# Does the Stanford Health Assessment Questionnaire have potential as a monitoring tool for subjects with rheumatoid arthritis?

M C Greenwood, D V Doyle, M Ensor

**Abstract**

***Objective***—To assist in the interpretation of the Stanford Health Assessment Questionnaire (HAQ) score changes for individual patients with rheumatoid arthritis (RA), by determining the minimum size of score change that can confidently be considered to reflect a significant change in disability from the patient's perspective.

***Method***—HAQ score changes were calculated for 40 clinic patients with RA who had reported no change to health in general over two months. These were considered to reflect both inconsistencies in questionnaire completion and any true but minor changes not considered significant enough by the patients to represent a change to their health in general. HAQ score changes over one year were also calculated for 207 clinic patients with RA.

***Results***—The range within which 95% of score changes would be expected to lie in the absence of significant change was estimated as ±0.48 points (±2SD of the score changes) and 80% within ±0.31 points (±1.29SD). A $\chi^2$ test showed no significant association between an HAQ score increase of >0.31 over one year and decline in health related to arthritis reported by the patient over the same period.

***Conclusion***—As a general guideline, an HAQ score needs to change by 0.48 points or more for 95% confidence that it reflects significant change (0.31 for 80% confidence). Although the value of HAQ as a group outcome measure is well established, this study questions the usefulness of monitoring individual HAQ scores in a clinical setting.

(*Ann Rheum Dis* 2001;**60**:344–348)

Rheumatoid arthritis (RA) is a progressive condition characterised by inflamed and painful joints. Typically, the patient experiences intermittent flares and increasing disability due to the cumulative effects of joint damage and the social and psychological effects of living with a painful, debilitating, and unpredictable condition. Minimising the impact of the disease on all areas of the patient's life is a major objective of care. However, assessing the degree to which this goal is being achieved for each patient is not always straightforward. Gradual change over many years may be overlooked even by the individual patient themselves, especially if over a period of several years their expectations and priorities have changed.

Gradual changes will be especially difficult to assess for new staff taking over care of the patient or if the disease is affecting aspects of a patient's life such as social function that are not readily apparent in a clinic setting. In addition, the patient's own perception of outcome may well differ markedly from that of the clinician. The regular formal assessment of outcome (metrology) has therefore been advocated to provide a long term record of change across a wide range of health domains.[1] The aim is to assist the clinician, nurse, physiotherapist, or therapist in the assessment of individual patient outcome

The Stanford Health Questionnaire (HAQ) was designed to measure disability in arthritis and is widely used in rheumatology in the United Kingdom.[2] It asks patients to rate degree of difficulty in performing 24 everyday activities and to indicate if they use certain aids and devices or need help in certain areas of activity. It generates a score on an ordinal scale from 0 (minimum disability) to 3 (maximum). It has been well validated as an outcome measure for groups of patients in clinical trials and at a group level is sensitive to change[3] and predictive of long term outcome.[4]

At Whipps Cross Hospital in East London we invite all patients with RA to attend an annual metrology appointment for outcome assessment. This includes completion of the HAQ, which enables the assessment of outcome for disability at the group level. However, HAQ has also been advocated for use as a measure of individual patient outcome.[1 5 6] In this department, changes in HAQ score have been calculated for individual patients over periods of up to seven years. The overall objective of this study was to establish whether it might be possible to infer reliable information about individual patient outcome from a change in their HAQ score. We aimed, firstly, at answering the question, what is the minimum level of HAQ score change that could confidently be considered to reflect a significant change in disability from the patient's perspective, and, secondly, would this have potential as a clinical tool for alerting clinicians to significant change that might otherwise be overlooked?

## Method

What constitutes significant change in disability can have a number of different interpretations depending on the context. However, here we wished to identify the minimum level of change that patients themselves would consider significant. Therefore we did not set out

**Rheumatology Department, Whipps Cross Hospital, Leytonstone, London E11 1NR, UK**
M C Greenwood
D V Doyle
M Ensor

Correspondence to: Dr Doyle
Mandy.Greenwood@
Forhc-tr.nthames.nhs.uk

Accepted 29 August 2000

|  | No change in health in general over 2 months (n=40) median (min–max) | Annually assessed clinic patients (n=207) median (min–max) |
|---|---|---|
| Age (years) | 64 (48–83) | 66 (32–90) |
| Disease duration (years) | 11 (1–29) | 12 (1–77) |
| First HAQ* score | 1.38 (0–2.625) | 1.50 (0–3) |
| Second HAQ score | 1.32 (0–2.625) | 1.50 (0–3) |
| No (%) With activities limited by comorbidity | 8 (20) | 43 (21) |
| No (%) Female | 30 (75) | 145 (70) |

*HAQ = Health Assessment Questionnaire.

to examine test-retest reproducibility or the size of HAQ score change associated with true but minor changes in disability. For the purpose of this study, significant change was defined as a change in level of disability that, over a period of two months, patients themselves considered significant enough to constitute a change to their health in general. It was considered that a period of two months would be short enough to minimise the confounding effects of any change in patients' expectations.

It is not possible to comment with certainty on the significance of a particular HAQ score change to any particular individual patient. However, one can make an assessment of the probability that it reflects true change, based on a knowledge of the observed distribution of score changes for similar patients in the absence of significant change over a short period. A guideline could be based on a 1%, 5%, or 20% probability of finding a score change so large in the absence of significant change, depending on the use for which it is required. If the looser criterion of 20% is used the clinician could go on to check its true significance by discussing the matter further with the patient concerned. We have used the distribution of score changes over two months in the absence of patient perceived change to health in general to estimate both the 20% and the 5% levels. To do this we applied the approach suggested by Bland and Altman[7] for quantifying the repeatability of a test. Although it was not strictly repeatability that we were assessing, the same principles applied. Assuming that the distribution of score changes, though discrete, can be approximated by a normal distribution with a mean of zero, 95% of score differences would be expected to lie within ±2SD and 80% within ±1.29SD (corresponding to probabilities of 5% and 20% of observing so large a score change in the absence of significant change in disability). It was expected that within these ranges would lie 80% and 95% of the score changes that were associated with inconsistencies in questionnaire completion and with any true change that patients did not consider significant enough to constitute a change to their health in general.

HAQ SCORE VARIABILITY OVER TWO MONTH INTERVALS
Fifty two consecutive rheumatology clinic patients with RA who were attending for routine annual outcome assessment were requested to complete an HAQ every two months for a year. Each time patients were asked to rate their health in general now as compared with at the time of the previous

assessment two months before. Options for answering were "much worse, somewhat worse, about the same, somewhat better, and much better". For the final six months subjects were also asked to rate the severity of their arthritis now compared with at the time of completion of the previous questionnaire. The options for answering this question were the same as for the change in health in general question. All but two agreed to participate. Table 1 shows the patients' characteristics.

Three of the patients were excluded from the analysis because they had consistently scored zero throughout the year. Any disability that they might have been experiencing was thus too low to be assessable by the HAQ. To assess the level of score change that might typically occur among patients reporting no significant change in condition, the score changes were calculated for the first two month period over which each of the subjects reported no change in health. Forty of the remaining subjects reported at least one such period. To check whether there was a relation between the size of score difference and position on the scale, score change was plotted against the midpoint between the two scores. The assumption that the mean difference was zero was also checked using a one sample *t* test. With a mean difference of zero, the standard deviation could be calculated as suggested by Bland and Altman by squaring the differences, adding them up, dividing by n, and then taking the square root.

For comparison, the HAQ score changes were also calculated for the second two month period over which patients reported no change in health in general (n=32) as well as for the first two month period over which each patient had reported that the severity of their arthritis had remained about the same (n=32). Again, the estimates for the limits within which 80% and 95% of cases would be expected to lie, in the absence of a change in the severity of their arthritis reported by the patient, were calculated as ±1.29 and ±2SD, respectively.

ASSOCIATION BETWEEN HAQ SCORE CHANGE AND PATIENT PERCEIVED, ARTHRITIS RELATED CHANGE IN HEALTH OVER ONE YEAR
The HAQ score changes over one year were calculated for all 207 rheumatology clinic patients who had attended for routine annual metrology assessment on two consecutive years (1997–98 and 1998–99) (table 1). Using the SF-36 questionnaire health transition question patients were asked to compare their health in general now with one year previously. Those who reported better or worse health were asked whether the change was due to their arthritis. Score changes were also calculated with a 10 cm visual analogue scale for patient assessment of pain over the past week, anchored at one end by "No pain" and at the other by "Pain as bad as it could be". Over a period of a year a patient's assessment of change in health in general may be affected by other factors than simply change in disability. For instance, other circumstances in their lives may have led them to modify their expectations. However, it was
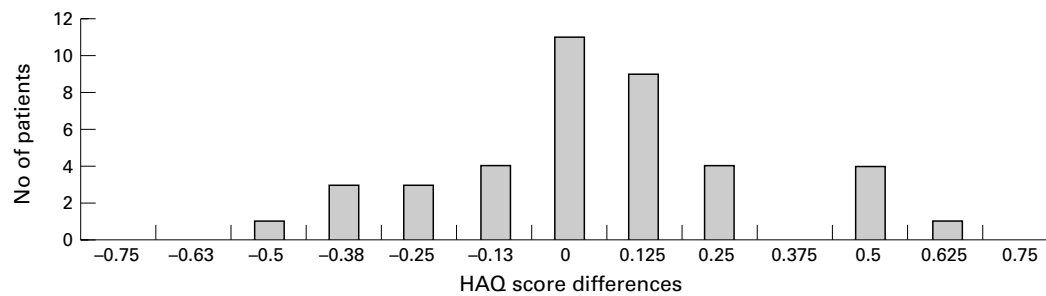
*Figure 1   Health Assessment Questionnaire (HAQ) score differences over two months—40 patients with rheumatoid arthritis reporting no change in health in general.*

reasoned that as a group, those who had truly had a significant increase in disability would be more likely than others to have reported a decline in health related to arthritis. Using the $\chi^2$ test, we therefore tested the hypothesis that among clinic patients with RA there would be a statistical association between an increase in HAQ score of 0.31 points or more over one year and self reported, arthritis related decline in health over the same period. We had earlier estimated by the method described above that only 10% of patients would be expected to have score increases of this magnitude in the absence of a significant change in disability.

To assess the sensitivity of HAQ to change, the annual HAQ score changes of all the patients who had attended for metrology on both years were grouped according to their assessment of arthritis related change in health in general over the year. Nineteen patients who had reported a change in health in general that was not due to their arthritis were excluded from this analysis so that the groups containing those most likely to have experienced a change in disability—that is, those reporting an arthritis related improvement or decline in health in general, could be compared with the group reporting no change. The Kruskal-Wallis one way analysis of variance was used to test for the presence of significant differences between the groups in the distributions of their HAQ score changes.

## Results

HAQ SCORE VARIABILITY OVER TWO MONTH INTERVALS

Figure 1 shows the calculated HAQ score change for the first two month period that each
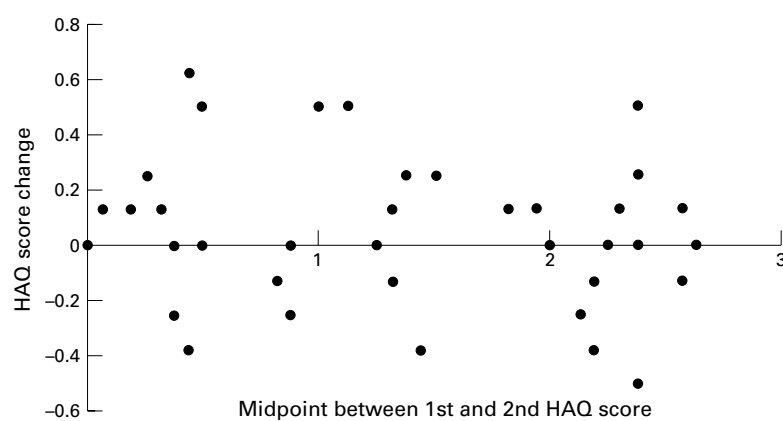


*Figure 2   Size and direction of Health Assessment Questionnaire (HAQ) score change was not related to position on the scale. HAQ score changes over two months (with no patient perceived change in health in general) plotted against the midpoint between the two scores of each subject.*

subject reported their health to have remained about the same (n=40). There was no significant relation between the size of the score change and position on the scale (see fig 2). The estimated limits within which 95% of score changes would be expected to lie in the absence of a significant change in disability were calculated as ±0.48 points and the 80% limits as ±0.31 points. For comparison these limits were also calculated using the score changes over the second reported period that health was reported to have remained about the same (n=32). This gave the figures of ±0.42 points and ±0.27 points for 95% and 80% respectively. Also, for comparison, the estimated limits within which 80% and 95% of differences would be expected to lie in the absence of a patient perceived change in the severity of their arthritis were calculated as ±0.28 and ±0.44 points respectively (n=32).

ASSOCIATION BETWEEN HAQ SCORE CHANGE AND PATIENT PERCEIVED, ARTHRITIS RELATED CHANGE IN HEALTH OVER ONE YEAR

Of the 207 patients with RA who had attended for routine annual metrology assessment in both years, 1997–98 and 1998–99, the HAQ score had increased by more than +0.31 points (the 80% confidence limits of repeatability over two months) in only 29 cases. In practice, because HAQ is an ordinal scale it is not possible to score 0.31, so effectively a change of ±0.375 would need to be used. Seventy six patients had reported an arthritis related decline in health over the year (table 2). Compared with other patients, this group had significantly higher increases in pain on the visual analogue scale (two tailed significance = 0.000). However, despite the fact that pain is a major cause of disability in RA, their changes in HAQ were not significantly different (two tailed significance = 0.262) from the 131 not reporting a decline in health due to their arthritis (table 3). In addition, there was no significant association between an HAQ score

*Table 2   Cross tabulation of patient perceived, arthritis related decline in health against a Health Assessment Questionnaire (HAQ) score increase of 0.31 points or more over the same year*

| | Self reported, arthritis related decline in health | | |
|---|---|---|---|
| *Counts* | *Yes* | *No* | *Row total* |
| HAQ score increase of 0.31 points or more | | | |
| Yes | 12 | 17 | 29 |
| No | 64 | 114 | 178 |
| Column total | 76 | 131 | 207 |

*Table 3   Changes in pain visual analogue scale (VAS) and Health Assessment Questionnaire (HAQ) over one year. A comparison between those who did and did not report an arthritis related decline in health over the same year*

| Patient reported, arthritis related decline in health | Pain VAS Median (lower quartile, upper quartile) | HAQ Median (lower quartile, upper quartile) |
|---|---|---|
| Yes (n=76) | +0.80 (−0.20, +2.6) | 0.00 (−0.12, +0.25) |
| No (n=131) | 0.00 (−1.90, +0.93) | 0.00 (−0.12, +0.13) |
| Two tail significance for Mann-Whitney test after Bonferroni correction | 0.000 | 0.262 |

increase of 0.31 points or more over a year and the tendency for patients to report an arthritis related decline in health over the same period (two tailed significance for $\chi^2$ test = 0.574). Kruskal-Wallis one way analysis of variance did show a significant relation between annual HAQ score change and the patient's perception of arthritis related change in health in general (two tailed p=0.004), but it can be seen from fig 3 that it was the distribution of score changes of the 28 who reported an arthritis related improvement in health which differed significantly from the distributions of those who reported either no change or an arthritis related decline in health. The Mann-Whitney test gave a two tailed significance of 0.001 for the difference between the distribution of score changes of the 28 who reported an arthritis related improvement in health and that of the 160 reporting either no change or an arthritis related decline.

Agreement between the subjects' perception of change in general health and their perception of change in the severity of their arthritis was tested by calculating the κ statistic for each of the three periods on which subjects were asked about both. This gave the agreement as a proportion, after correcting for the amount of agreement that would have been likely to have occurred by chance. The κ statistics for the three periods were 0.72 (n=38), 0.69 (n=37), and 0.76 (n=39).
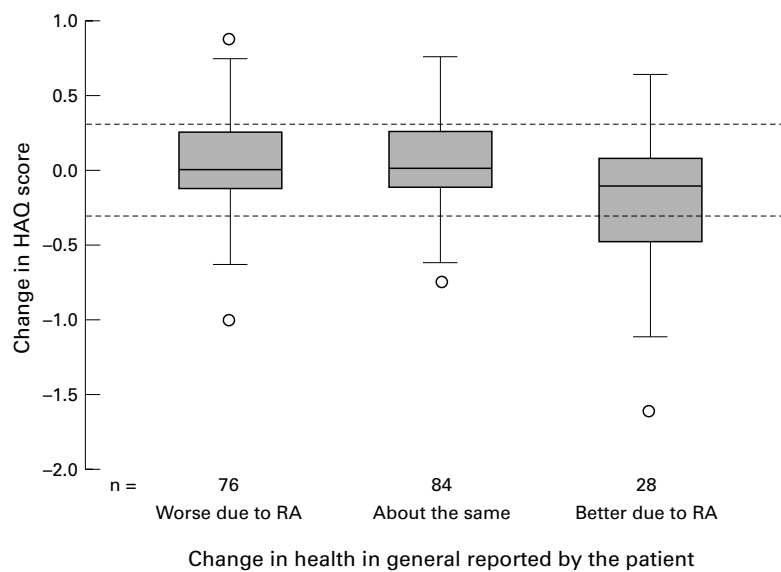


*Figure 3   Box plots of annual Health Assessment Questionnaire (HAQ) score changes for patients with rheumatoid arthritis (RA) attending for routine outcome assessment (excluding 19 who reported change not related to their arthritis). The box plots show the median, the upper and lower quartiles (upper and lower edges of the box), the range excluding outliers (whiskers), and outliers (small circles). The dashed lines indicate the calculated limits within which 80% of differences would be expected to lie based on the score changes over two months for no patient perceived change in health in general.*

## Discussion

It is not possible to comment with certainty on the significance of a particular HAQ score change to any individual patient. However, one can make an assessment of the probability that it reflects true change, based on the knowledge that only x% of similar patients would be likely to have a change in score of as much as y over two months in the absence of a patient perceived change in health.

This study was concerned with evaluating the amount of score variation that can occur in the absence of any change that the individual patient would consider a significant change to their health in general. Thus it was not concerned with the amount of agreement in the absence of a patient perceived change in disease severity. However, it is worth noting that with κ scores of from 0.69 to 0.76, there was a good level of agreement between the subjects' assessments of change in health and their assessments of change in the severity of their arthritis, indicating that severity of arthritis is a major factor in determining a patient's perception of change in health in general. This is also reflected in the observation that there was only slightly more agreement in HAQ scores over two months when subjects reported no change in the severity of their arthritis than when they reported no change in their health in general.

This study has found that a significant number of people can show marked changes in HAQ score over just two months while considering their health in general to have remained about the same. Whether using the 80% or 95% level or the first or second period for which there was no reported change in health in general, the minimum level of change necessary to be able confidently to consider a score change significant is high relative to our previously reported five year mean group increase in HAQ (0.23 points) for 46 patients with RA attending clinic.[8]

One reason for this high level of score change over such a short period of apparent stability may be the subjectivity of the HAQ questionnaire. The perception of degree of difficulty in performing an activity is highly subjective and thus likely to be influenced by extraneous factors such as mood. Other factors acting on the patient, such as problems at home or work, might cause them to rate as "much difficulty" what on another day in a better mood they might have described as "some difficulty".

A further explanation may lie in the way that some patients may perceive the short term fluctuations in pain and disability that are characteristic of RA. Possibly, some of the subjects had become accustomed to a certain level

of variation in disability month to month or even day to day and had allowed for this when judging whether they had experienced a change to health in general. Such people would only have considered a change in disability to contribute significantly to a change in their health in general if it was greater than the day to day or month to month fluctuation that they had become used to. The large score changes found for some patients who had reported no change in health in general may therefore have arisen because HAQ was sensitive to changes that many patients considered as normal variation and therefore not significant.

For assessing long term patient outcome, it is not the short term daily or monthly variations that are of interest but rather any underlying long term trends towards increasing disability. If the amount of short term fluctuation is large relative to the underlying rate of progression then this will limit the value of HAQ in identifying those who have experienced a significant progression in disability over time. Consider a patient who has regularly experienced short term fluctuation in disability (the good and bad days that many patients report) but who has actually experienced no overall trend towards increasing disability over several years. A large increase in HAQ score over this period—for example, 0.5 points, might be interpreted as a poor outcome. However, possibly, the score increase could equally well result from the chance assessment of the patient, initially on one of their "good days" and finally on a relatively "bad day".

When defining a minimum level of score change that might alert clinicians to significant change in disability from a patient's perspective, we would not wish to identify patients who have simply experienced fluctuation within their accustomed level. Therefore, it is quite valid to take into account the score changes of all the patients who had reported that their health had remained about the same, including any who might have experienced some fluctuation.

One limitation of using standardised questionnaires such as HAQ to monitor change in individual patients is that it is only possible to develop general guides to assist clinicians in interpretation of the results. A large change in score is required to be confident of real change, but for many patients a smaller score change might well be meaningful to them individually.

Another serious limitation is that although a questionnaire might adequately assess a concept such as the level of disability experienced by a group of patients, it cannot be assumed that it will equally well assess this concept in an individual patient. The standardised set of questions might very well be inappropriate for a particular individual patient and yet omit other activities that could be causing them great difficulty. Thus the use of a questionnaire such as the HAQ might be misleading in some situations.

For assessment of individual patient outcome there is no substitute for a sensitive and thorough discussion with the patient themselves and a thorough clinical examination. This may not yield quantifiable information but is more likely to fulfil the ultimate aim of optimising outcome by identifying the specific needs of the individual patient and tailoring care and support to meet these needs.

Owing to the limitations of using standardised questionnaires for assessment of individual patient outcome, a number of alternative, patient centred approaches are now being used. Two examples are the Patient Generated Index[9] and the Disease Repercussion Profile.[10] Rather than requiring standardised answers to standardised questions these allow the individual patient room to identify their own particular problems and priorities.

Wiles *et al* found considerable within-patient variation when HAQ was assessed annually in patients with early inflammatory polyarthritis and concluded that it is not possible in the early years to track disability using centile reference charts.[11] Our study, on the other hand, looked at within-patient HAQ score variation over a much shorter period of two months, and in patients with the much longer median disease duration of 11 years. Overall, although the HAQ is a well established and well validated measure of disability of groups, neither study supports the routine use of the HAQ in clinical practice for identifying significant change in individual patients.

1 Long AF, Scott DL. Measuring health status and outcome in rheumatoid arthritis within clinical practice. Br J Rheumatol 1994;33:682–5.
2 Carr A, Thompson P, Young A. Do health status measures (HSM) have a role in rheumatology [abstract]? Br J Rheumatol 1997;36(suppl):286.
3 Hawley D, Wolfe F. Sensitivity to change of the Health Assessment Questionnaire (HAQ) and other clinical and health status measures in rheumatoid arthritis. Arthritis Care and Research 1992;5:130–6.
4 Kaziz LE, Anderson JA, Meehan RF. Health status as a predictor of mortality in rheumatoid arthritis. J Rheumatol 1990;17:609–13.
5 Pincus P. Documenting quality management in rheumatic disease: are patient questionnaires the best (and only) method)? Arthritis Care and Research 1996;9:339–48.
6 Wolfe F. Data collection and utilisation: a methodology for clinical practice and clinical research. In: Wolfe F, Pincus T, eds. *Rheumatoid arthritis: pathogenesis, assessment, outcome, and treatment.* New York: Marcel Decker, 1994:463–514.
7 Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;i:307–10.
8 Greenwood M, Doyle D. Can long-term trends in the progression of disability be identified by serial assessment with the Stanford Health Assessment Questionnaire [abstract]? Ann Rheum Dis (EULAR Abstract Supplement) 1999:90.
9 Ruta DA, Garratt AM, Leng M, Russell IT, Macdonald LM. A new approach to the measurement of quality of life—the patient-generated index. Med Care 1994;32: 1109–26.
10 Carr AJ. A patient-centred approach to evaluation and treatment in rheumatoid arthritis: the development of a clinical tool to measure patient-perceived handicap. Br J Rheumatol 1996;35:921–32.
11 Wiles N, Barrett J, Barrett E, Silman A, Symmons D. Disability in patients with early inflammatory polyarthritis cannot be "tracked" from year to year: an examination of the hypothesis underlying percentile reference charts. J Rheumatol 1999;26:800–4.