Individualized Functional Priority Approach to the Assessment of Health Related Quality of Life in Rheumatology

JENNIFER CLINCH, PETER TUGWELL, GEORGE WELLS, and BEVERLEY SHEA

ABSTRACT. Over the past several years an interest has developed in health related quality of life questionnaires that address the specific concerns of each individual rather than the entire range of potential concerns for all patients. Even disease-specific questionnaires do not capture the concerns of all individuals and some items in these questionnaires are irrelevant for a fairly large minority of people. This paper focuses on individualized functional priority questionnaires that allow patients to specify and prioritize their own personal disease related problems. Of particular interest are the scoring methods and responsiveness of these questionnaires, both necessary variables for defining minimal clinically important differences. (J Rheumatol 2001;28:445–51)

Key Indexing Terms:
OUTCOME AND PROCESS ASSESSMENT QUESTIONNAIRES
MINIMAL CLINICALLY IMPORTANT DIFFERENCES

INTRODUCTION

Health related quality of life (HRQOL) instruments are usually classified as either generic or disease-specific. The advantage of generic instruments is that they can be used to compare the quality of life of those with different conditions, which may inform health policy and economic evaluations. Disease-specific instruments can be used to compare the outcomes of different treatments within a condition such as rheumatoid arthritis (RA). In the OMERACT 4 proceedings, Ortiz, *et al*[1] reviewed the use of generic quality of life instruments in randomized controlled trials (RCT) of rheumatic diseases with special emphasis on their responsiveness. Most of the reviewed studies also included a disease-specific measure. As part of the ongoing work of the OMERACT quality of life task force, the purpose of this paper is to focus on questionnaires that are not only disease-specific but patient-specific in that they take an individualized functional priority approach to the assessment of HRQOL in rheumatic diseases. As part of the OMERACT 5 module on minimal clinically important differences (MCID), this paper will address the responsiveness of HRQOL instruments that emphasize the individual. This will be a first step in considering the MCID in a time when patient-specific perspectives are of great concern. Within the context of the OMERACT filter, questionnaires are assessed for their truth, discrimination, and feasibility. These key issues will be addressed here.

The McMaster Toronto Arthritis Patient Preference Disability Questionnaire (MACTAR) is an individualized functional priority questionnaire that was first published in 1987[2]. Of the few questionnaires of this type, it has been the one used most frequently in clinical trials of various treatments in rheumatology. Since 1987 the instrument has undergone a number of revisions and is now also known as the problem elicitation technique (PET). In creating the instrument, its authors wished to address directly the issues that are important to patients with RA. Hence, the questionnaire asks patients to describe specific activity limitations caused by their arthritis and to rank these problem areas. In order to ensure that the list of affected activities is as comprehensive as possible, a standard series of probes are read to the patient after the spontaneously generated problems. Conventional questionnaires ask a standard set of questions that may or may not apply to a particular patient — some areas may be irrelevant and some relevant areas may not be included. Irrelevant questions create unwanted noise in the responses to such questionnaires, and lack of questions in relevant areas may reduce both sensitivity to change and the ability to discriminate between patients.

Scoring systems have varied considerably in the hands of different authors who have used the questionnaire to study patients with RA undergoing various forms of treatment. Originally, problem areas were identified at the first assessment but no score was assigned. After initiating treatment, change was measured *directly* at followup assessments by asking patients if each of their problems had become worse, remained the same, or improved. Scoring of only the top 3 or top 5 problems and weighting by the inverse of the

From *the Clinical Epidemiology Unit, The Ottawa Hospital, and the Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada.*

J. Clinch, MA, Clinical Epidemiology Unit; P. Tugwell, MD, MSc, FRCPC; G. Wells, PhD, Faculty of Epidemiology and Community Medicine, University of Ottawa; B. Shea, BScN, Clinical Epidemiology Unit, Loeb Health Research Institute.

Address reprint requests to J.J. Clinch, The Ottawa Hospital, General Campus, 501 Smyth Road, Box 201, Ottawa, Ontario, Canada K1H 8L6.

problem's rank order were suggested as possible scoring schemes. Alternatively, some authors have used a visual analog scale (VAS) at both initial and followup assessments and determined change by subtracting baseline scores from followup scores.

A modified version of the MACTAR includes additional conventional questions on physical, social, and emotional functioning, overall quality of life, and a global rating of arthritis. These questions are rated on Likert scales both at baseline and at followup. An assessment of change for each is also included but has not been used by all authors.

In the version of the questionnaire known as the PET, patients continue to identify and rank their problems but are asked to rate on a VAS scale the degree of difficulty, frequency, or severity of the problem, depending on its type. Patients are also asked to rate the importance of each item on a VAS scale, and these values are used as weights. A VAS global health assessment is the final question. Change is assessed by comparing baseline to followup scores.

Definition of an MCID for the MACTAR/PET and other individualized functional priority questionnaires will depend on the scoring method used. If one scoring method is more sensitive to change over time due to treatment, this will influence the choice of an MCID.

METHODS

A MEDLINE search was conducted to find studies that have used the MACTAR or PET as an outcome measure in the treatment of rheumatic diseases. Search terms were included until all studies known to the authors appeared in the search results. These were "physical function," "MACTAR," "McMaster," "PET," "problem elicitation," "quality of life," or "pooled index" combined with "arthritis" and "treatment," "therapy," "responsiveness," or "clinical trial." A total of 530 references were retrieved and the abstracts were reviewed to determine if the studies had used the appropriate type of instruments. Some secondary searches were done based on information contained in the abstracts. A total of 27 articles that described instrument development and/or the use of these instruments in clinical trials were selected for review. These articles are described below with the intent of evaluating the overall experience with the instruments and of comparing the performance of the instruments to other disease-specific instruments and outcomes. The majority of the studies used the MACTAR or PET questionnaire but a number of conceptually related questionnaires are also described; however, most of these have not yet been used extensively in clinical trials.

RESULTS

The original paper describing the MACTAR[2] compared the MACTAR to a conventional questionnaire evaluating both discrete functions and global functions. The questionnaire comprised 17 questions from the functional index developed by Lee, *et al* (discrete functions) and 8 questions taken from the McMaster Health Index (global functions). At the end of the study, patients were asked a global question on whether their arthritis had improved since the baseline questionnaire. Patients were being treated with a variety of therapies (nonsteroidal antiinflammatory drugs, aspirin, gold, penicillamine, corticosteroids). Three unweighted scoring methods were used based on the single most important disability, the top 3 disabilities, and the top 5 disabilities. The response format was the change from baseline, i.e., whether the problem had become worse, remained the same, or improved. Regardless of the number of problems included, the MACTAR showed substantially higher improvement than the conventional questionnaire. This was attributed mainly to the individualized approach to problem assessment, although the authors thought that the MACTAR's emphasis on change might have contributed to its increased sensitivity. Many of the problems identified by patients were not on the conventional questionnaire and even the most frequently identified disabilities were absent or unimportant to many other patients.

In a double-blind, randomized trial of methotrexate versus placebo in RA[3], the effect of treatment on physical, social, and emotional function was measured using 2 standard-item questionnaires, the Lee Functional Index and the McMaster Health Index Questionnaire (MHIQ), and the MACTAR individualized questionnaire. Ten standard clinical endpoints were also used. The relative efficiency (square of the ratio of the t-statistics) of all outcome measures in detecting differences between groups was compared to tender joint count (TJC). Standard-item quality of life questionnaires showed statistically significant but modest improvements varying from 5% to 12%, whereas the MACTAR score improved by 29% over that in the placebo group and had the highest relative efficiency of all the function/quality of life measures. All the function questionnaires had relative efficiencies greater than 4 of the traditional clinical endpoints did. Again, several of the problems identified on the MACTAR were not included in either of the standard-item conventional questionnaires (68% not addressed by the Lee Functional Index and 55% not addressed by the MHIQ).

In a placebo controlled study of low dose cyclosporine in RA[4,5] the primary outcomes were joint count, a VAS pain assessment, and the PET. In addition to the PET, half the patients completed the Health Assessment Questionnaire (HAQ) and half the Arthritis Impact Measurement Scale (AIMS), enabling a comparison of the PET to each of these other quality of life measures. To elicit problems patients were questioned about physical, social, emotional, occupational, communication, and sleep components. Patients were asked to identify the level of difficulty in each area on a 7 point scale and to rank their problems in order of importance. The total score was the sum of difficulty × importance

for the top 5 problems. Responsiveness of these measures and TJC was compared using the standardized effect size (SES) approach. The SES is the ratio of the treatment effect (mean differences in treatment group minus differences in placebo group) to the pooled standard deviation of these differences. For the HAQ cohort the SES were rank ordered as follows: TJC, HAQ disability, PET, and HAQ pain; and for the AIMS cohort the SES were rank ordered: AIMS pain, TJC, AIMS psychological, PET, and AIMS physical function. All measures were judged to have similar responsiveness, except for this latter scale, which was the only one that failed to distinguish between treatment and placebo groups. In a further article Buchbinder, *et al*[6] calculated the relative efficiency of the outcome measures and determined if there were significant differences in the ability to detect a treatment effect compared to TJC. None of the quality of life measures was significantly different from TJC, but pain assessed on a 5 point scale and erythrocyte sedimentation rate (ESR) were both significantly inferior to TJC.

In a RCT comparing femoral head prostheses that were inserted with and without cement, Laupacis, *et al*[7] evaluated the effect of hip replacement on health related quality of life using the Harris hip score, the Merle d'Aubigne hip score, the Sickness Impact Profile (SIP), the Western Ontario McMaster Osteoarthritis index (WOMAC), the MACTAR, a time trade-off utility measure and the 6 minute walk test. The MACTAR form identified the 5 most important physical and social activities adversely affected by hip disease and assessed each pre and postoperatively by 10 cm VAS. Assessments were done preoperatively and at 3, 6, 12, and 24 months postoperatively. One hundred eighty-eight were followed for 3 months; 179 of them for 6 months; 156 for one year; and 90 for 2 years. The 2 treatment groups were combined since relatively small differences in quality of life were expected during the first few years of followup. Pre-post comparisons done on the available patients at each time point were all significant except for the work section of the SIP at 3 months. Responsiveness of the various instruments was not compared. Consistent with the other studies reviewed, 9 of the 20 most frequently chosen problems on the MACTAR were not included in the WOMAC and 6 were not included in the SIP. A later followup study[8] showed similar quality of life results.

Several outcome measures including the PET were used to assess construct validity and sensitivity to change of the PET in 2 trials comparing nonmedical therapies for ankylosing spondylitis (AS) and fibromyalgia (FM)[9]. In patients with AS daily exercises at home were compared to daily exercises plus weekly physical therapy, whereas in patients with FM, low impact fitness training, biofeedback, or no treatment were compared. Overall the AS patients showed significant improvement on the PET, but there was no difference between the treatment groups. In FM patients there was no improvement on the PET when all groups were combined, but the score improved in the fitness group and deteriorated in the biofeedback and controls. Comparative sensitivity to change of all outcome measures was assessed by calculating the efficiency of each measure in the 2 disease groups with all treatments combined. Efficiency is the mean change in the measure divided by the standard deviation of the change. In the AS patients physical fitness, thoracolumbar flexion and extension, patient assessed improvement, and cervical rotation were the most sensitive measures (efficiencies from 3.5 to 0.69). Of the other self-report measures used the PET was by far the most sensitive, with an efficiency of 0.6: the AIMS scales, SIP, HAQ for FM, pain, stiffness, and utilities all had efficiencies below 0.25. In FM patients the highest efficiency was shown by the patient global assessment at 0.48, but no other self-report measure including the PET was responsive. Construct validity of the PET was confirmed by significant correlations ranging from 0.22 to 0.48 with the mobility (FM only), physical activity, social role, activities of daily living (FM only), pain, depression, anxiety, health perception, and arthritis impact scales of the AIMS, and also the SIP, HAQ, pain (VAS), stiffness, utilities, and patient's global health assessment.

Esdaile, *et al*[10] conducted a double blind, placebo controlled trial of hydroxychloroquine in early RA. Four indices comprised the primary outcomes: a joint index, a pain index, a physical function index, and the psychological function subscale of the AIMS. The MACTAR was part of the physical function index with the physical disability dimensions of the HAQ and AIMS. The MACTAR was scored using the S5R method, which weights the scores for the top 5 problems with the inverse of the rank. A significantly greater improvement was noted in the joint, pain, and physical function indices at the end of the study period (36 weeks) for the hydroxychloroquine group compared to placebo. When the scales contributing to the physical function index are examined individually, the MACTAR was the only component showing a significantly greater improvement for the treatment group at 36 weeks. The pain index and its components — the AIMS pain dimension and the HAQ VAS pain scale — all showed a significantly greater improvement in the treatment group. Of the individual joint index components the swollen joint count and grip strength were significantly more improved in the treatment group but there was no difference for the TJC and duration of morning stiffness.

In another study of the quality of life effects of hip arthroplasty, Wright and Young[11,12] compared several questionnaires for responsiveness — 2 disease-specific questionnaires (the Harris Hip Scale and WOMAC), one generic health status scale [Medical Outcome Survey Short Form-36 (SF-36)] and 2 patient-specific indices (the MACTAR and the Patient Specific Index or PASI). In the PASI, patients are asked to rate 21 complaints for severity and

importance plus any additional complaints. Both severity and importance are rated on a 7 category ordinal rating scale. The score on the PASI includes only those items that the patient identifies as problems and is the sum of the products of severity and importance for each item. Since the number of problems can vary from patient to patient, scores are standardized by dividing by the maximum possible score for that patient and multiplying by 100. Several responsiveness measures were used: (a) responsiveness statistic; (b) standardized response mean (SRM); (c) relative efficiency statistic; (d) effect size; and (e) correlation with patient's global rating of change in hip function. Because of the way it was scored (i.e., no available preoperative score) the MACTAR could only be compared using (b) and (e). In fact the calculated SRM for the MACTAR is not strictly comparable to that obtained for the other measures because the denominator or standard deviation of change scores is determined directly from the change indicated postoperatively and is not determined by taking *differences* between a pre and postoperative score. This may tend to overemphasize the responsiveness of the MACTAR in before and after studies. The different responsiveness indices provided different rank ordering of the 14 scales (the WOMAC has 3 subscales and the SF-36 has 8), with the actual ranks differing significantly up to 5 levels. Disease-specific scales were the most responsive, with the MACTAR appearing the most responsive on the SRM. The 2 physical scales of the SF-36 were of intermediate responsiveness and its general and mental health scales were the least responsive.

Boers, *et al*[13] evaluated the effect of combined step-down therapy with prednisolone, methotrexate, and sulfasalazine versus sulfasalazine alone. Outcome was assessed using a pooled index that summarized the change in 5 measures after 28 weeks of treatment. For each group and measure, a standardized change score was calculated by dividing the mean change for that group by the pooled standard deviation of change at week 28. The pooled index was the mean of the standardized scores. A constant was added to all index values so that the value at baseline was zero. Five measures were selected for maximum sensitivity to change: TJC, overall assessment by the independent assessor, grip strength, ESR, and MACTAR. The modified version of the MACTAR was used in this study but without the global assessment of arthritis question (Verhoeven A.C., personal communication). The scores on this questionnaire reflect change, increase as disability improves, and vary from 10 (maximum deterioration) to 40 (maximum improvement). Since there is no baseline score for this instrument, mock change items were added at baseline and scored as unchanged. As well as the items included in the pooled index, swollen joint count, pain, patient's overall assessment, and the HAQ were measured.

At 28 weeks, there were significant differences between the 2 treatment groups on all measures except the patient's overall assessment. After tapering of prednisolone and methotrexate, no significant differences remained between the 2 groups, although there was still a maintained improvement in both groups from baseline.

Verhoeven, *et al*[14] reported further on the validity of the modified MACTAR as demonstrated in this clinical trial. Scores on the MACTAR were partitioned into the "status" questions on physical, social, and emotional functioning, overall quality of life, and global rating of arthritis, and the "transitional" questions that assess change directly. High correlations were found between the total MACTAR score, its partitions, and other functional indices, i.e., the HAQ, AIMS mobility, and grip strength. Correlations with nonfunctional outcomes (i.e., ESR, various pain scales, joint counts, and global assessments) were also high.

Within-group responsiveness of separate items and parts of the MACTAR was assessed by SRM (or mean change within one treatment group divided by the standard deviation of this change). Compared to other outcome measures the total MACTAR showed high responsiveness, with an SRM in the combined treatment group of 2.2 and in the sulfasalazine group of 1.2; the next highest SRM were 1.5 (HAQ, pain VAS, and patient global assessment) and 0.9 (patient global assessment and ESR) in the combined treatment and sulfasalazine groups, respectively. As noted by others, the transitional or direct change items had the best responsiveness. Weighting of these items gave somewhat better SRM for the total MACTAR of 2.8 and 1.3, respectively.

Kirkley, *et al*[15] used the MACTAR as one of several outcome measures to assess the effect of knee bracing on varus gonarthrosis. An unloader brace was compared to a neoprene sleeve and a no-treatment control group. The version of the MACTAR used was the top 3 problems scored on 10 cm VAS. Patients were assessed at baseline and at 6 week and 3 and 6 month followups. Other outcome measures were the WOMAC and pain/distance/stairs climbed on 6 minute walking and 30 second stair climbing. At 6 months significant differences were found between the unloader brace group and control on the WOMAC aggregate, pain, stiffness, and physical function, and on the MACTAR and pain on 6 minute walk and 30 second stair climb, but the distance walked and number of stairs climbed were not different. Comparing the unloader brace versus neoprene sleeve groups only the pain on 6 minute walk and 30 second stair climb were significantly different. Finally, in comparing the neoprene sleeve and control group only the WOMAC stiffness and the 2 pain measures were significantly different. On all measures the greatest improvement was in the order unloader brace, neoprene sleeve, control. The authors attributed the success of both the neoprene sleeve and unloader brace in reducing pain to improved joint proprioception and the greater improvement seen in the unloader brace group to the additional effect of a decrease in compartment loading.

In a RCT of leflunomide versus methotrexate or placebo[16,17], several measures of function and health related quality of life were used to assess patient outcome. These included the Modified Health Assessment Questionnaire (MHAQ), the HAQ, the PET, and the SF-36. To reduce the time required for questionnaire completion the PET problems were limited to those addressed in the HAQ. Tugwell, *et al*[18] compared the outcome measures in terms of their ability to detect a treatment effect by calculating the relative efficiencies (RE) of the measures in comparison to TJC. When comparing leflunomide to placebo the RE in decreasing order were patient global assessment of disease (1.88), HAQ disability index (1.84), SF-36 bodily pain (1.63), MHAQ (1.37), physician global assessment of disease (1.33), PET top 5 and SF-36 physical component (both 1.29), pain intensity scale (1.21), and SF-36 physical functioning scale (1.10). Based on an approximate Z statistic the only 2 RE that were significantly different from TJC were patient global assessment and SF-36 mental component, with a RE of only 0.01.

When comparing methotrexate to placebo, patient and physician global assessments, pain intensity, and C-reactive protein had RE greater than one, whereas none of the disease-specific and generic measures of function were better at detecting treatment effects than TJC. Strand, *et al*[16,17] noted that compared to the methotrexate group, the leflunomide group had significantly greater improvement in the MHAQ scores ($p \leq 0.01$) 5 of 8 scales, and the disability index of the HAQ ($p \leq 0.01$), weighted top 5 score of the PET ($p \leq 0.001$), and 2 of 8 scales (bodily pain and vitality) and the physical component score of the SF-36 ($p \leq 0.01$). Tugwell, *et al* did not report the RE for the leflunomide-methotrexate comparison.

A minimal clinically important difference has not yet been defined for the MACTAR or PET. Kosinski[19] has described an approach that may be useful in defining MCID for quality of life questionnaires in RA. Categories of change were defined in disease severity measures that are familiar to clinicians, such as tender and swollen joint counts and global assessments. The mean changes in quality of life outcomes were then calculated for each category of change in the disease severity measures. These values then attain a clinical meaning for the clinician by virtue of their relationship to more familiar outcomes.

There are a number of questionnaires that are similar to the MACTAR in that they allow individual selection of items and some method of priority ranking. Some are disease-specific and others global. Some representative questionnaires are described below.

There are 2 questionnaires that include a priority function section similar to the MACTAR. A revised and expanded version of the Arthritis Impact Measurement Scales, or AIMS2[20], has a priority function section limited to a choice of 3 of the areas covered by the AIMS2 scales. This section is not scored but has been used in the validation of the remainder of the questionnaire. Wright, *et al*[21,22] have included a priority function section, using the MACTAR framework, in a juvenile arthritis functional status index (JASI). Part I of the JASI is a standard structured item section and part II is the priority function section with the standard set of probes modified to be appropriate for children/adolescents. The 5 chosen activities are scored on a 7 point response option scale going from "Someone has to do it for me" to "As easily as friends without juvenile rheumatoid arthritis." Change was evaluated in 2 ways by using both the original MACTAR format of asking if the functions had improved, worsened, or remained the same and the score difference on the 7 point scale. The estimated kappa for agreement between the 2 scoring methods was 0.22 for a 2 to 3 week interval and 0.09 for a 3 month interval. Aside from the low agreement, only one activity was considered to be worse when change was assessed directly as opposed to 16.2% of activities at 2 to 3 weeks and 11.9% at 3 months for the score difference method, indicating the possibility of bias in the direct change responses. Further work is in progress to evaluate responsiveness to change since the validation study sample was expected to be stable over the study period.

Duffy, *et al*[23] have developed a Juvenile Arthritis Quality of Life Questionnaire (JAQQ) that is similar in concept to the MACTAR/PET. Affected activities have been divided into 4 areas: gross motor function (17 items), fine motor function (16 items), psychosocial (22 items), and general symptoms (19 items). Patients or parents, depending on the child's age, select up to 5 items from each dimension and may add their own items. Each item is scored on a 7 point Likert scale. A mean score is computed for each dimension and an overall mean of the 4 dimensions. Thus, in contrast to the MACTAR/PET each dimension is equally represented in the overall score. The questionnaire is individualized within dimensions but unlike the MACTAR/PET all dimensions must be included. Moderate correlations were found between the JAQQ scores and measures of joint disease activity, but high correlations were found with parent assessed pain (VAS). Similar correlations were found for change scores. Physician assessment of change was significantly related to physician global assessment of change in all JAQQ dimensions except the psychosocial.

Carr's Disease Repercussion Profile (DRP)[24,25] is an individualized measure that gives a profile of perceived handicap in 6 domains: functional activities, social activities, socioeconomic status, relationships, emotional well being, and body image. Patients specify the handicap they are experiencing in each of the domains and rate its severity on a 10 point graphic rating scale. The instrument is designed to help choose an intervention to suit patients rather than to assess outcomes of chosen interventions in groups of patients. Like the JAQQ, the instrument is individualized

within dimensions and all dimensions are rated. No total score is calculated and instead a patient profile is produced. The instrument has been used to assess handicap in groups of patients with RA, osteoarthritis, and low back pain.

A quality of life questionnaire tailored to the individual has been developed by O'Boyle, *et al*[26]. The schedule for the evaluation of individual quality of life (SEIQOL) asks patients to list the 5 areas of life that they judge to be most important to their overall quality of life, which they then rate on a VAS. To quantify the relative weights for each area in the individual's overall judgment of quality of life, patients are asked to rate the overall quality of life of 30 randomly generated profiles of hypothetical people labeled with the same 5 areas chosen by the patient. Multiple regression analysis is then used to estimate the weight attached to each area. This questionnaire has been used to evaluate change in quality of life from pre to postoperatively in patients undergoing hip replacement. A significant improvement was noted in the SEIQOL, the AIMS total score, and the total and physical score of the MHIQ. Changes were smaller in the SEIQOL than the other measures, which may be explained by noting that this is a global rather than disease-specific outcome measure.

Recently a direct weighting method[27] has been used for the chosen areas in the SEIQOL. The apparatus consists of 5 interlocking colored, laminated circular disks that can be rotated around a central point to form a type of pie chart. Each segment is labeled with a life area nominated by the patient who then adjusts the disks until the size of each colored segment corresponds to the relative importance of the life area represented by that segment. Weights derived from the direct weighting and full judgment analysis procedure are similar but there are some differences. The authors suggest that this implies that there are some elements of judgment that are implicit and about which the respondent is unaware, but which may have a bearing on overall judgment.

The Patient Generated Index[28] is very similar to the SEIQOL in structure. Patients are asked to list the 5 most important areas or activities of their life affected by their condition and to rate them on a scale from 0 to 100, where 0 represents the worst they can imagine and 100 represents exactly as they would like to be. A sixth rating covers all other areas of life not previously mentioned. Weights are provided by giving patients 60 points that they can choose to "spend" on improvement across one or more areas. A final score out of 100 is calculated by weighting each area by the proportion of the 60 points assigned to it and summing over all 6 areas. This instrument has been used in cohorts of patients with RA, atopic dermatitis, and low back pain[25].

DISCUSSION

Overall, the experience with the MACTAR or PET patient preference disability questionnaire has shown the instrument to be as responsive as other disease-specific quality of life questionnaires for RA and osteoarthritis. Generally, patient and physician global assessments are more responsive but they provide no information on the specific functional changes that are occurring as a result of treatment. It could be argued that concentrating on what is most important to the patient may provide a more accurate picture of the effects of treatment and certainly one that is more relevant for directing clinical care. There are several scoring procedures for the questionnaire, and it appears that the schemes that concentrate on change rather than assessing current status might be more responsive. As was done with the JASI, these different scoring methods should be compared directly as the subject of further research on the instrument. This is a necessary step before moving forward to define a MCID. For quality of life questionnaires that use a 7 point response scale, Juniper, *et al*[29] have shown that an average scale change of 0.5 points corresponds to a global rating of change of a little or somewhat better or worse, which they defined as a MCID. Their asthma quality of life questionnaire also includes an individualized functional priority section. If a 7 point scale becomes the chosen method of scoring instruments such as the MACTAR/PET then it will be interesting to see if its MCID becomes defined close to 0.5 points.

REFERENCES

1. Ortiz Z, Shea B, Dieguez MG, et al. The responsiveness of generic quality of life instruments in rheumatic diseases. A systematic review of randomized controlled trials. J Rheumatol 1999;26:210-6.
2. Tugwell P, Bombardier C, Buchanan WW, Goldsmith C, Grace E, Hanna B. The MACTAR patient preference disability questionnaire — An individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. J Rheumatol 1987;14:446-51.
3. Tugwell P, Bombardier C, Buchanan WW, et al. Methotrexate in rheumatoid arthritis: Impact on quality of life assessed by traditional standard-item and individualized patient preference health status questionnaires. Arch Intern Med 1990;150:59-62.
4. Tugwell P, Bombardier C, Gent M, et al. Low-dose cyclosporin versus placebo in patients with rheumatoid arthritis. Lancet 1990;335:1051-5.
5. Bombardier C, Buchbinder R, Tugwell P. Efficacy of cyclosporin A in rheumatoid arthritis: Long-term follow-up data and the effect on quality of life. Scand J Rheumatol 1992;21 Suppl 95:29-33.
6. Buchbinder R, Bombardier C, Yeung M, Tugwell P. Which outcome measures should be used in rheumatoid arthritis clinical trials? Arthritis Rheum 1995;38:1568-80.
7. Laupacis A, Bourne R, Rorabeck C, et al. The effect of elective total hip replacement on health-related quality of life. J Bone Joint Surg Am 1993;75:1619-26.
8. Rorabeck C, Bourne R, Laupacis A, et al. A double-blind study of 250 cases comparing cemented with cementless total hip replacement. Cost-effectiveness and its impact on health-related quality of life. Clin Orthop Rel Res 1994;298:156-64.
9. Bakker C, van der Linden S, van Santen-Hoeufft M, Bolwin P, Hidding A. Problem elicitation to assess patient priorities in ankylosing spondylitis and fibromyalgia. J Rheumatol 1995;22:1304-10.

10. The HERA Study Group. A randomized trial of hydroxychloroquine in early rheumatoid arthritis: The HERA study. Am J Med 1995;98:156-68.
11. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol 1997;50:239-4.
12. Wright JG, Young NL. The patient-specific index: asking patients what they want. J Bone Joint Surg Am 1997;79:974-83.
13. Boers M, Verhoeven AC, Markusse HM, et al. Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. Lancet 1997;350:309-18.
14. Verhoeven AC, Boers M, van der Linden S. Validity of the MACTAR questionnaire as a functional index in a rheumatoid arthritis clinical trial. J Rheumatol 2000; in press.
15. Kirkley A, Webster-Bogaert S, Litchfield R, et al. The effect of bracing on varus gonarthrosis. J Bone Joint Surg Am1999; 81:539-48.
16. Strand V, Tugwell P, Bombardier C, et al. Function and health-related quality of life: Results from a randomized controlled trial of leflunomide versus methotrexate or placebo in patients with active rheumatoid arthritis. Arthritis Rheum 1999;42:1870-8.
17. Strand V, Cohen S, Schiff M, et al, for the Leflunomide Rheumatoid Arthritis Investigators Group. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. Arch Intern Med 1999;159:2542-50.
18. Tugwell P, Wells G, Strand V, et al. Clinical improvement as reflected in measures of function and health related quality of life: Sensitivity and relative efficiency to detect a treatment effect in a 12 month placebo controlled trial comparing leflunomide with methotrexate. Arthritis Rheum 2000;43:506-14.
19. Kosinski M, Zhao SZ, Dedhiya S Osterhaus JT, Ware JE. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis Rheum 2000;43:1478-87.
20. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales health questionnaire. Arthritis Rheum 1992;35:1-10.
21. Wright VF, Law M, Crombie V, Goldsmith CH, Dent P. Development of self-report functional status index for juvenile rheumatoid arthritis. J Rheumatol 1994;21:536-44.
22. Wright VA, Kimber JL, Law M, Goldsmith CH, Crombie V, Dent P. The juvenile arthritis functional status index (JASI): A validation study. J Rheumatol 1996;23:1066-79.
23. Duffy C, Arsenault L, Watanabe Duffy K, Paquin J, Strawczynski H. The juvenile arthritis quality of life questionnaire — development of a new responsive index for juvenile rheumatoid arthritis and juvenile spondyloarthritides. J Rheumatol 1997;24:738-46.
24. Carr AJ. A patient-centred approach to evaluation and treatment of rheumatoid arthritis: The development of a clinical tool to measure patient-perceived handicap. Br J Rheumatol 1996;35:921-32.
25. Carr AJ. Beyond disability: measuring the social and personal consequences of osteoarthritis. Osteoarthritis Cartilage 19997:230-8.
26. O'Boyle CA, McGee H, Hickey A, O'Malley K, Joyce CRB. Individual quality of life in patients undergoing hip replacement. Lancet 1992;339:1088-91.
27. Hickey AM, Bury G, O'Boyle CA, Bradley F, O'Kelly FD, Shannon W. A new short form individual quality of life measure (SEIQoL-DW): application in a cohort of individuals with HIV/AIDS. BMJ 1996;313:29-33.
28. Ruta DA, Garratt AM, Leng M, Russell IT, MacDonald M. A new approach to the measurement of quality of life — The patient generated index. Med Care 1994;32:1109-26.
29. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. J Clin Epidemiol 1994;47:81-7.