King's Research Portal

Validation of the Musculoskeletal Health Questionnaire (MSK-HQ) in inflammatory arthritis: a psychometric evaluation

Sam Norton,[1,2] Benjamin Ellis,[3] Beatriz Santana Suárez,[1] Samana Schwank,[1] Ray Fitzpatrick,[4] Andrew Price,[5] James Galloway[1]

1. Academic Rheumatology, Faculty of Life Sciences and Medicine, King's College London, London, United Kingdom

2. Psychology Department, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

3. Imperial College Healthcare NHS Trust, London, United Kingdom

4. Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

5. Nuffield Department of Orthopaedics, Rheumatology & Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Oxford, United Kingdom

Corresponding author:

James Galloway, Academic Rheumatology, King's College London. james.galloway@kcl.ac.uk

Abstract

Objectives: The Musculoskeletal Health Questionnaire (MSK-HQ) is a recently developed Patient Reported Outcome Measures (PROM) for use across patients with different musculoskeletal conditions. This study provides a validation of the MSK-HQ examining construct validity and reliability in inflammatory arthritis.

Methods: 287 adults with inflammatory arthritis completed the MSK-HQ and other PROMs at baseline and again after three-months. Construct validity was assessed using item response theory (IRT) methods. Concurrent validity of the MSK-HQ was considered in relation to the Health Assessment Questionnaire (HAQ) and EQ5D in all patients, as well as the Rheumatoid Arthritis (RA) Impact of Disease (RAID), Psoriatic Arthritis (PsA) Impact of Disease (PsAID) scales, and Ankylosing Spondylitis (AS) Quality of Life Questionnaire (ASQoL) in disease subtypes.

Results: The MSK-HQ was approximately normally distributed with no evidence of floor or ceiling-effects. A unidimensional structure was confirmed though two items were less weakly related to the latent musculoskeletal health variable, providing some evidence of multidimensionality. Reliability across the range of responses was high (α=0.93). The total scale score correlated highly with the HAQ (r=-0.81), EQ5D index (r=0.80) and EQ5D visual analogue scale (r=0.60) in all patients, RAID in RA patients (r=-0.88), PsAID in PsA patients (r=-0.88), and ASQoL in AS patients (r=-0.86). Test-retest reliability over 3 months was high (r_{ICC}=0.73).

Conclusions: This study provides evidence for the validity and reliability of a new musculoskeletal health PROM in people with inflammatory arthritis. The major advantage of the MSK-HQ is that it is not disease specific and has high content validity in rheumatological conditions. The MSK-HQ score will be of value in clinical rheumatology practice, providing a measure which can be used across disease areas.

Keywords
Quality of life; outcomes measures; inflammatory arthritis

Key messages
- The MSK-HQ is a valid and reliable measure of musculoskeletal health related quality of life.
- It is not disease specific and can be used across inflammatory arthritis subtypes.
- A shortened version, excluding two items assessing understanding and self-efficacy, is potentially optimal.

Introduction

Collecting patient views about their own health using Patient Reported Outcome Measures (PROMs) is central to evaluating clinical care in routine practice and research.[1,2] For instance, PROMS can act as a catalyst for organisational change by improving standards, as evidenced by the NHS England National PROMs Programme.[3] A large number of PROMs currently exist, many addressing aspects of musculoskeletal health. Generic PROMs assessing health related quality of life (e.g. 36-Item Short Form Health Survey [SF-36], EuroQol-5D [EQ-5D]) allow for broad comparisons to be made across health conditions. However, generic instruments may have limited sensitivity to compare between musculoskeletal conditions since they do not necessarily capture constructs pertinent to musculoskeletal health.[4] Within rheumatology, PROMs assessing musculoskeletal health related quality of life have, in general, been developed specific to individual diseases (e.g. Rheumatoid Arthritis Impact of Disease (RAID), Psoriatic Arthritis Impact of Disease (PsAID); until recently there was no musculoskeletal health PROM validated across the spectrum of musculoskeletal conditions.

The Arthritis Research UK Musculoskeletal Health Questionnaire (MSK-HQ) is a recently developed PROM assessing musculoskeletal health related quality of life for use by patients with different musculoskeletal conditions, across a range of settings.[5] This scale incorporates the assessment of physical symptoms such as pain and fatigue, plus the impact of disease on psychological well-being.

The initial development of the MSK-HQ included a detailed scoping exercise and qualitative development phase involving both patients and clinicians.[5] In a sample of 570 people with osteoarthritis, the MSK-HQ was demonstrated to have good internal consistency between items, as well as high test-retest reliability and convergent validity compared to other measures of musculoskeletal health used in osteoarthritis populations. While initial findings are promising, additional validation studies are needed to confirm the utility of the MSK-HQ as an outcome measure for more widespread use.[6] Specifically, though content (e.g. face) validity for the MSK-HQ is high, further research is needed to examine the construct validity of the scale, by confirming the assumed unidimensional structure, and its reliability in other musculoskeletal conditions.

The aim of the present study was to examine the construct validity and test-retest reliability of the MSK-HQ in people with inflammatory arthritis.

Methods

Sample

Patients were recruited from six centres between September 2015 and July 2016. Eligibility criteria were adults with an established diagnosis of inflammatory arthritis (rheumatoid arthritis (RA), undifferentiated inflammatory arthritis (UIA), psoriatic arthritis (PsA), ankylosing spondylitis (AS)) who were initiating a new synthetic or biologic anti-rheumatic medication as part of routine care. Follow up data were collected by post at three months, following baseline visit.

Outcomes measures

MSK-HQ

The MSK-HQ consists of 14 items relating to facets of musculoskeletal health including pain, fatigue, physical function, symptom interference, sleep, self-efficacy and psychological well-being.[5] Respondents rate how much their musculoskeletal condition in the previous two weeks has affected each facet using a five-point ordinal scale from "not at all" to "extremely", respectively scored from four to zero. Items 12, relating to understanding, and 13, relating to self-efficacy, have their response scale reversed where "not at all" is scored zero. Items are summed to generate a total score with a range from 0 to 56, where higher scores indicate better musculoskeletal health status. A

fifteenth item assessing physical activity in the past week is not included in the total score and was therefore not considered in the current analysis. Previous analysis in non-inflammatory musculoskeletal disease has demonstrated good reliability (Cronbach's α = 0.88, test-retest = .84).[5] Sample copies of the MSK-HQ and a license for use can be obtained at https://innovation.ox.ac.uk/clinical-outcomes/patient-reported-outcome-measures

Other PROMs

All participants completed the Health Assessment Questionnaire (HAQ),[7] which assesses functioning across eight domains. This is the most widely used measure of functional ability in musculoskeletal conditions and 20 items assessing the difficulty of completing specific daily activities across eight domains. A further 18 items with a binary response format relating to the use of assistive devices or help from others. Scale scores range between 0 and 3, increasing in increments of 0.125, where higher scores indicate worse functional limitation.

All participants also completed the EQ5D-5L a generic measure of health related quality of life. The version used includes response categories with five levels of severity for each of five items assessing mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The responses to these items are combined against value sets to generate an index value (utility) that reflects the preferences compared to other responses. An additional item asks respondents to rate the health on a 0 to 100 visual analogue scale (VAS).

Patients with RA and UIA completed the RAID questionnaire.[8] This scale consists of seven items scored between 0 and 10. The total score is calculated as a weighted mean across all items where weights are based on patient rating of importance. The scale score ranges between 0 and 10, where higher scores indicate greater impact of disease. A minimum clinically important difference (MCID) of an absolute change of 3 has been recommended.[9]

Patients with PsA completed the nine item version of the PsAID.[10] This scale consists of nine items scored between 0 and 10. As with the RAID, the total score is calculated as a weighted mean across all items providing a scale score ranging between 0 and 10, where higher scores indicate greater impact of disease. An MCID of 3 has also been proposed.

Patients with AS completed the Ankylosing Spondylitis Quality of Life questionnaire (ASQoL) [11] and Bath Ankylosing Spondylitis Disease Activity Index (BASDAI).[12]

Data analysis

Construct validity was assessed using item response theory (IRT) methods.[13] These approaches assume responses to the set of items forming the scale are due to a single underlying latent variable (e.g. musculoskeletal health). Therefore, unidimensionality was initially examined using principal components analysis with parallel analysis to determine the number of underlying latent variable explaining more than chance variance in item response correlations.[14] Since the response scale is ordinal, the polychoric method was used to estimate the correlation between items.[15] After examining unidimensionality, IRT methods explored the functioning of individual items. A graded response model for ordinal items was estimated.[16] Preliminary analyses confirmed a graded response model provided a better fit to the data than a partial credit or generalised partial credit models. The graded response model estimates a discrimination parameter for each item and a difficulty parameter per response category for each item. Respectively, these parameters represent the strength of the association (correlation) between the item and the level of the latent musculoskeletal health variable being assessed, and the level of the latent variable where an individual is estimated to have 50% probability of responding positively to the response category for

the item. Items with low discrimination that do not correlate highly with the latent musculoskeletal health variable are considered poor indicators that should not be included in the scale.[14] Here we use the rule of thumb from the factor analytic literature that items that correlate less than 0.6 are weak.

To test the assumption that scores on the MSK-HQ are comparable across disease groups, differential item functioning (DIF) was assessed using ordinal logistic regression models. This approach, which is a reparameterisation of the graded response model,[17] sequentially compares models for differences in both difficulty and discrimination parameters (uniform and non-uniform DIF). A false discovery rate correction was applied to control for multiple testing.

Reliability, representing the precision with which the scale score assesses an individuals' true level of musculoskeletal health, was assessed using four methods. Using the baseline assessment the association between items was summarised using Cronbach's estimate of internal consistency α. Measurement precision across the range of musculoskeletal health was further assessed by examining the test information function from the graded response model. A change greater than measurement error was computed based on the reliable change index,[18] which uses alpha and the baseline standard deviation to determine the 95% interval where changes that may be due to chance alone would lie. Test-retest reliability for the scale score was assessed using the intra-class correlation coefficient and for individual items using Kendall's concordance coefficient, a descriptive measure indicating agreement between ordinal rankings. All analyses were conducted in Stata 14.2.

Results

The sample included people with RA (208, 72%), PsA (45, 16%), AS (22, 8%) and UIA (12, 4%). All patients were included in the main analyses; disease specific analyses were performed for RA and PsA only. Of the 287 providing data at baseline, 178 (62%) provided data again after three weeks to allow test-retest reliability to be assessed. Baseline information is provided in Table 1.

Information about the functioning of the 14 items is given in the supplementary data. High acceptability was indicated by low levels of missing data per item (range 0.0% to 1.7%). All items showed responses across the range of categories with positive inter-item correlations ranging from weak to strong (range 0.07 to 0.82, mean inter-item correlation 0.57), though the majority of correlations were strong (i.e. r>0.5).

The distribution of the MSK-HQ total score was approximately normally distributed with participants scoring across the entire range from 0 to 56. No floor or ceiling effect was observed. RA patients had the most favourable MSK-HQ total scores, with a mean score of 26.8 (SD=11.9) compared to 24.6 (13.1) for PsA patients and 23.2 (8.6) for those with other inflammatory arthritis conditions.

Structural validity
The first principal component of the item correlation matrix explained 62.8% of the total variance, indicating that a single underlying latent musculoskeletal health variable explained the majority of the variability in item responses. However, parallel analysis suggested the pattern of correlations between item responses was likely to result from two underlying latent variables. Specifically, an additional minor latent variable explained a further 9.5% of the item variance, which was greater than expected by chance. Examining inter-item and item-total correlations (supplementary materials) revealed that items 12 "understanding" and 13 "self-efficacy" were less related to overall musculoskeletal quality of life compared to other items and potentially reflect a separate construct.

The IRT graded response model confirmed this observation. Figure 1 plots the item characteristic curves for each response category separately for each item. These plots show the probability that each response category is endorsed given an individuals' level of the latent musculoskeletal health variable. The gradient of the item characteristic curve is defined by the items discrimination parameter, and the point at which the crosses 0.5 by its difficulty parameter. Transforming the discrimination parameters to the more familiar correlation metric indicates that the correlation between the response to each item and the latent musculoskeletal health variable was between r=0.75 and 0.91 for all variables, except for understanding (r=0.19) and self-efficacy (r=0.52). Further analyses considers the total scale score across all items and a reduced scale score excluding these two items.

Additional analysis considered differential item functioning by inflammatory arthritis subtype for those with RA and PsA. Due to low numbers those with AS and UIA were not included in this analysis. For the total score across all items, item 9 "sleep" (χ^2(2)=11.37, p=0.003) and item 14 "overall impact" (χ^2(2)=7.79, p=0.020) demonstrated non-uniform differential item functioning between those with RA and PsA. This was also observed for the reduced scale score. In both instances, responses to the items were better at discriminating between levels of musculoskeletal health in the PsA group than the RA group. The magnitude of bias for the "overall impact" item was small but more considerable for the "sleep" item.

Concurrent validity

The MSK-HQ was compared against five instruments assessing similar constructs as a test of construct validity (Figure 3). A strong positive correlation was observed with the EQ5D Index (n=285, r=0.80) and VAS (n=285, r=0.60) for all patients. Strong negative correlations were observed with the HAQ for all participants (n=284, r=-0.81), the RAID for those with RA or UIA (n=218, r=-0.88), the PsAID for those with PsA (n=43, r=-0.88), and the ASQoL for those with AS (n=22, r=-0.86). This supports the assumption that the MSK-HQ total score measures a highly related latent variable to these measures.

There was some indication of non-linearity in the relationship between the MSK-HQ with the HAQ and EQ5D Index, which is suggestive of the HAQ and EQ5D Index being more sensitive at differentiating between those with extremely poor musculoskeletal health. However, the MSK-HQ performs better in differentiating between those with better function since there is no floor effect observed, as is present for HAQ.

Reliability

Reliability was high for the both the total scale score including all items and the reduced scale score excluding items 12 and 13: α=0.93 and α=0.95, respectively. The IRT construct of information is related to reliability and has the advantage that measurement precision can be assessed at different points across the range of the latent variable. The test information curve (Figure 2) demonstrates that reliability was high (>0.8) between around 3 standard deviations below and above the mean of the latent musculoskeletal health variable. Converted to the metric of the MSK-HQ, reliability is above 0.8 for people scoring between approximately 3 and 54, where the range 0 to 56. A reliable change above measurement error was estimated to be a change of 9 points or more for the full scale, and 7 points of more for the reduced scale, given its higher reliability and reduced range.

Test-retest reliability

Data at follow up were provide by 178 (67.9%) participants with RA, PsA or UIA. Of these, 143 (80.3%) had stable symptoms as indicated by differences in either RAID or PsAID scores less than the minimum clinically important difference. Within those with stable symptoms, the MSK-HQ total

score test-retest reliability had an intraclass correlation coefficient r_{ICC}=0.73 (95% CI: 0.67 to 0.81) and for the reduced scale score r_{ICC}=0.75 (95% CI: 0.65 to 0.80). Examining test-retest agreement at the item level, Kendall's concordance coefficient ranged from 0.54 to 0.81 across all items and was greater than 0.7 for all items except 12 "understanding" and 13 "self-efficacy".

Discussion

This study has confirmed that the MSK-HQ performs well as a PROM for assessing musculoskeletal health related quality of life within an inflammatory arthritis cohort, with acceptable psychometric properties. The MSK-HQ is simple to administer and has good acceptability with patients. The tool includes seven domains including pain, fatigue, physical function, symptom interference, sleep, self-efficacy and psychological well-being. These domains were selected through a process of extensive consultation with patients to ensure the measure captured items that matter most [5]. The total score is reached by summing up each individual item, making the tool also simple for clinicians to use. The result ranges from 0 (worst musculoskeletal health) to 56 (best). We estimate a statistically reliable change, where measurement error can be ruled out for 95% of cases, to be a change in the full scale score of 9 points or higher.

The MSK-HQ was highly related to other measures of musculoskeletal health in the sample, with correlations with the HAQ, EQ5D index, RAID PsAID, ASQoL all greater than 0.8. Compared to the HAQ, the MSK-HQ will perform better at differentiating between those where the impact of disease is relatively mild due to the lack of floor effect. Furthermore, while there have been attempts to improve the validity of the HAQ as a tool for capturing musculoskeletal health,[19,20] the HAQ predominantly considers functional limitation related to the upper limbs and as such assesses only one important facet of musculoskeletal health.

Given the high degree of content overlap relating to both the RAID and PsAID, the high correlations with the MSK-HQ and these instruments are not surprising. The RAID and PsAID perform well psychometrically,[8,10] and there is little reason to recommend the use of the MSK-HQ over these two recently and well developed PROMs in disease specific studies. In our view, each will adequately assess health related quality of life; defined by the impact of inflammatory arthritis on a range of lower level constructs including physical function, pain, fatigue sleep, and psychological well-being. The advantage of the MSK-HQ, over the RAID and PsAID, is that it can potentially be used as a universal questionnaire across musculoskeletal conditions. This not only allows for the comparison of disease burden across conditions but reduces the logistical burden on clinicians and researchers having to manage multiple PROMs; reducing the risk that incorrect PROMs are completed. Although the MSK-HQ, RAID and PsAID assess a broad health related quality of life construct, they are still specific to populations with a musculoskeletal disorder. In situations where health economic analysis requires the calculation of utilities (e.g. for comparison across conditions) generic health related quality of life measures, such as the EQ5D, will also need to be administered.

The psychometric properties of the MSK-HQ are important to consider. Whilst the overall analysis supports the pragmatic use of the total score as a single measure of musculoskeletal health, interrogation of the measure reveals some interesting, if not unexpected, results. Two of the 14 items, relating to understanding and self-efficacy (illness perception items), appear to behave differently to the other items: the relationship between overall musculoskeletal health and individual items is weaker for the illness perception items. In particular the understanding item was only weakly related to musculoskeletal health, indicating that an individuals reported level of understanding is not particularly related to the severity of their condition. Understanding, as a construct, reflects something distinct from, though related to musculoskeletal health. It is likely that

the understanding item assesses the concept of illness coherence from the self-regulation model of illness.[21] Similarly, the self-efficacy item is related to the concept of personal control within this theoretical framework. The self-regulation model of illness, which includes other important illness perceptions of consequences, chronicity, stability, and treatment control, describes the process whereby beliefs influence the individual's emotional response to their illness and their coping behaviours, such as adherence to treatment.[22,23] In inflammatory arthritis, patient's perceptions of their illness have been shown to be associated with treatment adherence, function, pain and mood even after controlling for underlying disease.[22,24–26]

The relevance of these psychometric characteristics depends upon how the tool is intended to be used. As a measure for a clinical trial or assessing the standard of care in routine practice where a high degree of reliability is desired, it might be prudent to analyse the 12-items that performed best independently (items 1-11, & 14). When using the tool in the clinic, the impact of the two items is unlikely to make meaningful difference. In contrast, having insight into an individual's illness perceptions has substantial clinical utility. Supporting people to develop understanding and self-efficacy is an important goal for health services to achieve given their relationship to self-management and outcomes. Furthermore, given the method of development of the MSK-HQ, the inclusion of illness perceptions highlights their importance to patients. The use of validated PROMs for assessing the broader range of illness perceptions alongside the MSK-HQ may provide additional information about the mechanisms by which the disease impacts musculoskeletal health.[27,28]

The study must be considered alongside its limitations. The sample was predominantly RA and PsA, and so the conclusions across the other disease areas needs further study. We found some indication of differential item functioning across inflammatory arthritis subtypes. While the level observed is unlikely to bias comparisons between different inflammatory arthritis subtypes, it is difficult to draw robust inferences given the size of the sub groups. There is a clear need for further study with considerably larger sample sizes. In addition, while we provide some information potentially relevant to sensitivity to change, in terms of a change that is greater than the expected natural variation in scores, more robust assessment is required to determine whether such a change is clinically meaningful to patients. Calculation of the minimum clinically important difference may reveal a change that is smaller than the statistically reliable change is meaningful.

In this study, we have provided further support for the validity and reliability of a new musculoskeletal health PROM in people with inflammatory arthritis. The tool compares well to existing PROMs, with the added advantage of lacking the floor effect of HAQ. The major advantage of the MSK-HQ is that it is not disease specific and has high content validity in rheumatological conditions. We believe that the MSK-HQ score will be of value in clinical rheumatology practice, providing a measure which can be used across disease areas.

Acknowledgements

Funding

References

1. Black N. Patient reported outcome measures could help transform healthcare. BMJ. 2013;346:f167.

2. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, D'Agostino M-A, et al. Developing Core Outcome Measurement Sets for Clinical Trials: OMERACT Filter 2.0. J Clin Epidemiol. 2014;67(7):745–53.

3. NHS Information Centre. National Patient Reported Outcomes Programme [Internet]. Available from: www.ic.nhs.uk/proms

4. Patrick DL, Burke LB, Powers JH, Scott JA, Rock EP, Dawisha S, et al. Patient-Reported Outcomes to Support Medical Product Labeling Claims: FDA Perspective. Value Heal. 2007;10:S125–37.

5. Hill JC, Kang S, Benedetto E, Myers H, Blackburn S, Smith S, et al. Development and initial cohort validation of the Arthritis Research UK Musculoskeletal Health Questionnaire (MSK-HQ) for use across musculoskeletal care pathways. BMJ Open. 2016;6(8):e012331.

6. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539–49.

7. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980;23(2):137–45.

8. Heiberg T, Austad C, Kvien TK, Uhlig T. Performance of the Rheumatoid Arthritis Impact of Disease (RAID) score in relation to other patient-reported outcomes in a register of patients with rheumatoid arthritis. Ann Rheum Dis. 2011;70(6):1080–2.

9. Dougados M, Brault Y, Logeart I, van der Heijde D, Gossec L, Kvien T. Defining cut-off values for disease activity states and improvement scores for patient-reported outcomes: the example of the Rheumatoid Arthritis Impact of Disease (RAID). Arthritis Res Ther. 2012;14(3):R129.

10. Gossec L, de Wit M, Kiltz U, Braun J, Kalyoncu U, Scrivo R, et al. A patient-derived and patient-reported outcome measure for assessing psoriatic arthritis: elaboration and preliminary validation of the Psoriatic Arthritis Impact of Disease (PsAID) questionnaire, a 13-country EULAR initiative. Ann Rheum Dis. 2014;73(6):1012–9.

11. Doward LC, Spoorenberg A, Cook SA, Whalley D, Helliwell PS, Kay LJ, et al. Development of the ASQoL: a quality of life instrument specific to ankylosing spondylitis. Ann Rheum Dis. 2003;62(1):20–6.

12. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol. 1994;21(12):2286–91.

13. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum. 2007;57(8):1358–62.

14. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods. 1999;4(3):272–99.

15. Holgado–Tello FP, Chacón–Moscoso S, Barbero–García I, Vila–Abad E. Polychoric versus

Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. Qual Quant. 2008;44(1):153–66.

16. Samejima F. ESTIMATION OF LATENT ABILITY USING A RESPONSE PATTERN OF GRADED SCORES1. ETS Res Bull Ser. 1968;1968(1):1–169.

17. Crane PK, Gibbons LE, Jolley L, van Belle G. Differential item functioning analysis with ordinal logistic regression techniques. DIFdetect and difwithpar. Med Care. 2006;44(11 Suppl 3):S115–23.

18. Christensen L, Mendoza J. A method of assessing change in a single subject: An alteration of the RC index. Behav Ther. 1986;17(3):305–8.

19. Pincus T, Sokka T, Kautiainen H. Further development of a physical function scale on a MDHAQ [corrected] for standard care of patients with rheumatic diseases. J Rheumatol. 2005;32(8):1432–9.

20. Wolfe F, Michaud K, Pincus T. Development and validation of the health assessment questionnaire II: A revised version of the health assessment questionnaire. Arthritis Rheum. 2004;50(10):3296–305.

21. Leventhal H, Brissette I, Leventhal EA. The common-sense model of self-regulation of health and illness. In: Cameron LD, Leventhal H, editors. The self-regaulation of health and illness behaviour. New York: Routledge; 2003.

22. Hale ED, Treharne GJ, Kitas GD. The common-sense model of self-regulation of health and illness: how can we use it to understand and respond to our patients' needs? Rheumatology (Oxford). 2007;46(6):904–6.

23. Hope HF, Bluett J, Barton A, Hyrich KL, Cordingley L, Verstappen SMM. Psychological factors predict adherence to methotrexate in rheumatoid arthritis; findings from a systematic review of rates, predictors and associations with patient-reported and clinical outcomes. RMD Open. 2016;2(1):e000171.

24. Norton S, Hughes LD, Chilcot J, Sacker A, van Os S, Young A, et al. Negative and positive illness representations of rheumatoid arthritis: a latent profile analysis. J Behav Med. 2013;37(3):524–32.

25. van Os S, Norton S, Hughes LD, Chilcot J. Illness perceptions account for variation in positive outlook as well as psychological distress in Rheumatoid Arthritis. Psychol Health Med. 2011;17(4):37–41.

26. Groarke AM, Curtis R, Coughlan R, Gsel A. The impact of illness representations and disease activity on adjustment in women with rheumatoid arthritis: A longitudinal study. Psychol Health. 2005;20(5):597–613.

27. Moss-Morris R, Weinman J, Petrie KJ, Horne R, Cameron L, Buick D. The revised illness perception questionnaire (IPQ-R). Psychol Health. 2002;17(1):1–16.

28. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. J Psychosom Res. 2006;60(6):631–7.

Tables

Table 1. Sample characteristics

	All patients n=287	RA n=208	PsA n=45	AS n=22	UIA n=12
Age, mean (SD)	55 (18)	56 (17)	52 (19)	47 (14)	58 (18)
Female, n (%)	191 (67)	148 (71)	22 (49)	13 (59)	8 (75)
Years since diagnosis, median (IQR)	1 (0,7)	1 (0,7)	1 (0,6)	2 (0,18)	1 (0,11)
HAQ, mean (SD)	0.80 (0.59)	0.79 (0.61)	0.82 (0.58)	0.95 (0.48)	0.55 (0.41)
MSK-HQ, mean (SD)	27 (12)	27 (12)	25 (13)	21 (7)	29 (11)

RA=rheumatoid arthritis; PsA=psoriatic arthritis; AS=ankylosing spondylitis; UIA=undifferentiated arthritis; SD=standard deviation; IQR=interquartile range; HAQ=health assessment questionnaire
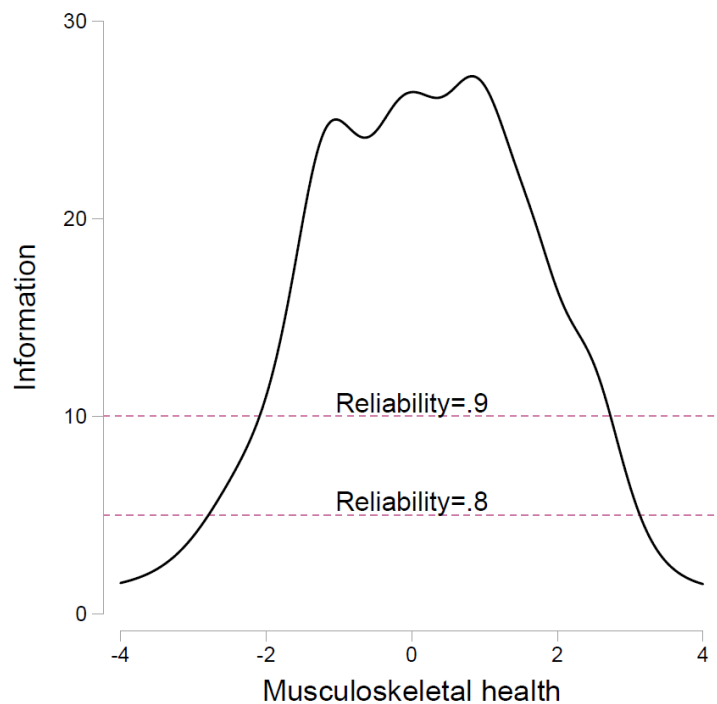
Figures

Figure 1. Test information function for the MSK-HQ

Test information (y-axis) relates to the precision of the scale (i.e. reliability) given the level of latent musculoskeletal health variable (x-axis). Dashed lines indicate levels of information where reliability on the familiar metric (e.g. that of Chronbach's α) is 0.8 and 0.9. Reliability is high from around 3 standard deviations below to 3 standard deviations above the mean of the latent musculoskeletal health variable.
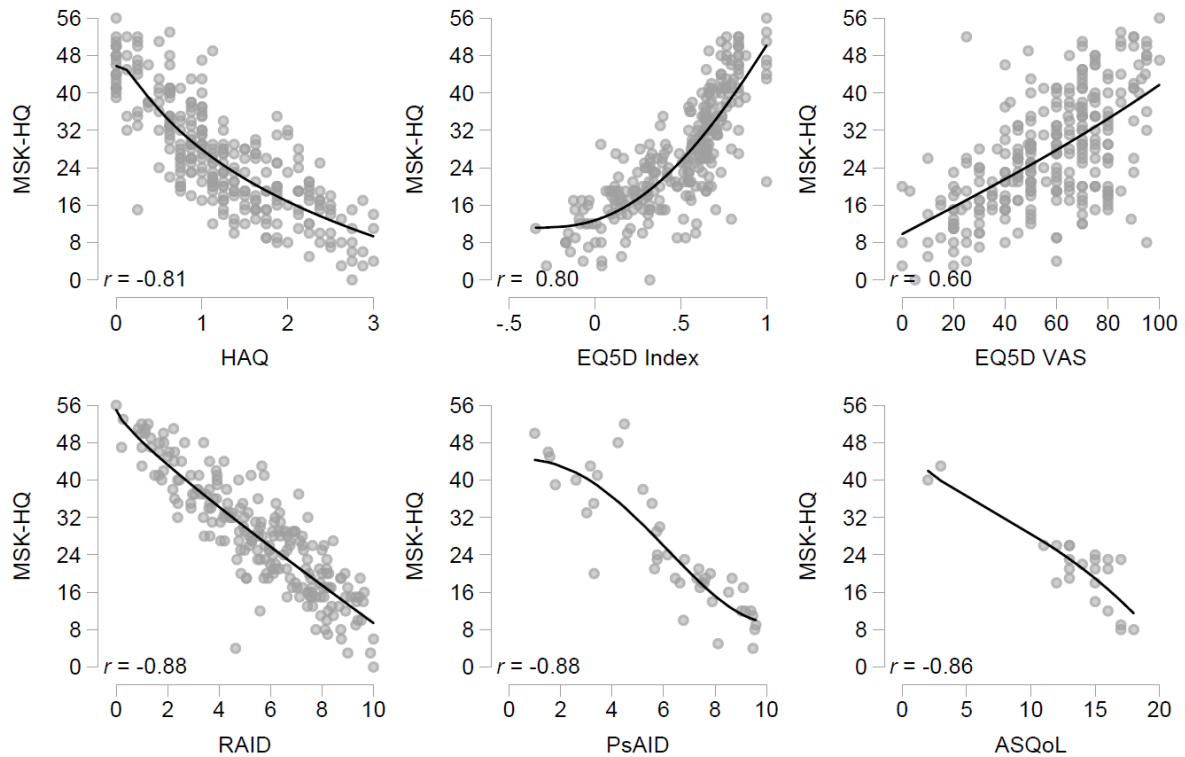
Figure 2. Concurrent validity of the MSK-HQ against other PROMs

Scatterplots indicating strong negative associations between the MSK-HQ total score and scores for other general and disease specific patient reported outcome measures (PROMs): Health Assessment Questionnaire (HAQ) and EQ5D in all patients, Rheumatoid Arthritis Impact of Disease (RAID) in RA patients, Psoriatic Arthritis Impact of Disease (PsAID) in PsA patients, and Ankylosing Spondylitis Quality of Life Questionnaire (ASQoL) in AS patients. Lines are fractional polynomial regression of MSK-HQ on the other PROM as an indicator of potential non-linearity in the association.