporate the missing ADLs.

Twenty eight consecutive inpatients with confirmed RA were invited to take part and interviewed using a structured format. Inpatients were used in order to find patients with high disability, as to ask patients about values for non-existent disabilities would have been an abstract question. Before being shown an HAQ, patients were asked to generate physical functions they considered personally important in response to the question: “Some people with arthritis have difficulties physically doing everyday things. When you think about your ability to do things in everyday life, what one thing bothers or upsets you the most?” Patients were then asked to place a personal value on their generated items, and on the 20 ADLs of the HAQ and the eight domains of the MHAQ in response to the question: “I would like to find out how important this is to you, or how much you value it. For example, someone who lives in a bungalow may have difficulty climbing stairs but that might not be important to them” (not at all important, a little, quite, very important, 0–3). Patients then completed an HAQ.

Fifty non-arthritic controls, primarily lay friends of colleagues, were invited to take part.
Patients with RA, 25 rheumatology health professionals, and 25 non-arthritic controls agreed to take part. Both patients and professionals had considerable personal or professional experience of RA, and only one professional had no personal experience of the HAQ (table 1). Patients and controls were of a similar age, and patients included 17 women (health professionals 20, controls 16). Patients had a high mean HAQ score (2.16). Professionals comprised nine nurses, five physiotherapists, five medical doctors, four occupational therapists, and two psychologists.

Patients gave different values to the different items, with each item having a range of importance of 0 or 1 (not at all/a little important) to 3 (very important) (tables 2 and 3). This was also true for professionals and controls, though controls designated all domains as being of some level of importance. Items were given maximum importance scores most frequently by controls (69.6% of HAQ ADLs, 82% of MHAQ domains) and least frequently by professionals (36.8% and 54%, respectively). Controls’ ratings for the importance of either ADLs or domains were generally the highest and professionals’ values were usually the lowest, with patients’ ratings lying in between (figs 1, 2, and 3).

When the group values for the full 20 HAQ ADLs were totalled (that is, 0–3 × 25 subjects) four of the most highly valued ADLs were the same for all groups (toilet, dress, wash, rise) and the least important was the same for every group (reach up) (table 4, ordered by patient scores). However, other ADLs varied widely in position (for example, rise from chair, 8th, 15th, and 19th) or in the magnitude of their importance (for example, reach up, 18, 33, and 49 points). When the group values for each of the eight domains on the MHAQ were totalled in the same manner, they were ordered identically for all groups (table 5, ordered according to patient scores). Professionals gave the bottom three domains lower values than the other groups.

There was less than slight agreement within each group for importance of both the overall HAQ and the overall MHAQ, which was no
stronger when any combination of groups was analysed (κ values <0.17, table 6). When quadratic weighted κ values were examined, agreement was slightly higher and had not arisen by chance, but was still only fair (κ values <0.38, table 6).

To see whether the HAQ adequately covered items of function important to patients and whether it contained redundant items, patients were asked to generate items of difficulty that were important to them, before seeing the HAQ (table 7). Of the 54 disability items of importance generated by patients, 37 (69%) were already included in the HAQ. Most of the remaining 17 items related to leisure activities (seven items or 13% of total patient generated items).

Discussion

The first aim of the study was to establish whether there is agreement between and within patients with RA, health professionals, and healthy controls on the importance of disability items measured by the HAQ. The mean disability (HAQ) score for these patients with RA was intentionally high, but as this group may not be representative it might have been appropriate to have included some patients with lesser disability.

Individual patients could clearly identify different levels of importance for different functions. The ratings varied between patients and they used a wider range of values than controls. Professionals gave the lowest importance to functions, healthy controls gave the highest, and patients fell between the two. One possible explanation for this difference may be that professionals attach less importance to functions because their training encourages them to teach alternative ways round problems. Healthy controls may be unable to conceive of not being able to do everything and take function for granted. Patients may be making a transition from the inexperienced views of healthy subjects towards the more pragmatic views of professionals. However, professionals may not appreciate that even if there are ways round problems, patients may still desire to do their usual tasks (for example, a shower may be a practical substitute for a bath, but some people prefer bathing to showering). It seems clear, therefore, that with the low k levels, neither a population, professional, nor even a patient mean rating for the importance of disabilities is an appropriate weighting system. The continued low level of agreement when groups of subjects are combined shows that not only do patients not have uniform values but also they do not agree with professional or control values. The MHAQ does not yield any better agreement, despite containing broader functional domains than the very specific HAQ questions.

Consequent use of the HAQ may have made professionals lose sight of the basic tenet—namely, that it was designed to measure disability alone. Familiarity may lead professionals to assume that an HAQ score has a universal value (for example, a score of 2.5 will mean the same to all patients in practical and...
emotional terms). Secondly, for most professionals an HAQ score conjures up a mental picture of what such a disability means—that is, we feel we know what that universal value is. Finally, in the absence of any other information we assume that the disability is important to the patient. By failing to assess the way a particular functional loss affects the individual patient, we may be guilty of imposing our own judgments upon factual disability scores. The study shows that HAQ items do not have a universal value and that health professional assumptions about values do not accord with patient values.

The second aim of the study was to see whether a well respected disability questionnaire (HAQ) could be used as a basic tool to calculate an impact score. When patients were asked to generate items of disability important to them, most items generated (69%) were included in the HAQ and although up to 10% of patients overall place no value on some HAQ items, no HAQ item was consistently deemed unimportant. Leisure activities were the only significant items of importance to patients omitted from the HAQ. The items in the HAQ were appropriately chosen to represent function, but it appears to contain a large number of the items deemed important by patients, reassuring us of its validity in measuring relevant disability, and a sufficient number for it to be used as a basis for capturing the impact of disability.

The study data suggest three ways in which a measure of the impact of disability on patients might be developed. These are (a) to create a new measure that includes leisure items; (b) to add an open question to the HAQ for the patient to insert their important leisure items; or (c) to accept this limitation of the HAQ for measuring impact but still use it as a basis for the new impact tool. It would be inappropriate to develop a new functional measure as a basis for calculating impact when the HAQ is a well validated and internationally respected functional measure covering 70% of items important to patients. Adding an open question for the purpose of calculating impact would cause interpretation problems as there would be no corresponding difficulty score, the topic of the open question would have to be inserted when the scale was re-administered, and it is not obvious how the open question should be handled if the previously designated topic was no longer important but a new problem had arisen. This has been recognised as a difficulty with other scales which include items specified by individual patients. On balance it would seem appropriate to use the HAQ as a basis for an impact measure as it covers most items important to patients, but acknowledging the gap left by leisure activities.

A small change in a function score can be clinically significant for the patient but not statistically significant. Clinical significance for the patient must mean that something of personal importance to them has occurred. The personal impact of disability has been described as “the degree to which it disrupts the life situation”, though life disruption may not depend on disability alone. Assessment of disability is currently measurement of fact (assuming the HAQ to reflect actual disability), but for the patient self assessment of difficulty is made in the context of personal circumstances, something that has been lost in standardised disability measurement.

We therefore propose a model whereby the personal impact of disability is an interaction between disability for ADLs and the belief that the ability to perform those ADLs is important (or valued) by the patient. One method of calculating personal impact would be to weight the level of difficulty for each ADL by the patient’s opinion of the importance of being able to perform that ADL (fig 4). Creating separate profiles of disability and importance (rather than a weighted scale) would not aid the interpretation of disability, as it is only when difficulty occurs in a personally important ADL that it is likely to impact upon the patient.

The principle of weighting has been explored elsewhere. The problem elicitation technique was developed from the MacMaster Toronto arthritis patient preference disability questionnaire but is administered by an interviewer, making it less convenient. The PARIS sectorgram weights three symptoms in osteoarthritis (although the weights are relative to each other), whereas weighted scales have been used in general rehabilitation. Conceivably, the addition or removal of mean weights in a scale may not make large differences to the validity and sensitivity, but the use of individualised weights may be more important in this respect.

Using a personal impact score to complement (but not replace) a disability score will provide a method of differentiating between the impact of disability in patients with similar disability levels and allow recognition of the impact of a relatively minor disability in some patients. It will help to inform decisions on intervention, in the assessment of clinical trials of treatments, and to determine whether a treatment affects disability in a meaningful way for patients. It will allow us to justify the expense of treatments according to the benefit perceived by a patient rather than the assumed benefit and may help to identify patients at risk from reduced psychological wellbeing or depression. More importantly, it will help us to understand the patient better.

Concepts of outcome measurement underwent a major shift into patient centred assessment by the publication of the HAQ and it is now time to place the measurement of

![Figure 4 Proposed model for calculating the personal impact of disability](http://ard.bmj.com/content/46/9/2905/F4.large.jpg)
disability within the patient’s value system. The development of tools measuring patient values or important changes in function has long been advocated in arthritis,19,20 and this study provides further evidence of the need to measure the impact of disability on individual patients and justifies using the HAQ as an appropriate basic tool. A model of the personal impact of disability is proposed: disability weighted by the importance or value for that disability to an individual patient. Further work is under way to develop and validate such a measure of the personal impact of disability.

The authors thank statisticians Dr Barnaby Reeves and Rosemary Greenwood (University of Bristol Research and Development Support Unit) for their advice on the use of k measures of agreement; the Arthritis Research Campaign for its support; and patients, professionals and volunteers for their participation.

13 Hewlett S, Kirwan JR. Discrepancies between actual and perceived change in function in rheumatoid arthritis are not a function of memory [abstract]. Br J Rheumatol 1998;37(suppl 1):177.