Validity and Responsiveness of the Dutch McMaster Toronto Arthritis Patient Preference Questionnaire (MACTAR) in Patients with Osteoarthritis of the Hip or Knee

DI-JANNE J.A. BARTEN, MARTIJN F. PISTERS, TIM TAKKEN, and CINDY Veenhof

ABSTRACT. Objective. To determine the content validity, the construct validity, and the responsiveness of the Dutch McMaster Toronto Arthritis Patient Preference Questionnaire (MACTAR) in patients with osteoarthritis (OA) of the hip or knee.

Methods. The MACTAR comprises 2 parts: a transitional part and a status part. Content validity was investigated by comparing patient-elicited activities to items on the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) and the Medical Outcomes Study Short-Form 36 (SF-36). Construct validity was determined by correlating MACTAR outcomes with WOMAC/SF-36 outcomes. Responsiveness was investigated by correlating MACTAR, WOMAC, and SF-36 change scores with patient global assessment (PGA) scores and plotting a receiver-operating characteristics (ROC) curve.

Results. Eleven percent of the 894 impaired activities identified by 192 patients were not represented in either the WOMAC or the SF-36. The correlations (r) investigated for the MACTAR transitional part varied between 0.27 and 0.40; the status part correlated moderately with the general health scale of the SF-36 (r = 0.44). MACTAR change scores correlated better with PGA than with WOMAC/SF-36 change scores. The area under the ROC curve amounted to 0.90.

Conclusion. Our results suggest that the MACTAR exhibits moderate construct validity and good responsiveness in a population of patients with OA of the hip or knee. The MACTAR is potentially better able to detect changes over time in activities that are important to individual patients compared to other tools measuring physical function (WOMAC, SF-36). Clinicians could use the MACTAR to evaluate clinically relevant changes over time in patient-specific physical functioning. (First Release April 1 2012; J Rheumatol 2012;39:1064–73; doi:10.3899/jrheum.110876)

Key Indexing Terms: OSTE OA HIP KNEE PHYSICAL FITNESS REPRODUCIBILITY OF RESULTS

Osteoarthritis (OA) is a common chronic musculoskeletal disorder, which can result in moderate to severe limitations in physical functioning. Limited physical functioning can lead to a diminished quality of life. OA treatment guidelines recommend exercise therapy to reduce impairments in physical function due to OA. Exercise therapy can thus enable individuals to better meet the demands of daily living.

A number of tools are available to clinicians to evaluate the effect of exercise therapy on physical function. General, disease-specific, and patient-specific tools can be applied as either (self-reported) questionnaires or performance-based tests. A systematic review of the psychometric quality of both questionnaires and performance-based tests in patients with OA of the hip or knee has been published. The reviews recommended the application of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), the Medical Outcomes Study Short-Form 36 (SF-36), and multiactivity tests when evaluating physical function in patients with OA.

Standardized tools, applied to all patients in an identical manner, are recommended for the evaluation of physical functioning. Data produced by these tools may be conveniently and relatively easily categorized and compared between patients and across settings. However, standardized tools are often difficult to interpret at the individual level and fail to take account of individual preferences and variation in the performance of particular activities. Patient-specific tools measuring physical function have been developed based on...
the need for a more patient-centered approach as set out in healthcare policies and to enable clinicians to measure changes in activities that really matter to individual patients\(^1\). In contrast with standardized tools, patient-specific instruments can identify the relevant issues at an individual level and allow evaluation to focus on what is important to each individual patient\(^2\). Although the possibilities to compare statistical data between patients are minimal, the application of patient-specific tools may improve the validity and responsiveness for the assessment of physical function\(^3\).

The McMaster Toronto Arthritis Patient Preference Disability Questionnaire (MACTAR) is one example of a patient-specific scale measuring physical function\(^4\) (Appendix). The objective of the MACTAR is to identify individual disabilities due to the disease and their relative importance to the patient, complemented by questions on general health status\(^5\). The MACTAR has been described as a highly responsive and valid tool for the evaluation of physical function in patients with rheumatoid arthritis (RA)\(^6\). A recent psychometric evaluation of the questionnaire in patients with chronic lower back pain and patients with systemic sclerosis (SSc) showed moderate correlations with general and disease-specific tools that measure physical function\(^7\).

To enable clinicians to use the MACTAR when evaluating physical function in patients with OA of the hip or knee, the psychometric properties of the questionnaire in this specific population must be determined. Therefore, our objective was to determine the content validity, construct validity, and responsiveness of the MACTAR in patients with OA of the hip or knee.

**Materials and Methods**

**Study design.** Data reported in this study were collected from a cluster-randomized controlled trial of 200 patients with OA of the hip or knee over a 12-week period (maximum 18 sessions) that compared behavioral graded activity with usual care in accord with the Dutch physical therapy guidelines\(^8\). The consent of the interventions has been described elsewhere\(^9\). The Medical Ethics Committee of the VU University Medical Center, Amsterdam, approved the study. For the purposes of this validation study, data on "physical function" were used, as well as descriptive data on the study population.

**Study population.** Participants were recruited between November 2001 and May 2003 through participating physical therapists and local newspapers. Dutch-speaking patients with OA of the hip or knee (based on the criteria of the American College of Rheumatology\(^10\)) aged between 50 and 80 years who experienced diminished physical function were included in the study\(^11\). Participants who completed both baseline and followup (Week 13) measurements were eligible for inclusion in the present psychometric evaluation.

**Measurements.** Demographic and clinical data, Demographic and clinical data, including age, sex, duration of symptoms, OA location, and OA grade according to Kellgren and Lawrence\(^12\), were collected from participating patients.

**Physical function.** Dutch MACTAR: The objective of this interview-based measurement tool is to evaluate changes in patient-specific physical function over time. It comprises 2 parts. The baseline interview starts with a transitional part. In this part, a trained interviewer asks the patient to identify up to 10 activities in which he/she experiences difficulties because of OA, such as activities in domestic care, professional life, and social interaction. The identified activities are ranked by the patient from 1 to 10 in order of importance: 1 for the activity the patient most wishes to be able to do without pain or discomfort due to OA, 2 for the next most important activity and so on. The top 5 prioritized activities are evaluated at followup. The second part of the MACTAR (status) contains information on health status. Perceived overall health, as well as psychological, emotional, and social well-being is measured by 5 questions (Likert-type rating scale); when a question obtains a less than optimal score, a followup question probes whether this is due to OA.

**Statistical analyses.** Descriptive statistics were applied to describe the study population. PGJ ratings were dichotomized as "improved" (PGJ score 5, 6, 7, or 8) versus "not improved" (PGJ score 1, 2, 3, or 4). For continuous data, independent t tests were used to calculate differences at baseline between these patients who improved and those who did not. For categorical data, Mann-Whitney U tests were used to compare between groups.

**Construct validity.** Construct validity examines the extent to which the domain in question is comprehensively represented by the items in the questionnaire\(^22\). To determine whether the items in the MACTAR refer to relevant aspects of the construct and are relevant to the purpose of the Instrument, the impaired activities mentioned by patients were compared with items on the WOMAC and the SF-36 physical functioning subscale\(^4\).

**Construct validity.** There is currently no "gold standard" for attributes such as validity and functional status\(^3\). Therefore, construct validity rather than criterion validity was assessed. Construct validity refers to the extent to which scores on a particular instrument relate to other assessment tools in a manner that is consistent with theoretically derived hypotheses\(^3\).

To investigate the construct validity of the MACTAR in patients with OA,
<table>
<thead>
<tr>
<th></th>
<th>Transitional Part</th>
<th>Status Part</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td>Min: 5 (poor health status)</td>
</tr>
<tr>
<td>Followup</td>
<td>Min: 6 (maximum deterioration)</td>
<td>Max: 25 (good health status)</td>
</tr>
<tr>
<td>Change score</td>
<td>Max: 22 (maximum improvement)</td>
<td>Max: 25 (good health status)</td>
</tr>
<tr>
<td></td>
<td>The same as the followup score</td>
<td>Min: −20 (maximum deterioration)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Max: +20 (maximum improvement)</td>
</tr>
</tbody>
</table>

Change scores on the transitional part of the MACTAR were correlated with change scores on both the WOMAC and the SF-36 physical function subscales, as well as the SF-36 role-physical subscale. Further, followup scores on the status part of the MACTAR were correlated with followup scores on the SF-36 general health subscale. For normally distributed data, Pearson correlation coefficients (r) were used to express these correlations. Spearman’s rank correlation coefficients (rs) were applied when data were not distributed normally. The following hypotheses were tested: (1) The change score on the physical function subscale of the WOMAC is negatively correlated (r ≥ −0.5; p < 0.05) with the change score on the transitional part of the MACTAR; the correlation was expected to be negative, because the WOMAC and the MACTAR use reverse scales. (2) Change scores on the physical functioning and role-physical subscales of the SF-36 are positively correlated (r ≥ 0.5; p < 0.05) with the change scores on the transitional part of the MACTAR. (3) Followup scores on the general health subscale of the SF-36 are positively correlated (r ≥ 0.4; p < 0.05) with the followup scores on the status part of the MACTAR.

Responsiveness can be assessed in many different ways. However, one can distinguish 2 definition groups. The first describes responsiveness as “the ability to detect clinically important change”. In this group, an instrument is indicated as high-responsive if it is able to distinguish real change from measurement error. Responsiveness is calculated as the magnitude of a treatment effect in which the standardized response mean (SRM) and effect size could be very useful. The second group defines responsiveness as “the ability to detect changes over time in the construct to be measured”. In this case, responsiveness is independent from any treatment effect and is interpreted as longitudinal validity. It should be assessed in analogy to construct validity. Therefore, predefined hypotheses concerning change scores on the MACTAR, WOMAC, and SF-36 in relation to PGA scores were tested. In the case of normally distributed change scores, parametric statistics were applied; nonparametric variants were applied for data that were not distributed normally. It was hypothesized that (1) the correlation between change scores on the MACTAR (transitional part) and the PGA will be better than that between change scores on the PGA and the WOMAC physical function subscale, the SF-36 physical functioning subscale, and the role-physical subscale, respectively (p < 0.05). (2) Change scores on the MACTAR (transitional part) for patients who have improved according to PGA will differ significantly (p < 0.05) from change scores for those who have not improved according to PGA.

Second, responsiveness was determined by plotting a receiver-operating characteristics (ROC) curve. The first step in this construction was to calculate sensitivity and specificity statistics for MACTAR change scores in patients identified as improved (PGA score > 4) and patients identified as nonimproved (PGA score ≤ 4). Next, the true-positive rate (sensitivity) was plotted in functions of the false-positive rate (1 − specificity) for different cut-off points. The best possible cut-off point would yield a point in the upper left corner of the ROC space, representing 100% sensitivity, 100% specificity, and an area under the curve (AUC) of 1.0. An instrument is indicated as highly responsive if its AUC is ≥ 0.90, moderately responsive if the AUC is between 0.70 and 0.90, and poorly responsive if the AUC is between 0.50 and 0.70.

All analyses were performed using PASW Statistics 18.0. If patients were unable to identify at least 5 impaired activities on the transitional part of the MACTAR, missing activity scores were filled with a score indicating a "no-change situation" (2 points); data from patients who mentioned fewer than 3 impaired activities were excluded from the responsiveness analyses. Further, in cases of just 1 missing followup item for the status part of the MACTAR, the score obtained on the equivalent question in the baseline interview was also used for the followup.

Following the initial analyses, a sensitivity analysis was performed on various cutoff points of the dichotomized PGA score, the aim of which was to determine whether the chosen cutoff point was the optimal point to dichotomize.

RESULTS

Study population. A total of 192 patients participated in both the baseline and the first followup assessment and were included for content and construct validity analyses. The median PGA score of these 148 women and 44 men was 5, representing “slightly improved.” Baseline characteristics of the study population are presented in Table 2.

Outcomes. Table 3 shows absolute scores on the MACTAR, WOMAC, SF-36, and PGA at baseline and followup for both the total population and improved/nonimproved patients. At baseline, there were no differences on any of the outcome measures between patients who indicated that they had improved and patients who indicated that they had not improved. At followup, MACTAR scores (both transitional and status parts), WOMAC physical function scores, and SF-36 physical functioning scores differed significantly between improved and nonimproved patients. The measurement variation was higher in the WOMAC and SF-36 compared with the MACTAR, at both baseline and followup (Table 3).

Content validity. The study population (n = 192) identified a total of 894 impaired activities, a mean of 4.6 impaired activities per patient. Seventy-one patients (37%) were unable to identify at least 5 impaired activities; 1 patient could name only 1 impaired activity; 10 patients identified 2 impaired activities; and 33 patients were able to name a maximum of 3 impaired activities. Walking was most frequently mentioned as the most impaired activity (43%). Overall, 72% of the impaired activities that were identified comprised activities in the category of mobility. Table 4 summarizes all the activities mentioned, ranked by category. All items from both the WOMAC and the SF-36 physical function subscales were represented in the impaired activities list based on the MACTAR questionnaire. However, 27% of the activities mentioned by patients during the MACTAR interview were not represented in the WOMAC and 41%
Table 2. Baseline characteristics of total population and subgroups. Values are number (%), unless otherwise indicated.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total Population, n = 192</th>
<th>Improved Group, n = 144*</th>
<th>Nonimproved Group, n = 48*</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>148 (77)</td>
<td>118 (82)</td>
<td>30 (63)</td>
<td>0.01</td>
</tr>
<tr>
<td>Age, mean (± SD), yrs</td>
<td>64.7 (7.9)</td>
<td>65.5 (7.9)</td>
<td>62.3 (7.6)</td>
<td>0.02</td>
</tr>
<tr>
<td>Location of OA</td>
<td></td>
<td></td>
<td></td>
<td>0.58</td>
</tr>
<tr>
<td>Knee</td>
<td>126 (66)</td>
<td>96 (67)</td>
<td>30 (63)</td>
<td>0.62</td>
</tr>
<tr>
<td>Hip</td>
<td>49 (26)</td>
<td>36 (25)</td>
<td>13 (27)</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>17 (9)</td>
<td>12 (8)</td>
<td>5 (10)</td>
<td></td>
</tr>
<tr>
<td>Radiologic severity of OA*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No features/doubtful</td>
<td>44 (23)</td>
<td>35 (24)</td>
<td>9 (19)</td>
<td></td>
</tr>
<tr>
<td>Minimal/moderate</td>
<td>91 (47)</td>
<td>67 (47)</td>
<td>24 (51)</td>
<td></td>
</tr>
<tr>
<td>Severe or prosthesis</td>
<td>6 (3)</td>
<td>5 (4)</td>
<td>1 (2)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>51 (27)</td>
<td>37 (26)</td>
<td>14 (29)</td>
<td></td>
</tr>
<tr>
<td>Duration of symptoms, yrs</td>
<td></td>
<td></td>
<td></td>
<td>0.55</td>
</tr>
<tr>
<td>&lt; 1</td>
<td>44 (23)</td>
<td>36 (25)</td>
<td>8 (17)</td>
<td></td>
</tr>
<tr>
<td>1–5</td>
<td>71 (37)</td>
<td>52 (36)</td>
<td>19 (40)</td>
<td></td>
</tr>
<tr>
<td>&gt; 5</td>
<td>75 (39)</td>
<td>55 (38)</td>
<td>20 (42)</td>
<td></td>
</tr>
</tbody>
</table>

* Kellgren-Lawrence score. † Improved group: patient global assessment score (PGA) > 4; nonimproved group PGA ≤ 4. OA: osteoarthritis.

Table 3. Baseline, followup, and change scores on the outcome measures (mean ± SD).

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Total Population, n = 192</th>
<th>Improved Group, n = 144</th>
<th>Nonimproved Group, n = 48</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACTAR transitional</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Followup</td>
<td>16.6 (2.8)</td>
<td>17.6 (2.0)</td>
<td>13.6 (2.6)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>WOMAC physical function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>28.6 (11.0)</td>
<td>28.3 (10.9)</td>
<td>29.8 (11.3)</td>
<td>0.42</td>
</tr>
<tr>
<td>Followup</td>
<td>23.0 (11.4)</td>
<td>21.3 (10.6)</td>
<td>27.9 (12.1)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>SF-36 physical function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>48.7 (20.1)</td>
<td>49.1 (19.5)</td>
<td>47.6 (22.1)</td>
<td>0.68</td>
</tr>
<tr>
<td>Followup</td>
<td>56.0 (21.7)</td>
<td>58.0 (21.2)</td>
<td>50.1 (22.1)</td>
<td>0.03</td>
</tr>
<tr>
<td>SF-36 role-physical subscale</td>
<td></td>
<td></td>
<td></td>
<td>0.11</td>
</tr>
<tr>
<td>Baseline</td>
<td>42.5 (41.1)</td>
<td>41.2 (41.0)</td>
<td>46.1 (41.5)</td>
<td>0.48</td>
</tr>
<tr>
<td>Followup</td>
<td>55.9 (42.0)</td>
<td>59.1 (41.5)</td>
<td>46.1 (42.5)</td>
<td>0.07</td>
</tr>
<tr>
<td>SF-36 general health</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>50.8 (19.4)</td>
<td>51.8 (19.7)</td>
<td>47.9 (18.8)</td>
<td>0.24</td>
</tr>
<tr>
<td>Followup</td>
<td>48.9 (16.5)</td>
<td>50.0 (17.2)</td>
<td>45.7 (14.1)</td>
<td>0.13</td>
</tr>
<tr>
<td>PGA, median (range)</td>
<td>5 (6)</td>
<td>6 (3)</td>
<td>4 (2)</td>
<td></td>
</tr>
</tbody>
</table>

MACTAR: McMaster Toronto Arthritis Patient Preference Questionnaire; WOMAC: Western Ontario McMaster Universities Osteoarthritis Index; SF-36: Medical Outcomes Study Short-Form 36; PGA: patient global assessment.

were not represented in the SF-36. Eleven percent of the impaired activities mentioned were not covered by items of the WOMAC or the SF-36: examples of these include gardening and activities related to professional life.

Construct validity. Correlations (r) between change scores on the transitional part of the MACTAR and change scores on the physical function subscales of the WOMAC and the SF-36 were −0.40 (p < 0.01) and 0.27 (p < 0.01), respectively.
Table 4. Patient-mentioned impaired activities on the MACTAR questionnaire.

<table>
<thead>
<tr>
<th>Category Activity</th>
<th>Mentioned as Most Impaired Activity, n (%)</th>
<th>Mentioned in Total, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housekeeping activities*</td>
<td>4 (2.3)</td>
<td>70 (8.5)</td>
</tr>
<tr>
<td>Vacuum cleaning, mopping, washing windows or dishes, lifting buckets, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leisure activities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gardening</td>
<td>19 (11.0)</td>
<td>102 (12.4)</td>
</tr>
<tr>
<td>Remaining leisure activities (cultural activities, shopping)</td>
<td>5 (3.2)</td>
<td>30 (3.6)</td>
</tr>
<tr>
<td>Sports (jogging, tennis, swimming, fitness, riding a horse, dancing)*</td>
<td>4 (2.3)</td>
<td>17 (2.2)</td>
</tr>
<tr>
<td>Mobility</td>
<td>134 (77.9)</td>
<td>591 (71.5)</td>
</tr>
<tr>
<td>Bicycling (including getting up/off)*</td>
<td>19 (11.0)</td>
<td>81 (9.8)</td>
</tr>
<tr>
<td>Climbing stairs*</td>
<td>13 (7.6)</td>
<td>106 (12.8)</td>
</tr>
<tr>
<td>Driving (including getting in/out of car)*</td>
<td>4 (2.3)</td>
<td>52 (6.3)</td>
</tr>
<tr>
<td>Getting up from the floor's chair, getting out of bed*</td>
<td>10 (5.8)</td>
<td>53 (6.4)</td>
</tr>
<tr>
<td>Inability to stand for long*</td>
<td>1 (0.6)</td>
<td>54 (6.5)</td>
</tr>
<tr>
<td>Kneeling down, bending over, reaching down*</td>
<td>12 (7.0)</td>
<td>77 (9.3)</td>
</tr>
<tr>
<td>Remaining mobility activities</td>
<td>1 (0.6)</td>
<td>39 (4.7)</td>
</tr>
<tr>
<td>Walking*</td>
<td>74 (43.0)</td>
<td>129 (15.6)</td>
</tr>
<tr>
<td>Professional activities</td>
<td>1 (0.6)</td>
<td>13 (1.6)</td>
</tr>
<tr>
<td>Remaining activities</td>
<td>1 (0.6)</td>
<td>13 (1.6)</td>
</tr>
<tr>
<td>Self-care activities</td>
<td>1 (0.6)</td>
<td>13 (1.6)</td>
</tr>
<tr>
<td>Dressing (socks, underwear, trousers)*</td>
<td>6 (3.5)</td>
<td>74 (9.0)</td>
</tr>
<tr>
<td>Remaining self-care activities*</td>
<td>1 (0.6)</td>
<td>13 (1.6)</td>
</tr>
<tr>
<td>Sexuality</td>
<td>1 (0.6)</td>
<td>13 (1.6)</td>
</tr>
<tr>
<td>Sleeping and resting, including turning around in bed*</td>
<td>5 (2.9)</td>
<td>18 (2.2)</td>
</tr>
<tr>
<td>Social roles</td>
<td></td>
<td>11 (1.3)</td>
</tr>
<tr>
<td>Total</td>
<td>172 (100)</td>
<td>894 (100)</td>
</tr>
</tbody>
</table>

1 Item is represented in Western Ontario McMaster Universities Osteoarthritis Index. * Item is represented in Medical Outcomes Study Short-Form 36.

MACTAR: McMaster Toronto Arthritis Patient Preference Questionnaire.

Change scores on the transitional part of the MACTAR and the role-physical subscale of the SF-36 were also moderately correlated ($r = 0.27$, $p < 0.01$). Spearman's $r_s$ between follow-up score of the MACTAR status part and the general health subscale of the SF-36 was 0.44 ($p < 0.01$; Table 5).

Responsiveness. Data from 133 patients (82% women, mean age 64.0 ± 8.1 yrs) were used in the responsiveness analyses. Seventy-seven percent of these patients indicated that they had improved following treatment (PGA score ≥ 4), while 23% reported that they had not improved (PGA score ≤ 4). With the exception of age, the improved and nonimproved groups had similar baseline characteristics. Absolute change scores on physical function outcomes are presented in Table 6. Change scores for patients who indicated that they had improved differed significantly from patients who indicated that they had not improved on all outcome measures (Table 6).

Correlations between change scores on the physical function outcomes and the PGA score are also presented in Table 5.

Table 5. Correlation of change scores/followup score outcome measures ($n = 189$) (Spearman's rank correlation coefficient). All correlations were significant at the 0.01 level.

<table>
<thead>
<tr>
<th>MACTAR Transitional</th>
<th>WOMAC Physical Function</th>
<th>SF-36 Physical Functioning</th>
<th>SF-36 Role-physical</th>
<th>MACTAR Status</th>
<th>SF-36 General Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MACTAR transitional</td>
<td>1.00</td>
<td>-0.40</td>
<td>0.27</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>WOMAC physical function subscale</td>
<td>1.00</td>
<td>-0.36</td>
<td>-0.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 physical functioning subscale</td>
<td>1.00</td>
<td>0.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 role-physical subscale</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Followup scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MACTAR status</td>
<td>1.00</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 general health subscale</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MACTAR: McMaster Toronto Arthritis Patient Preference Questionnaire; WOMAC: Western Ontario McMaster Universities Osteoarthritis Index; SF-36: Medical Outcomes Study Short-Form 36.
Table 6. Change scores on physical function measures and correlation coefficients (r) with patient global assessment (PGA).

<table>
<thead>
<tr>
<th></th>
<th>Absolute Change Score, mean (95% CI)</th>
<th>Correlation with PGA Score**</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACTAR transitional*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total population, n = 133</td>
<td>17.7 (17.0 to 18.2)</td>
<td>0.69</td>
</tr>
<tr>
<td>Improved group, n = 102</td>
<td>13.7 (12.7 to 14.7)</td>
<td></td>
</tr>
<tr>
<td>WOMAC physical function*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total population</td>
<td>-6.1 (-7.7 to -4.5)</td>
<td>-0.39</td>
</tr>
<tr>
<td>Improved group</td>
<td>-7.2 (-9.0 to -5.3)</td>
<td></td>
</tr>
<tr>
<td>Nonimproved group</td>
<td>-2.7 (-5.8 to 0.4)</td>
<td></td>
</tr>
<tr>
<td>SF-36 physical functioning*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total population</td>
<td>8.3 (5.2 to 11.3)</td>
<td>0.26</td>
</tr>
<tr>
<td>Improved group</td>
<td>9.7 (6.0 to 13.5)</td>
<td></td>
</tr>
<tr>
<td>Nonimproved group</td>
<td>3.5 (-0.6 to 7.6)</td>
<td></td>
</tr>
<tr>
<td>SF-36 role-physical*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total population</td>
<td>14.9 (6.8 to 22.9)</td>
<td>0.25</td>
</tr>
<tr>
<td>Improved group</td>
<td>19.7 (10.1 to 29.2)</td>
<td></td>
</tr>
<tr>
<td>Nonimproved group</td>
<td>-0.9 (-14.2 to 12.4)</td>
<td></td>
</tr>
</tbody>
</table>

* Significant differences between improved and nonimproved group (p < 0.05). ** Significant at p < 0.05, unless otherwise indicated. MACTAR: McMaster Toronto Arthritis Patient Preference Questionnaire; WOMAC: Western Ontario McMaster Universities Osteoarthritis Index; SF-36: Medical Outcomes Study Short-Form 36.

As hypothesized, change scores on the MACTAR correlate better with PGA (r = 0.69) than change scores on the WOMAC (r = -0.39) and SF-36 (r = 0.26 and 0.25, respectively; Table 6).

Figure 1 presents an ROC curve of the change scores for the MACTAR (transitional part), in which the sensitivity of the MACTAR amounted to its 1 — specificity. The AUC was 0.90 (95% CI 0.89—0.96) with a standard error of 0.03.

The sensitivity analysis showed that the cutoff point for the PGA score dichotomization (> 4) was chosen correctly. Higher and lower cutoff points resulted in less optimal responsiveness values.

**DISCUSSION**

Our aim was to investigate the content validity, construct validity, and responsiveness of the MACTAR in patients with OA of the hip or knee.

The content validity of the MACTAR seems to be good. Specifically, the majority of the impaired activities identified correlate with items on the WOMAC and/or SF-36, which also aim to assess physical function. However, the MACTAR fits better with the WOMAC questionnaire than with the SF-36. This is not surprising, since the WOMAC is aimed specifically at patients with OA, whereas the SF-36 has a more general purpose. Data for 11% of the activities are gathered only by the MACTAR and are not represented in either the WOMAC or the SF-36. These comprised activities in the areas of leisure, professional life, and social interaction. Indeed, participation in these fields varies widely among individuals. Disease-specific and general instruments do not take account of individual limitations, but patient-specific measures such as the MACTAR allow clinicians to evaluate physical functioning at the individual level.

The majority of the activities identified by the MACTAR questionnaire comprised activities in the mobility domain, which corresponds with the majority of activities in daily life. Recent validation studies on the MACTAR questionnaire in patients with chronic low back pain and RA showed comparable results. The most frequently mentioned impaired activity in patients with chronic low back pain was taking part in sports activities; in patients with hip/knee OA, walking was the most commonly cited impaired activity.

Although the content validity of the MACTAR seems to be good in patients with OA, the construct validity is less convincing. Moderate associations between the transitional part of the MACTAR and presumed comparable outcomes (r = 0.4).
might be explained by an unbalanced distribution of impaired activities across the various activity categories. Specifically, the mobility category comprised almost 72% of all reported impaired activities, whereas the mobility domain in the WOMAC contains only 58% and in the SF-36, 60% of the total questionnaire. Thus, the transitional part of the MAC-
TAR covers one specific part of the physical function domain extensively, whereas disease-specific and general tools account for a broader spectrum of this domain. Another explana-
tion for the moderate construct validity could be the narrow variance around the mean on the MAC-TAR, compared with a wide variance in WOMAC and SF-36 scores. The variance is caused by patients who tend to assign the same disability score to very different impaired activities. The difference in variance impedes a comparison between a patient-specific instrument on the one hand and a disease-specific/generic instrument on the other.

As hypothesized, the status part of the MAC-TAR was moderately correlated with the general health subscale of the SF-36 (r = 0.44). Previous studies identified comparable correlation coefficients between the MAC-TAR and other physical function measures. Sanchez, et al.25 found a correlation (r_p) of 0.40 between the MAC-TAR and the Quebec Back Pain Disability Scale37 in patients with chronic low back pain, and a correlation (r_p) of 0.38 (p = 0.002) was found between the MAC-TAR and the Health Assessment Questionnaire (HAQ)45 in patients with SSC22. Verhoeven, et al.21 showed a correlation coefficient (r) of 0.73 (p < 0.0003) between the MAC-TAR and the HAQ in patients with RA.

The MAC-TAR was developed to evaluate patient-specific physical function over time. With this goal in mind, responsiveness is the most important psychometric property. For that reason, we evaluated the responsiveness of the questionnaire. As hypothesized, change scores for the MAC-TAR correlated better with the PGA score than change scores on the WOMAC and SF-36 do, leading to the conclusion that the MAC-TAR is better able to detect changes over time in patients with hip/knee OA than the WOMAC or SF-36. It has also been demonstrated that the MAC-TAR is capable of distinguishing patients who reported an improvement from those patients who reported no improvement. An AUC of 0.90 confirms the high responsiveness of the MAC-TAR in patients with hip/knee OA. Verhoeven, et al.21 also investigated the respons-
siveness of the MAC-TAR, concluding that it showed a high degree of responsiveness, based on an SRM of 3.5. However, an SRM is not an appropriate measure of assessing responsi-

The reproducibility of a single-item transitional scale is probably lower than that for a more extended measurement tool. Finally, "a little better" is not, as a matter of course, equiva-

tent to an important change. However, better external criteria to discriminate between improved and nonimproved patients have not yet been elaborated.

Although the MAC-TAR appears to have some advantages over the WOMAC and the SF-36 in assessing physical function in individual patients, it also has some limitations. The need for a trained interviewer to apply the MAC-TAR, as well as its complicated scoring method, may reduce the likelihood that the MAC-TAR will become the instrument of first-choice in clinical practice. Further, patient-specific measures, including the MAC-TAR, do not take account of shifts in patient pri-
orities that can occur over time in cases of change in disease status. Therefore, further studies should take account of the application of patient-specific measures at long-term followup.

Our results suggest that the MAC-TAR exhibits moderate construct validity and good responsiveness in a population of patients with OA of the hip or knee. Further, the MAC-TAR is potentially better able to detect changes over time in activities that are important to individual patients compared to other tools measuring physical function (WOMAC and SF-36). Therefore, clinicians could use the MAC-TAR to evaluate clinically relevant changes over time in patient-specific physical functioning.

REFERENCES


APPENDIX. McMaster Toronto Arthritis Patient Preference Questionnaire (MACTAR). From J Rheumatol 1987;14:446-50, with permission.

Baseline interview
Interviewers note:
Read the questions and response categories provided. Tick the response given.
1a. How would you say your overall health has been during the last 2 weeks?
   You think of it as
   □ 1. very good
   □ 2. pretty good
   □ 3. not too good

2. Osteoarthritis may cause restrictions in several areas of your daily life. For different people the impact of osteoarthritis is also different. We will ask you to name activities in which you experience difficulties because of your osteoarthritis. What matters here, is what your personal experience has been. Please, think of what became a problem, now that you have osteoarthritis.

Interviewers note:
In order to elicit a comprehensive list of activities: First, give the patient opportunity to react spontaneously. Then read the probes. Record the exact phrases of the patient on the line hereunder.
To support you in naming any problems caused by osteoarthritis, I will read you a number of areas of your daily life that might be affected.

Does your osteoarthritis limit:
   • any (other) activities around the house such as cooking, housework etc.?
   • any activity related to dressing such as buttoning, pulling a sweater over your head etc.?
   • any (other) activities at your work (outside the home), drive a car or other transportation?
   • any (other) leisure activities. Either sports such as bowling, swimming, golf, etc. or non-sports such as needlework, woodwork etc.?
   • any (other) social activities. Such as visiting, playing cards, going to church, etc.?
   • sexuality?

Are there changes in the relationship with your family?
If you live together with a husband/ wife/ partner, are there changes in the relation with him/ her?
If you have children living at home, are there changes in your relationship with them?

[The line above is printed 10 times]

Interviewers note:
To rank the activities in order of importance to the patient, ask the following questions:

Which of these activities would you most like to be able to do without pain or discomfort of your osteoarthritis?
Show and read the list to the patient. Write '1' next to the activity the patient chose.

After (read activity 1), which activity would you next most like to be able to do without pain or discomfort of your osteoarthritis?
Show and read the list to the patient, with exception of the activity with priority 1.
Write '2' next to the activity the patient chose.
After (read activity 1 and 2), which activity would you next most like to be able to do without pain or discomfort of your osteoarthritis?
Show and read the list to the patient, with exception of the activity with priority 1 and 2.
Write '3' next to the activity the patient chose.
Continue like this, until all activities are ranked. The 5 with the highest priority will return in the follow-up interview.

3a. In general, how satisfying do you find the way you spend your life?
   Over the last week you think of it as:
   □ 1. completely satisfying ➔ go to Q4a
   □ 2. pretty satisfying
   □ 3. not very satisfying

4a. How would you say your overall physical functioning has been?
   Over the last week you think of it as:
   □ 1. good ➔ go to Q5a
   □ 2. fair to good
   □ 3. fair
   □ 4. fair to poor
   □ 5. poor
4b. Is your physical function not as good as it might be because of your osteoarthritis?
☐ 0 yes  ☐ 1 no

5a. How would you say your overall social functioning has been?
Over the last week you think of it as:
☐ 0 good  ☐ 1 go to Q6a
☐ 4 fair to good
☐ 3 fair
☐ 2 fair to poor
☐ 1 poor

5b. Is your social functioning not as good as it might be because of your osteoarthritis?
☐ 0 yes  ☐ 1 no

6a. How would you say your overall emotional functioning has been?
Over the last week you think of it as:
☐ 0 good  ☐ 1 stop here
☐ 4 fair to good
☐ 3 fair
☐ 2 fair to poor
☐ 1 poor

6b. Is your emotional functioning not as good as it might be because of your osteoarthritis?
☐ 0 yes  ☐ 1 no

Follow-up interview

1a. How would you say your overall health has been during the last 2 weeks?
You think of it as
☐ 0 very good
☐ 3 pretty good
☐ 2 not too good

1b. Have you noticed any change in your osteoarthritis since we talked during the first interview?
☐ 0 yes  ☐ 1 no  ☐ 2 'no change' at Q1d, go to Q2a

1c. Please describe how your osteoarthritis has changed?

1d. When you think of your osteoarthritis during the two weeks before the first interview, how much better or worse overall has your osteoarthritis become?
☐ 0 a great deal better
☐ 4 moderately better
☐ 3 slightly better
☐ 2 no change
☐ 1 slightly worse
☐ 6 moderately worse
☐ 5 a great deal worse

You may remember the first time we spoke. You told me which activities were at that time problems due to your arthritis. I will ask you again about the five most important.

2a. Since the first interview, have you noticed any change in your ability to (activity 1)?
☐ 0 less of a problem
☐ 3 the same
☐ 1 more of a problem

2b. Since the first interview, have you noticed any change in your ability to (activity 2)?
☐ 0 less of a problem
☐ 3 the same
☐ 1 more of a problem

2c. Since the first interview, have you noticed any change in your ability to (activity 3)?
☐ 0 less of a problem
☐ 3 the same
☐ 1 more of a problem

2d. Since the first interview, have you noticed any change in your ability to (activity 4)?
☐ 0 less of a problem
☐ 3 the same
☐ 1 more of a problem

2e. Since the first interview, have you noticed any change in your ability to (activity 5)?
☐ 0 less of a problem
☐ 3 the same
☐ 1 more of a problem

Questions 3a to 6b are equal to the equally numbered questions in the baseline interview.