**Main Title:** Validation of the Musculoskeletal Health Questionnaire (MSK-HQ) in Primary Care Patients with Musculoskeletal Pain.

**Short Title**: MSK-HQ Validation in Primary Care.

**Authors:** Dr Ian C Scott (MBChB, MSc, PhD)1,2,3, Dr Gareth McCray (MA, MRes, PhD)1, Prof Gillian Lancaster (MSc, PhD)1,2, Prof Nadine E Foster (BSc [Hons], DPhil)1, Dr Jonathan C Hill (BSc, MSc, PhD)1

**Affiliations:**

1. Primary Care Centre Versus Arthritis, School of Primary, Community and Social Care, Keele University, Keele, UK.
2. Clinical Trials Unit, School of Primary, Community and Social Care, Keele University, Keele, UK.
3. Haywood Academic Rheumatology Centre, Haywood Hospital, Midlands Partnership NHS Foundation Trust, High Lane, Burslem, Staffordshire, UK.

**Abstract Word Count:** 244

**Manuscript Word Count:** 4,416

**Corresponding Author:** Dr Ian Scott; i.scott@keele.ac.uk; ORCHID ID: 0000-0002-1268-9808; telephone 01782 733905; fax 01782 734719.

**ABSTRACT**

**Objective**

To evaluate the responsiveness, and concurrent validity of the Musculoskeletal Health Questionnaire (MSK-HQ) in UK primary care patients with common musculoskeletal (MSK) pain presentations.

**Methods**

A secondary analysis of a primary care pilot randomised trial (STarT MSK) was performed. In 524 people consulting with back, neck, shoulder, knee, or multi-site pain, the following were recorded at 0/6 months: MSK-HQ, EQ-5D-5L, Roland-Morris Disability Questionnaire (RMDQ; back pain), Neck Disability Index (NDI), Shoulder Pain and Disability Index (SPADI), Knee Injury and Outcome Score (KOOS), Short-Form-12 (SF-12; multisite pain). At 6-months, patients self-rated their global change in MSK condition, from -5 (“very much worse”) to +5 (“completely recovered”). Receiver operating characteristic curves evaluated abilities of 6-month changes in each patient reported outcome measure (PROM) to discriminate between patients improving/not improving on global change scores, with Minimal Clinically Important Differences (MCID) calculated.

**Results**

The MSK-HQ had a good ability to discriminate between MSK pain patients reporting global improvement vs. no improvement (area under the curve [AUC] 0.81; 95% CI 0.78, 0.85). Its discriminative ability was higher than the EQ-5D-5L (AUC 0.68; 95% CI 0.62, 0.73) and similar to site-specific PROMs. The MCID for the 6-month change in MSK-HQ was 5.5. The MSK-HQ had strong correlations with all PROMs, except SF-12 scores.

**Conclusion**

In primary care patients with common MSK pain presentations, the MSK-HQ was as good as existing pain-site specific PROMs at identifying people reporting global improvements in their MSK condition, and was better than the EQ-5D-5L.

**Key Words**

Musculoskeletal, patient reported outcome measures, primary care, quality of life.

1. **INTRODUCTION**

Musculoskeletal (MSK) pain is a major global health problem. Each year, ~14% of UK adults consult in primary care for MSK conditions (1); in Sweden, 24% of adults report chronic regional MSK pain (2); in the USA, 1 in every 2 people report MSK conditions (3); and in middle and low income countries, MSK pain accounts for 19% of all healthy years of life lost due to disability (4). Preventing and managing MSK pain are, therefore, international priorities.

A crucial barrier to delivering high-quality healthcare to patients with MSK pain is the absence of a standardised approach to assessing clinical outcomes. It is widely agreed that the impacts of MSK pain on patients’ lives are best captured using patient-reported outcome measures (PROMs), often captured through patient-completed questionnaires (5). Despite this, marked variation exists in PROMs used to assess patients with MSK pain, which differ by geography (across regions/countries), clinical setting (primary/secondary care), and MSK pain sites.

To address this, the UK charity, Versus Arthritis, funded a collaboration between the Universities of Oxford and Keele to develop the Versus Arthritis Musculoskeletal Health Questionnaire (MSK-HQ) (6). A generic MSK PROM was co-produced and designed to measure overall MSK health status in patients with MSK conditions (6). It comprises 14-items, including pain, fatigue, physical function, symptom interference, social activities, sleep, self-efficacy, and psychological well-being. It is anticipated the MSK-HQ will act as a standard PROM across the MSK clinical care pathway, enabling people to report key issues to clinicians, facilitating shared decision-making, and supporting quality improvement (6).

In research to date, the MSK-HQ has been evaluated in patients with MSK pain in two settings. First, in patients with hip, shoulder, and knee pain listed for orthopedic surgery in secondary care (6,7). Second, in patients with a range of MSK problems seen in community physiotherapy clinics (6,7). In these contexts, high completion rates, excellent test-retest reliability, strong convergent validity with existing PROMs, and responsiveness to change were observed. It has not yet been evaluated in people with MSK pain consulting in UK primary care, where the vast majority of NHS MSK-related health contacts occur. Furthermore, its concurrent validity against accepted “reference” standards for MSK pain has not been assessed. Evaluating these issues prior to more widespread use of the MSK-HQ is needed.

We addressed these knowledge gaps, through secondary analysis of the STarT MSK pilot randomised trial, enrolling 524 patients consulting with one of the five commonest MSK pain presentations (low back, neck, shoulder, knee, and multi-site pain) in 8 general practices (8). We evaluated the MSK-HQ’s responsiveness (extent to which it detects changes over time that matter to patients (9)) using patients’ own global assessments of changes in their MSK condition as the reference criterion. This is a widely-used approach in psychometric assessments of MSK outcome measures (10). We also evaluated its concurrent validity (extent to which it correlates with previously validated MSK site-specific reference measures) (11–15).

1. **PATIENTS AND METHODS**

**2.1 Patients**

*2.1.1 STarT MSK Pilot Randomised Controlled Trial (RCT)*

This was conducted in 8 primary care practices. It examined the feasibility of a future main, cluster RCT comparing stratified primary care for patients consulting with one of the five commonest MSK pain presentations, compared with usual care (8). Stratified care comprised: (1) prognostic assessment to allocate patients into low-, medium- or high-risk strata (using the Keele STarT MSK tool (16)); and (2) recommended matched treatment options in each strata. Four practices were randomised to offer stratified care and four to continue usual care.

*2.1.2 Inclusion/Exclusion Criteria*

Full details are provided in the trial’s primary publication (8). In brief, patients who were included were ≥18 years old, and registered at a participating practice. During their consultation the clinician was asked to complete a recruitment template on the computer system and confirm if the patient was consulting with low back, neck, shoulder, knee or multi-site MSK pain. Patients were excluded if they had serious pathology requiring urgent attention, were unable to communicate in English, or were considered vulnerable (e.g. had a diagnosis of dementia/terminal illness).

*2.1.3 Identification/Assessment*

Patients were identified by weekly sceening of electronic medical records using appropriate Read codes (coded clinical terms). Eligible patients were sent invitation letters from their general practice (containing a participant information sheet, and baseline questionnaire including the MSK-HQ, EQ-5D-5L, a pain intensity measure [11-point numeric rating scale] and pain-site specific PROM focusing on function) shortly after their consultation. The pain-site specific PROM depended on which pain-site the clinician recorded as the primary reason for their visit. Patients consenting to data collection were sent a questionnaire at 6-months follow-up, capturing the same PROMs alongside a global rating of change in their MSK condition.

*2.1.4 Pooled Analysis*

Analysis of the results from the pilot RCT primary analysis focused on feasibility outcomes (e.g. recruitment, use of stratified care), alongside describing patients’ clinical outcomes over 6 months. (8). For this secondary analysis, the two intervention arms were pooled to make one dataset.

**2.2 Patient Reported Outcome Measures (PROMs)**

*2.2.1 All Participants*

All participants completed the MSK-HQ and EQ-5D-5L. The MSK-HQ uses 14 items to capture domains prioritised as important by patients and clinicians (6). These domains span pain, mobility, physical activity, sleep, social interaction, work/daily routine, independence, understanding of condition/treatment, confidence at managing symptoms, washing/dressing, fatigue, and overall impact. Each item has a 5-point scale (coded from ‘not at all’ [4 points] to ‘extremely’ [0 points], except for items 12 and 13, which have response options in the reverse order). The total summed score ranges from 0 to 56, with 56 being the best possible MSK health state. The EQ-5D-5L is a widely-used, generic health-related quality of life (HRQoL) measure, capturing information from five domains (mobility, self-care, usual activities, pain/discomfort, anxiety/depression). These can be combined in a summary index score; higher scores indicate better HRQoL (17). At 6-months, all patients were asked to provide a global rating of change in their MSK condition, by answering the following question on a scale of -5 (“very much worse”) to +5 (“completely recovered”): “With respect to your [pain site inserted] pain, how would you describe yourself now compared to how it was when you saw your doctor around 6 months ago?” This provided an 11-point scale ranging from -5 to 5, as recommended by Kamper et al (10) with “0” representing “no change”, and values greater than and less than “0” representing an improvement, and deterioration, respectively.

*2.2.2 Low Back Pain*

Patients with low back pain were asked to complete the Roland and Morris Disability Questionnaire (RMDQ). This includes 24-items likely to be affected by low back pain, summed to provide a score ranging from 0 (no disability) to 24 (maximum disability). It has been extensively validated in patients with low back pain across many settings (12). The minimal clinically importance difference (MCID) in RMDQ scores has been reported in a primary care population as a reduction of ≥30% from baseline score at 6 months, alongside patients rating their back pain as ‘better’ on a global rating scale (18). Other studies have reported the MCID as being a change score of ~5 (19–21).

*2.2.3 Neck Pain*

Patients with neck pain were asked to complete the Neck Disability Index (NDI) (13). This includes 10-items covering issues such as pain intensity, headache, lifting, and work. Each item is measured on a scale from 0 (no pain/disability) to 5 (maximum pain/disability), with a total score generated ranging from 0 (no disability) to 100 (maximum disability) (22). The NDI is the most widely used and validated instrument for assessing self-rated disability in patients with neck pain (23). Various MCIDs have been reported across studies (24). In patients with non-specific chronic neck pain, the MCID has been estimated at 3.5 (25); in patients with cervical radiculopathy it has been estimated at 7 (26).

*2.2.4 Shoulder Pain*

Patients with shoulder pain completed the Shoulder Pain and Disability Index (SPADI). This includes two subscales (pain and functional activities), combined to give a total score out of 100, with higher scores indicating greater shoulder-related disability (27). The SPADI has been validated across a range of shoulder problems and clinical settings, including those consulting in primary care, and those having shoulder surgery (27). The MCID has been estimated at 8-13 (28).

*2.2.5 Knee Pain*

Patients with knee pain completed the Knee injury and Osteoarthritis Outcome Score (KOOS). This includes five patient subscales, combined in a summative score ranging from 0 (extreme knee problems) to 100 (no knee problems). It has been extensively validated in patients with knee pain attributable to many causes (29). The KOOS MCID is most commonly considered to be 8-10 (14).

*2.2.6 Multi-Site Pain*

Patients with multi-site pain completed the short-form version 12 (SF-12) (30). This measures eight health domains (four physical; four mental), each scored 0 to 100, with higher scores indicating better health. Health domains can be combined in a physical component summary score (PCS), also ranging 0 to 100. The SF-12 is commonly used in studies of multi-site pain (31,32). The MCID for the domain and component summary scores are unknown for patients with multi-site MSK pain, although for patients with subacute and chronic low back pain, the MCID for the PCS is estimated to be 3.29 (15).

**2.3 Statistical Analysis**

*2.3.1 Missing Data*

Pre-analysis, missing MSK-HQ item data were imputed for participants with a small (n<3) number of missing item responses. At baseline, 14 participant questionnaires had 1 missing item response, 1 had 2 missing item responses, and 2 had 7 missing item responses. At 6-months, 13 participant questionnaires had 1 missing item response, 3 had 2 missing item responses, and 1 had 3 missing item responses. Missing item values were replaced with the mean item score for the individual in whom they were missing.

*2.3.2 Responsiveness*

To evalute the responsiveness of the MSK-HQ, we used the patient global assessment of change in MSK condition as an external anchor transition instrument, and assessed the ability of the MSK-HQ to discriminate between patients reporting and not reporting an improvement in their MSK condition. Global rating of change scales offer efficient methods of charting self-assessed progress in research/clinical settings, demonstrating clinical relevance, adequate reproducibility, and sensitivity to change (10). We dichotomised the patient global assessment score into those reporting any improvement in their MSK condition versus those reporting no change/worsening symptoms. Receiver operating characteristic (ROC) curves were constructed, and the area under the curve (AUC) calculated for 6-month changes in MSK-HQ, EQ-5D-5L, and site-specific disability PROM scores at discriminating between patients reporting and not reporting an improvement in their MSK condition.

We also assessed the ability of the MSK-HQ and EQ-5D-5L to identify patients reporting varying levels of global change in their MSK condition using two ordinal logistic regression models (proportional odds). Patient global assessment of change in MSK condition scores were grouped into four categories: -5 to -2 representing “worse”; -1 to 1 representing “similar”; 2 to 3 representing “improved”; 4 to 5 representing “much improved”. These groupings ensured sufficient patient numbers in each category. Predicted cut-offs for 6-month changes in MSK-HQ and EQ-5D-5L scores that assigned patients to each global change category were calculated, with 95% CIs constructed using bootstrap resampling (33). Agreement between known and predicted score cut-offs were examined using Kappa statistics with quadratic weights. This analysis was only conducted on the full sample, due to insufficient numbers in global change categories on stratifying by MSK pain-site.

*2.3.3 Calculating Minimal Clinically Important Differences*

The MCID is broadly defined as a difference in a score large enough to have an implication for patient management (34). No consensus exists on the best methodology to calculate an MCID. We used the commonest anchor-based approach to calculate the MCID for the MSK-HQ, EQ-5D-5L, and each of the site-specific PROMs (35). This involved choosing the ROC curve point for PROM change scores optimising the sensitivity and specificity of discriminating between patients with improved and unchanged/worsening patient global assessment of MSK condition scores. As MCID values can be affected by baseline scores, a sensitivity analysis used the 6-month PROM change score, expressed as a percentage of the baseline score (36). As recommended by Revicki et al (37), we calculated correlation coefficients between 6-month changes in MSK-HQ scores, and patient global assessment of change in MSK condition scores, to confirm “non-trivial” associations.

*2.3.4 Concurrent Validity*

Concurrent validity was assessed by evaluating Spearman’s correlations between 6-month change scores in the MSK-HQ, and the EQ-5D-5L, and site-specific disability PROMs.

*2.3.5 Programme Used*

All analyses were conducted in R, version 3.4.1.

**2.4 Ethical Approval**

All participants provided informed consent for data collection. Research ethical approval was obtained for the STarT MSK pilot trial from the East Midlands Research Ethical Committee, Nottingham (Ref:16/EM/0257).

**2.5 Trial Registration**

The STarT MSK pilot trial was registered on the ISRCTN registry (reference number 15366334).

1. **RESULTS**

**3.1 Participant Characteristics**

Of 524 participants at baseline, most had low back (n=155), knee (n=144), or shoulder (n=124) pain, with fewer consulting for neck (n=59) or multi-site (n=42) pain (Table 1). Mean age of participants was 61.1 years, with similar mean ages reported by participants with different pain conditions (58.1-64.5 years). Overall, 39% of participants were male. Most reported that their current episode of MSK pain had lasted for less than 12 months, apart from those reporting multi-site pain, in whom 69% reported pain lasting for >1 year, and 19% pain for >10 years.

**3.2 Pain Scores**

Overall, the mean baseline MSK pain score was 6.21 (out of 10), reducing to 4.07 at 6 months follow-up (Table 1). Mean baseline, 6-month, and 6-month changes in pain scores were similar in patients reporting different MSK pain sites, with the exception of those with knee and multi-site pain, who reported smaller 6-month improvements of only 1.53 and 1.14, respectively. Many (60%) patients reported global improvements in their MSK condition. The proportion of patients reporting global improvements in neck (71%) and shoulder (70%) MSK pain were higher than those reporting knee (50%), low back (58%), or multi-site (54%) MSK pain.

* 1. **Generic PROM Scores**

Overall, the mean baseline MSK-HQ score was 29.36 (potential maximum of 56), with a mean increase (representing improvement) by 6 months of 7.46, resulting in a mean 6-month score of 37.35 (Table 1). Similar mean baseline MSK-HQ scores were seen in participants with MSK pain in different sites (28.12-32.19). Greater mean increases in MSK-HQ scores were observed in those with neck (9.41), low back (9.01), and shoulder (7.81) pain than knee (5.58) or multi-site (4.38) pain. Similar patterns were observed for EQ-5D-5L.

**3.4 Responsiveness**

The MSK-HQ had a moderate ability to discriminate between participants reporting a global assessment of resolution/improvement versus no change/worsening at 6 months in their MSK condition (Table 2; Figure 1). The AUC for the absolute 6-month change in MSK-HQ scores in all participants was 0.81 (95% CI 0.78, 0.85). An AUC of 0.70 is considered sufficient in this context (38). The lowest AUC was observed for neck (0.75; 95% CI 0.59, 0.81) and highest for multi-site (0.92; 95% CI 0.82, 1.00) pain, although the small number of participants in the latter group (n=35) limited the estimate’s precision. Similar AUC scores were observed for the relative 6-month change in the MSK-HQ.

The MSK-HQ had higher AUC values than the EQ-5D-5L. The AUC for the absolute 6-month change in EQ-5D-5L in all participants was 0.68 (95% CI 0.62, 0.73). The lowest AUC was observed for neck (0.49; 95% CI 0.32, 0.66) and highest for shoulder (0.72; 95% CI 0.60, 0.72) pain. Similar AUC scores were observed for the relative 6-month change in EQ-5D-5L.

The MSK-HQ had similar AUC values to reference standard PROMs measuring disability at discriminating between those reporting resolution/improvement versus no change/worsening (Table 2; Figure 1). For low back pain, AUC values for the absolute 6-month change scores in the RMDQ and MSK-HQ were 0.74 (95% CI 0.65, 0.82) and 0.81 (95% CI 0.74, 0.88), respectively. For neck pain, AUC values for the absolute 6-month change scores in the NDI and MSK-HQ were 0.74 (95% CI 0.59, 0.89) and 0.75 (95% CI 0.59, 0.81), respectively. For shoulder pain, AUC values for the absolute 6-month change scores in the SPADI and MSK-HQ were 0.74 (95% CI 0.64, 0.84) and 0.83 (95% CI 0.74, 0.91), respectively. For knee pain, AUC values for the absolute 6-month change scores in the KOOS and MSK-HQ were 0.80 (95% CI 0.72, 0.88) and 0.80 (95% CI 0.72, 0.87), respectively. Finally, for multi-site pain, AUC values for the absolute 6-month change scores in the SF-36 PCS, and MSK-HQ were 0.57 (95% CI 0.37, 0.68), and 0.92 (95% CI 0.82, 1.00), respectively. Similar findings were seen for relative 6-month PROM change scores. The EQ-5D-5L had lower AUC statistics than all reference standard PROMs, except the SF-12 PCS for multi-site pain (Table 2; Figure 1).

Using ordinal logistic regression models, predicted cut-offs in 6-month changes in MSK-HQ scores of -6.5 (95% CI -10.5, -3.5), 7.5 (95% CI 4.5, 10.5) and 17.5 (95% CI 13.5, 20.5) were optimal to identify patients that were worse/similar, similar/improved, and improved/much improved in their global assessment of change in MSK condition scores, respectively. For the EQ-5D-5L, these scores were -0.30 (95% CI -0.48, -0.16), 0.08 (95% CI -0.02, 0.19), and 0.26 (95% CI 0.16, 0.40), respectively. For the MSK-HQ, the Kappa statistic for agreement between the known and predicted score cut-offs for global change categories was 0.55 (95% CI 0.48, 0.61), with the lower 95% confidence interval indicating intermediate-to-good agreement (39). For the EQ-5D-5L, the Kappa statistic was 0.31 (95% CI 0.22, 0.40) with the lower CI indicating poor agreement.

**3.5 Minimal Clinical Important Differences (MCIDs)**

The 6-month change in MSK-HQ scores had a strong correlation with patient global assessment of change in MSK condition scores (r=0.60; 95% CI 0.53, 0.65) (Table 3). High correlations were also observed between patient global assessment of change scores, and the KOOS (r=0.59), NDI (r=-0.53), RMDQ (r=-0.43), and SPADI (-0.44), with weaker correlations observed with the EQ-5D-5L (r=0.37), and SF-36 PCS (r=0.15).

In all participants, the MCID in the absolute 6-month change score for the MSK-HQ was 5.5, which had a 72% sensitivity and 78% specificity for discriminating between patients reporting and not reporting resolution/improvement in their MSK condition (Table 2), replicating the MCID reported in another study (7). Some minor variation (within 2.0 units) was observed in the MCID in the MSK-HQ absolute 6-month change score across four of the five pain-sites. However, for multi-site pain the MCID was substantially lower at 0.50, although the limited sample size of this group makes the relevance of this finding uncertain.

The MCID for the relative 6-month change score for the MSK-HQ was 15.7% (Table 2), which had the same sensitivity as the absolute 6-month change score (72%), but slightly lower specificity (71%). There was some minor variation in the MCID for the relative 6-month change scores in three of the five pain-sites (ranging from 14.9-20.6%) but substantially lower scores were seen for shoulder (9.7%) and multi-site (2%) pain.

The MCIDs for the absolute 6-month change scores for site-specific PROMs were similar to those reported in the existing literature, with the exception of the SF-36 PCS, although this was a poor discriminator of those reporting or not reporting resolution/improvement. For knee pain, the KOOS MCID was 10.7; for neck pain, the NDI MCID was 3.5; for low back pain, the RMDQ MCID was 2.5; for shoulder pain, the SPADI MCID was 11.2; for multi-site pain, the SF-36 PCS MCID was 0.35.

**3.6 Concurrent Validity**

Strong positive correlations were observed between 6-month changes in the MSK-HQ and the EQ-5D-5L (r=0.59), and KOOS (r=0.61), and strong negative correlations between 6-month changes in the MSK-HQ and the NDI (r=-0.77), RMDQ (r=-0.69), and SPADI (r=-0.70) (Table 4). A weaker positive correlation was observed between 6-month changes in the MSK-HQ and the SF-36 PCS (r=0.17), which was not statistically significant. Correlations between 6-month changes in the EQ-5D-5L, and site-specific PROMs were of a smaller magnitude compared to the MSK-HQ for all PROMs except the SF-36 PCS.

1. **DISCUSSION**

Our study evaluated the responsiveness, and concurrent validity of the MSK-HQ in patients with MSK conditions managed in primary care. We found that, firstly, the MSK-HQ was able to discriminate well between patients reporting, and those not reporting global resolution/improvement in their MSK condition over 6 months, with AUC statistics ranging from 0.75-0.92 across the 5 pain-sites for absolute 6-month changes in MSK-HQ scores. Secondly, the MSK-HQ had similar capabilities to existing pain-site specific reference PROMs that measure disability, at identifying patients reporting global improvements in their MSK condition, with similar AUC statistics observed. As many patients in the primary care sector will have symptoms at more than one MSK pain site over time, its similar responsiveness to site-specific PROMs across different pain sites is particuarly advantagous. Thirdly, the MSK-HQ had higher AUC values than the EQ-5D-5L (widely-used, generic HRQoL measure) at identifying patients who reported global resolution/improvement in their MSK condition. Overall, these results provide further evidence that the MSK-HQ is a valid and responsive PROM for use in patients reporting a range of MSK pain presentations.

Our finding that the MSK-HQ had higher AUC values at identifying patients self-reporting a global improvement in their MSK condition, compared to the EQ-5D-5L index score, is of particular interest. The EQ-5D (encompassing its 3- and 5-level forms) is one of the most frequently used generic, preference-based HRQoL measures, with the EQ-5D-3L recommended by the National Institute for Health and Care Excellence for use in its technology appraisal economic evaluations (40). Responsiveness to detect clinical changes of importance to patients is a crucial property of a PROM; if it is insufficently sensitive, it will fail to detect treatment benefits, leading to imprecise evaluations. Payakachat et al recently systematically reviewed the ability of the EQ-5D to detect clinically meaningful changes in health status, reporting it was responsive in fewer than half of the 56 conditions evaluated across 145 studies, with mixed and no evidence of responsiveness seen in 48% and 7% of conditions, respectively (41). In patients with MSK conditions, the responsiveness of the EQ-5D has mainly been evaluated in patients having surgery, or those reviewed in orthopedic clinics. Several studies have assessed its internal responsiveness (ability to change over a specific time-period, usually pre-/post-surgery). In 60 patients undergoing carpal tunnel decompression surgery the EQ-5D-5L had “moderate” internal responsiveness (Cohen’s d score of 0.5 comparing mean post-operative and pre-operative EQ-5D-3L scores) (42). In 54,486 patients having knee replacement surgery, standardised response mean scores were 0.90 for EQ-5D-3L index scores, compared with 1.50 for Oxford Knee Scores (43). Limited data exists on its external responsiveness (extent to which its change relates to corresponding changes in a reference measure of clinical or health status), although Soer et al reported that in 151 patients with non-specific chronic lower back pain receiving treatment in a tertiary care spine centre, EQ-5D-3L index scores had an AUC of 0.71 at idenfifying patients reporting improvements in their global scores (44). The findings from our study suggest that in patients with common MSK pain presentations seen in primary care, the MSK-HQ is better at detecting clinically meaningful improvements in patients’ MSK conditions, and is a superior PROM than the EQ-5D in this setting.

We used the anchor-based method to calculate the MCID for the MSK-HQ, comparing changes in MSK-HQ scores with changes in a patient global assessment of MSK condition score in a ROC curve analysis. This is the most widely-used approach to calculating MCID (45), and analagous to the optimal cut-off when comparing scores at one time-point. We observed variability in the MCID for the MSK-HQ across the five pain-sites, ranging from 0.50-6.50. This is not a suprising finding, considering the small sample size of some of the MSK pain subgroups limiting the precision of our estimates (with only 35 patients in the multisite pain group having 6-month MSK-HQ change data). Variability of PROM MCIDs across clinical conditions and studies is a well-established challenge, consistently demonstrated in several systematic reviews (45,46). The reasons for this variability include heterogeneity in patient populations (varying by age, gender, and conditions), methodology (different global anchors and analysis methods), and study quality. Therefore, whilst it is accepted that the MCID is a useful and necessary tool to help interpret patient reported improvement, interpretation is needed that takes the range of available estimates into account (34,45). Whilst our overall MCID of 5.5 for the MSK-HQ is constistent with a previous study in a physiotherapy, and secondary care setting (7), further work is required to evaluate MCIDs for the MSK-HQ in different MSK patient populations, to provide a greater understanding of what changes would have implications for patients’ management in different healthcare settings, and MSK conditions.

In the START MSK Pilot RCT both the Keele STarT MSK Tool and MSK-HQ were used (8). These PROMs have important conceptual differences. The STarT MSK tool is a prognostic stratification tool, for use among patients with MSK pain to identify those at low, medium or high-risk of persistent MSK pain (16,47). It should be used to help clinicians understand patients’ likely prognoses, and guide initial treatment decisions e.g. early physiotherapy in those at high-risk. In contrast, the MSK-HQ is used to measure overall MSK health status using domains that patients and clinicians have prioritised as being important. Having now demonstrated its excellent responsiveness in the primary care setting, clinicians can be confident in using it as a standard MSK PROM to evaluate changes in MSK health following treatment in primary care.

The strengths of this secondary analysis include the sample size (>500 patients), inclusion of widely-accepted reference PROMs capturing disabilty related to different MSK pain-sites, and its primary care setting (since most patients with MSK pain are managed there). It also has several limitations. Firstly, participants were grouped into one of five common MSK pain presentations, rather than specific clinical diagnoses; clinical heterogeneity is likely to exist in the pathophysiology underlying pain at different sites, and the MSK-HQ could perform differently across specific diagnoses. Secondly, the sample size of some of the groups of participants (those with neck and multi-site pain) was small, and therefore the estimates were imprecise. Thirdly, whilst the SF-12 is frequently used in studies of multi-site pain, it was not developed for the purpose of evaluating multi-site pain specifically, and is therefore not analogous to the site-specific measures used for the other MSK-pain sites.

In conclusion, this analysis demonstrates that the MSK-HQ has good responsiveness, and concurrent validity in patients with the five most common MSK pain presentations consulting in primary care.

1. **FUNDING**

This paper presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research (PGfAR) Programme (Grant Reference Number RP-PG-1211-20010). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health. Nadine Foster is also NIHR Senior Investigator and was supported through an NIHR Research Professorship (NIHR-RP-011-015). Ian Scott receives funding from the NIHR West Midlands Clinical Research Network (Clinical Trials Scholarship). We also acknowledge Centre of Excellence funding from Versus Arthritis (Grant Reference Number 20202). The views expressed in this article are those of the authors and not necessarily those of the NHS, the NIHR, our funding bodies or the Department of Health and Social Care.

1. **CONFLICT OF INTEREST**

None declared.

**REFERENCES**

1. Jordan KP, Kadam UT, Hayward R, Porcheret M, Young C, Croft P. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. BMC Musculoskelet Disord. 2010;11:144.

2. Bergman S, Herrström P, Högström K, Petersson IF, Svensson B, Jacobsson LT. Chronic musculoskeletal pain, prevalence rates, and sociodemographic associations in a Swedish population study. J Rheumatol. 2001 Jun 1;28(6):1369 LP – 1377.

3. US Bone and Joint Initiative. The Burden of Musculoskeletal Diseases in the United States Available at: https://www.boneandjointburden.org/docs/BMUS Impact of MSK on Americans booklet\_4th Edition %282018%29.pdf. [Internet]. [Accessed 2019 May 26].

4. Hoy D, Geere J-A, Davatchi F, Meggitt B, Barrero LH. A time for action: Opportunities for preventing the growing burden and disability from musculoskeletal conditions in low- and middle-income countries. Best Pract Res Clin Rheumatol. 2014;28(3):377–93.

5. Food and Drug Administration (2009) Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. Available at: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. [Internet]. [Accessed 2018 Aug 22].

6. 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.

7. Price AJ, Ogollah R, Kang S, Hay E, Barker KL, Benedetto E, et al. Determining responsiveness and meaningful changes for the Musculoskeletal Health Questionnaire (MSK-HQ) for use across musculoskeletal care pathways. BMJ Open. 2019 Oct 1;9(10):e025357.

8. Hill JC, Garvin S, Chen Y, Cooper V, Wathall S, Saunders B, et al. Stratified primary care versus non-stratified care for musculoskeletal pain: findings from the STarT MSK feasibility and pilot cluster randomized controlled trial. BMC Fam Pract. In-Press (Accepted Dec 2019).

9. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess. 1998;2(14):i–74.

10. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther. 2009;17(3):163–70.

11. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. Arthritis Care Res. 1991;4(4):143–9.

12. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine (Phila Pa 1976). 2000 Dec;25(24):3115—3124.

13. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. J Manipulative Physiol Ther. 1991;14(7):409–15.

14. KOOS FAQs Available at: http://www.koos.nu/koosfaq.html. [Internet]. [Accessed 2019 May 26].

15. Diaz-Arribas MJ, Fernandez-Serrano M, Royuela A, Kovacs FM, Gallego-Izquierdo T, Ramos-Sanchez M, et al. Minimal Clinically Important Difference in Quality of Life for Patients With Low Back Pain. Spine (Phila Pa 1976). 2017;42(24):1908–16.

16. Campbell P, Hill JC, Protheroe J, Afolabi EK, Lewis M, Beardmore R, et al. Keele Aches and Pains Study protocol: validity, acceptability, and feasibility of the Keele STarT MSK tool for subgrouping musculoskeletal patients in primary care. J Pain Res. 2016;9:807–18.

17. EuroQol Group. EQ-5D-3L User Guide. Available at: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-3L\_UserGuide\_2015.pdf. [Internet]. [Accessed 2019 Mar 21].

 18. Jordan K, Dunn KM, Lewis M, Croft P. A minimal clinically important difference was derived for the Roland-Morris Disability Questionnaire for low back pain. J Clin Epidemiol. 2006 Jan 1;59(1):45–52.

19. Riddle DL, Stratford PW, Binkley JM. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 2. Phys Ther. 1998;78(11):1197–207.

20. Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. Phys Ther. 1998;78(11):1186–96.

21. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland-Morris questionnaire. Phys Ther. 1996;76(4):358–9.

22. Howell ER. The association between neck pain, the Neck Disability Index and cervical ranges of motion: a narrative review. J Can Chiropr Assoc. 2011;55(3):211–21.

23. Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. J Manipulative Physiol Ther. 2008;31(7):491–502.

24. MacDermid JC, Walton DM, Avery S, Blanchard A, Etruw E, McAlpine C, et al. Measurement Properties of the Neck Disability Index: A Systematic Review. J Orthop Sport Phys Ther. 2009 May 1;39(5):400–17.

25. Jorritsma W, Dijkstra PU, de Vries GE, Geertzen JHB, Reneman MF. Detecting relevant changes and responsiveness of Neck Pain and Disability Scale and Neck Disability Index. Eur Spine J. 2012/07/03. 2012 Dec;21(12):2550–7.

26. Cleland J, DPT OCS, Fritz J, PT P, Whitman J, PT Ds, et al. The Reliability and Construct Validity of the Neck Disability Index and Patient Specific Functional Scale in Patients With Cervical Radiculopathy. Spine (Phila Pa 1976). 2006;31(5):598–602.

27. Breckenridge JD, McAuley JH. Shoulder Pain and Disability Index (SPADI). J Physiother. 2011;57(3):197.

28. Roy J-S, MacDermid JC, Woodhouse LJ. Measuring shoulder function: A systematic review of four questionnaires. Arthritis Care Res (Hoboken). 2009 May 15;61(5):623–32.

29. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes. 2003 Nov 3;1:64.

30. Ware JJ, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996;34(3):220–33.

31. Finney A, Dziedzic KS, Lewis M, Healey E. Multisite peripheral joint pain: a cross-sectional study of prevalence and impact on general health, quality of life, pain intensity and consultation behaviour. BMC Musculoskelet Disord. 2017 Dec 16;18(1):535.

32. Thomas E, Belcher J, McBeth J, Lacey RJ, Wilkie R, Rathod T. Pain at multiple body sites and health-related quality of life in older adults: results from the North Staffordshire Osteoarthritis Project. Rheumatology. 2014 Jun 12;53(11):2071–9.

33. Venables WN, Ripley BD (2002). Modern Applied Statistics with S, Fourth edition. Springer, New York.

34. Sedaghat AR. Understanding the Minimal Clinically Important Difference (MCID) of Patient-Reported Outcome Measures. Otolaryngol Neck Surg. 2019 Jun 4;0194599819852604.

35. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4):407–15.

36. Olsen MF, Bjerre E, Hansen MD, Tendal B, Hilden J, Hrobjartsson A. Minimum clinically important differences in chronic pain vary considerably by baseline pain and methodological factors: systematic review of empirical studies. J Clin Epidemiol. 2018;101:87-106.e2.

37. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008;61(2):102–9.

38. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.

39. Fleiss, J. 1981. Statistical Methods for Rates and Proportions. John Wiley & Sons.

40. National Institute for Health and Care Excellence (NICE). Measuring and valuing effects for health interventions. Available at: https://www.nice.org.uk/process/pmg20/chapter/incorporating-economic-evaluation [Internet]. [Accessed 2019 Jul 22].

41. Payakachat N, Ali MM, Tilford JM. Can The EQ-5D Detect Meaningful Change? A Systematic Review. Pharmacoeconomics. 2015 Nov;33(11):1137–54.

42. Marti C, Hensler S, Herren DB, Niedermann K, Marks M. Measurement properties of the EuroQoL EQ-5D-5L to assess quality of life in patients undergoing carpal tunnel release. J Hand Surg. 2016 Jul 18;41(9):957–62.

43. Lin F-J, Samp J, Munoz A, Wong PS, Pickard AS. Evaluating change using patient-reported outcome measures in knee replacement: the complementary nature of the EQ-5D index and VAS scores. Eur J Health Econ. 2014;15(5):489–96.

44. Soer R, Reneman MF, Speijer BLGN, Coppes MH, Vroomen PCAJ. Clinimetric properties of the EuroQol-5D in patients with chronic low back pain. Spine J. 2012;12(11):1035–9.

45. Jayadevappa R, Cook R, Chhatre S. Minimal important difference to infer changes in health-related quality of life - a systematic review. J Clin Epidemiol. 2017 Sep 1;89:188–98.

46. Çelik D, Çoban Ö, Kılıçoğlu Ö. Minimal clinically important difference of commonly used hip-, knee-, foot-, and ankle-specific questionnaires: a systematic review. J Clin Epidemiol. 2019;113:44–57.

47. Keele University. Keele START MSK Tool. Available at: <https://www.keele.ac.uk/startmsk/>. [Internet]. 2017 [cited 2020 May 31].

**Table 1. Patient Characteristics and PROM Scores by Pain Site**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Characteristic** |  | **All (n=524)** | **Low Back (n=155)** | **Neck (n=59)** | **Shoulder (n=124)** | **Knee (n=144)** | **Multi-Site (n=42)** |
| Age (Years) |  | 61.14 (14.75) | 58.06 (15.57) | 58.75 (16.91) | 62.03 (13.86) | 64.49 (13.33) | 61.74 (13.33) |
| Male Gender, n (%) |  | 206 (39%) | 63 (41%) | 17 (29%) | 65 (45%) | 61 (42%) | 9 (21%) |
| Pain intensity on 11-point NRS | 0 months | 6.21 (2.25) | 6.43 (2.22) | 6.25 (2.23) | 5.94 (2.24) | 6.31 (2.33) | 5.86 (2.16) |
| 6 months | 4.07 (2.93) | 4.06 (2.93) | 3.02 (2.59) | 3.47 (2.93) | 4.83 (2.96) | 4.81 (2.42) |
| Change | -2.12 (2.97) | -2.38 (2.98) | -2.98 (2.79) | -2.37 (3.09) | -1.53 (2.96) | -1.14 (2.36) |
| Pain duration, n (%) | <3 months | 136 (26%) | 49 (30%) | 21 (36%) | 30 (24%) | 34 (24%) | 5 (12%) |
| 3-6 months | 77 (15%) | 19 (12%) | 14 (24%) | 26 (21%) | 14 (10%) | 4 (10%) |
| 7-12 months | 89 (17%) | 23 (15%) | 6 (10%) | 32 (26%) | 24 (17%) | 4 (10%) |
| 1-2 years | 75 (14%) | 16 (10%) | 5 (8%) | 19 (15%) | 26 (18%) | 9 (21%) |
| 3-5 years | 53 (10%) | 15 (10%) | 5 (8%) | 7 (6%) | 18 (12%) | 8 (19%) |
| 6-10 years | 48 (9%) | 15 (10%) | 4 (7%) | 6 (5%) | 19 (13%) | 4 (10%) |
| >10 years | 46 (9%) | 21 (14%) | 4 (7%) | 4 (3%) | 9 (6%) | 8 (19%) |
| Patient global change in MSK condition, n (%) | Worse | 102 (22%) | 30 (22%) | 9 (18%) | 18 (17%) | 38 (30%) | 7 (20%) |
| Same | 81 (18%) | 29 (20%) | 6 (12%) | 14 (13%) | 25 (20%) | 9 (26%) |
| Better | 273 (60%) | 79 (58%) | 36 (71%) | 76 (70%) | 63 (50%) | 19 (54%) |
| MSK-HQ | 0 months | 29.36 (10.41) | 28.14 (10.54) | 32.19 (10.19) | 31.24 (9.91) | 28.27 (10.73) | 28.12 (9.32) |
| 6 months | 37.35 (12.24) | 37.52 (12.76) | 43.14 (10.09) | 40.00 (11.12) | 34.06 (12.51) | 32.26 (10.17) |
| Change | 7.46 (10.8) | 9.01 (11.67) | 9.41 (9.65) | 7.81 (10.71) | 5.58 (10.77) | 4.38 (7.55) |
| EQ-5D-5L | 0 months | 0.56 (0.24) | 0.55 (0.24) | 0.64 (0.23) | 0.57 (0.23) | 0.53 (0.24) | 0.55 (0.23) |
| 6 months | 0.64 (0.26) | 0.66 (0.27) | 0.69 (0.21) | 0.66 (0.26) | 0.59 (0.25) | 0.59 (0.25) |
| Change | 0.08 (0.21) | 0.10 (0.23) | 0.05 (0.14) | 0.08 (0.21) | 0.07 (0.23) | 0.05 (0.13) |
| Site-Specific PROM | 0 months | - | 9.59 (5.5) | 16.08 (8.02) | 53.57 (21.84) | 42.88 (21.17) | 34.41 (9.52) |
| 6 months | - | 6.65 (6.11) | 9.73 (9.14) | 37.88 (27.36) | 52.79 (23.56) | 35.59 (12.21) |
| Change | - | -2.76 (5.25) | -6.29 (6.74) | -14.97 (27.74) | 9.40 (19.21) | 1.28 (8.33) |

All data are given as mean (SD) unless otherwise stated; MSK=musculoskeletal; NRS=numeric rating scale; site-specific PROMs are the Roland and Morris Disability Questionnaire for low back pain, Neck Disability Index for neck pain, Shoulder Pain and Disability Index for shoulder pain, Knee Injury and Osteoarthritis Outcome Score for knee pain, and Short-Form-12 Physical Component Summary Score (PCS) for multi-site pain.

**Table 2. Minimal Clinically Important Differences for the MSK-HQ Compared with Other Patient Reported Outcome Measures**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Absolute 6-Month Change Scores** | **Relative 6-Month Change Scores** |
| **No. Reporting Global Improvement** | **AUC****(95% CI)** | **MCID** | **Sensitivity/****Specificity** | **% Attained MCID** | **AUC****(95% CI)** | **MCID** | **Sensitivity/****Specificity** | **% Attained MCID** |
| **All Patients** |  |  |  |  |  |  |  |  |  |
| MSK-HQ(n=453) | 272 | 0.81(0.78, 0.85) | 5.50 | 0.72/0.78 | 52% | 0.78(0.73, 0.82) | 15.71% | 0.72/0.71 | 55% |
| EQ-5D-5L (n=437) | 271 | 0.68(0.62, 0.73) | 0.03 | 0.69/0.61 | 57% | 0.66(0.61, 0.71) | 3.18% | 0.68/0.61 | 57% |
| **Low Back Pain** |  |  |  |  |  |  |  |  |  |
| MSK-HQ(n=134) | 78 | 0.81(0.74, 0.88) | 6.50 | 0.76/0.73 | 55% | 0.75(0.67, 0.78) | 15.76% | 0.78/0.64 | 60% |
| EQ-5D-5L (n=128) | 79 | 0.69(0.60, 0.79) | 0.03 | 0.73/0.58 | 61% | 0.70(0.60, 0.70) | 3.03% | 0.74/0.64 | 59% |
| RMDQ(n=136) | 79 | 0.74(0.65, 0.82) | -2.50 | 0.63/0.74 | 52% | 0.79(0.71, 0.87) | -24.04% | 0.79/0.75 | 44% |
| **Neck Pain** |  |  |  |  |  |  |  |  |  |
| MSK-HQ(n=51) | 36 | 0.75(0.59, 0.81) | 4.50 | 0.81/0.67 | 67% | 0.70(0.52, 0.87) | 14.91% | 0.75/0.67 | 63% |
| EQ-5D-5L(n=50) | 36 | 0.49(0.32, 0.66) | 0.04 | 0.54/0.53 | 52% | 0.50(0.31, 0.68) | 4.90% | 0.54/0.60 | 50% |
| NDI(n=51) | 36 | 0.74(0.59, 0.89) | -3.50 | 0.69/0.67 | 41% | 0.80(0.68, 0.93) | -22.29% | 0.80/0.80 | 38% |
| **Shoulder Pain** |  |  |  |  |  |  |  |  |  |
| MSK-HQ(n=107) | 74 | 0.83(0.74, 0.91) | 3.50 | 0.79/0.81 | 62% | 0.82(0.73, 0.91) | 9.74% | 0.79/0.77 | 63% |
| EQ-5D-5L (n=103) | 76 | 0.72(0.60, 0.72) | 0.02 | 0.72/0.68 | 62% | 0.69(0.57, 0.81) | 2.15% | 0.71/0.61 | 62% |
| SPADI(n=105) | 76 | 0.74(0.64, 0.84) | -11.16 | 0.65/0.74 | 47% | 0.74(0.64, 0.84) | -29.91% | 0.62/0.81 | 50% |
| **Knee Pain** |  |  |  |  |  |  |  |  |  |
| MSK-HQ(n=126) | 63 | 0.80(0.72, 0.87) | 5.50 | 0.62/0.81 | 40% | 0.76(0.68, 0.84) | 20.61% | 0.62/0.76 | 43% |
| EQ-5D-5L (n=122)  | 62 | 0.71(0.62, 0.80) | 0.03 | 0.66/0.67 | 50% | 0.65(0.56, 0.75) | 4.15% | 0.62/0.64 | 50% |
| KOOS(n=123) | 60 | 0.80(0.72, 0.88) | 10.72 | 0.63/0.90 | 36% | 0.74(0.66, 0.83) | 31.68% | 0.58/0.84 | 37% |
| **Multi-Site Pain** |  |  |  |  |  |  |  |  |  |
| MSK-HQ(n=35) | 19 | 0.92(0.82, 1.00) | 0.50 | 0.81/0.95 | 60% | 0.9(0.79, 1.00) | 2.00% | 0.95/0.81 | 60% |
| EQ-5D-5L(n=34) | 19 | 0.61(0.61, 0.81) | 0.05 | 0.61/0.69 | 47% | 0.66(0.47, 0.86) | 6.44% | 0.67/0.69 | 53% |
| SF-12 PCS (n=35) | 19 | 0.57(0.37, 0.68) | 0.35 | 0.68/0.50 | 60% | 0.59(0.39, 0.78) | 0.97% | 0.68/0.50 | 60% |

No./n=number; No. reporting global improvement = number of patients with available data reporting improvement in their patient global score at 6 months; MCID = Minimal Clinically Important Difference; MCID calculated using anchor-based method; sensitivity and specificity given for MCID at differentiating between patients with and without an improvement in their patient global score at 6 months (health transition anchor); relative 6-month change score is the change in the PROM score as a proportion of the baseline score; % attained MCID = the proportion of patients attaining the MCID.

**Table 3. Correlations between 6-Month Changes in MSK-HQ, EQ-5D-5L and Site-Specific PROM scores, and Patient Global Assessment of Change in MSK Condition Scores**

|  |  |
| --- | --- |
| **PROM** | **Correlation Co-efficient****(95% Confidence Interval)** |
| MSK-HQ | 0.60  | (0.53, 0.65) |
| EQ-5D-5L | 0.37  | (0.28, 0.45) |
| KOOS | 0.59  | (0.46, 0.69) |
| NDI | -0.53  | (-0.70, -0.30) |
| RMDQ | -0.43  | (-0.56, -0.28) |
| SPADI | -0.44  | (-0.59, -0.28) |
| SF-36 PCS | 0.15  | (-0.20, 0.46) |

MSK-HQ=musculoskeletal health questionnaire; KOOS=, Knee Injury and Osteoarthritis Outcome Score; NDI=Neck Disability Index; RMDQ=Roland and Morris Disability Questionnaire; SPADI=Shoulder Pain and Disability Index; SF-36 PCS=Short-Form-12 Physical Component Summary Score.

**Table 4. Correlations Between 6-Month Changes in the MSK-HQ, EQ-5D-5L, and Reference Standard PROMs**

|  |  |  |
| --- | --- | --- |
|  | **MSK-HQ** | **EQ-5D-5L** |
| **MSK-HQ** | - |  |  - |  |
| **EQ-5D-5L** | 0.59  | (0.53, 0.65) |  - |  |
| **KOOS (Knee Pain)** | 0.61  | (0.49, 0.71) | 0.45 | (0.30, 0.58) |
| **NDI (Neck Pain)** | -0.77  | (-0.86, -0.63) | -0.39 | (-0.61, -0.13) |
| **RMDQ (Back Pain)** | -0.69  | (-0.77, -0.59) | -0.67 | (-0.75, -0.56) |
| **SPADI (Shoulder Pain)** | -0.70  | (-0.79, -0.59) | -0.38 | (-0.54, -0.20) |
| **SF-36 PCS (Multi-site Pain)** | 0.17  | (-0.17, 0.48) | 0.26 | (-0.09, 0.55) |

Correlations (95% confidence intervals) reported; all correlations are significant at *P*<0.001, except correlations between the MSK-HQ and EQ-5D-5L and SF-36 PCS. MSK-HQ=musculoskeletal health questionnaire; KOOS=Knee Injury and Osteoarthritis Outcome Score; NDI=Neck Disability Index; RMDQ=Roland-Morris Disability Questionnaire; SPADI=Shoulder Pain and Disability Index; SF-36 PCS=Short-Form-12 Physical Component Summary Score.

**Figure 1. Receiver Operating Characteristic Curves for the MSK-HQ and Other Patient Reported Outcome Measures at Discriminating Between Patients Improving and Not Improving on their Patient Global Assessment Score.**

Panel A: all patients; Panel B: patients with low back pain; panel C = patients with neck pain; panel D = patients with shoulder pain; panel E = patients with knee pain; panel F = patients with multi-site pain. MSK-HQ=musculoskeletal health questionnaire; KOOS=Knee Injury and Osteoarthritis Outcome Score; NDI=Neck Disability Index; RMDQ=Roland-Morris Disability Questionnaire; SPADI=Shoulder Pain and Disability Index; SF-36 PCS=Short-Form-12 Physical Component Summary Score.