Editorial

Patient Reported Outcomes for Rheumatoid Arthritis: Where Are We and Where Are We Going?



An increasing number of reports emphasize the importance of patient reported outcome (PRO) measures of health status and health-related quality of life (HRQOL) in rheumatoid arthritis (RA). The study by Linde, *et al* in this issue of *The Journal* examines a range of HRQOL instruments in patients with RA with the aim of assessing reliability, validity, and responsiveness[1].

The authors compared the measurement properties of the Medical Outcome Study Short-Form 36 (SF-36), the EuroQol (EQ-5D), the 15D, Rheumatoid Arthritis Quality of Life (RAQoL) questionnaire, Health Assessment Questionnaire (HAQ), and visual analog scales for pain, fatigue, and global RA. The main findings were that all instruments were able to discriminate between low, moderate, and severe RA activity as measured by the Disease Activity Score 28. Also, the authors found that while the RAQoL and the HAQ had the highest test-retest reliability, the SF-36 bodily pain and vitality subscales were the most responsive to improvement (and the HAQ had low responsiveness).

This study joins a number of articles that in recent years have compared HRQOL instruments directly[2-5] or reviewed them from separate studies[6]. With new instruments emerging for use in RA such as the CSHQ-RA (Cedars-Sinai Health-Related Quality of Life for Rheumatoid Arthritis Instrument)[7], PROMIS (Patient-Reported-Outcomes Measurement Information System)[8], and the computerized adaptive testing in back pain (CAT-5D-QOL)[9], the question remains for a researcher: which instrument to use in future studies?

The most important starting point is to carefully consider the question that you would like to answer with your study and then choose a questionnaire that meets this objective. For example, if one wants to compare the burden of a disease with other diseases or even population norms, then a generic questionnaire might be appropriate. Alternatively, if a specific attribute is the focus of the study, then a disease-specific instrument may be a better choice. If one wants to assess health utilities for use in calculations of cost per quality-adjusted life-year, then using a preference-based measure would be necessary.

Once the objective is defined, studies such as that presented by Linde, *et al* can be particularly useful, since they review the psychometric properties of many available instruments. However, it is important that only instruments meeting the defined objective are compared together. For example, to compare the responsiveness of the SF-36 directly to the EQ-5D is unlikely to be appropriate if the study objective is to assess general HRQOL. (The EQ-5D is purposely short to allow health utilities to be estimated.)

When looking at studies assessing the psychometric properties of instruments, some issues are worthy of consideration. First, any concept of validity or responsiveness is predominant on the reliability of a measure. The study by Linde, *et al* does include reliability but the conclusions of which measure to use should incorporate all these aspects. Second, there are numerous methods for assessing reliability, validity, and responsiveness, and then even more definitions and techniques used to derive results. For example, Linde, *et al* used the standardized response mean to assess responsiveness, but other distributional-based estimates such as the effect size or standard error of measurement are often utilized, along with minimal important difference calculations based on anchor calculations. The use of different methods can give different results[10]. Third, while a majority of the literature to date has focused on reliability, validity, and responsiveness, more recent techniques based on item-response theory can be applied to further assess the psychometric properties of instruments that derive an outcome on the interval scale[11]. Whether instruments meet these criteria should be equally considered. Last, the external validity of results should be considered since the results might be limited to only the population within the study. In

See HRQOL: validity, reliability, and responsiveness of PRO measures in patients with RA, *page 1528*

this respect, the study by Linde, *et al* has the unique result in finding the HAQ to be relatively nonresponsive to patient-reported changes in RA, whereas the majority of other studies find the opposite[4].

In summary, the study by Linde, *et al* adds to the evidence base for helping researchers decide on appropriate questionnaires for research studies.

NICK J. BANSBACK, MSc,
Centre for Health Evaluation and Outcomes Sciences,
Providence Health Care Research Institute;
ASLAM H. ANIS, PhD,
Professor of Health Economics,
School of Population and Public Health,
University of British Columbia;
Director, Centre for Health Evaluation and Outcomes Sciences,
Providence Health Care Research Institute;
CARLO A. MARRA, PharmD, PhD,
Assistant Professor and Director of CORE,
Faculty of Pharmaceutical Sciences,
University of British Columbia,
Vancouver, BC, Canada

Address reprint requests to Dr. C.A. Marra, 2146 East Mall, University of British Columbia, Vancouver, BC. E-mail: carlo.marra@ubc.ca

REFERENCES

1. Linde L, Sørensen J, Østergaard M, Hørslev-Pedersen K, Lund Hetland M. Health-related quality of life: validity, reliability, and responsiveness of SF-36, 15D, EQ-5D, RAQoL, and HAQ in patients with rheumatoid arthritis. J Rheumatol 2008;35:1528-37.
2. Marra CA, Woolcott JC, Kopec JA, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med 2005;60:1571-82.
3. Marra CA, Esdaile JM, Guh D, et al. A comparison of four indirect methods of assessing utility values in rheumatoid arthritis. Med Care 2004;42:1125-31.
4. Marra CA, Rashidi AA, Guh D, et al. Are indirect utility measures reliable and responsive in rheumatoid arthritis patients? Qual Life Res 2005;14:1333-44.
5. Conner-Spady B, Suarez-Almazor ME. Variation in the estimation of quality-adjusted life-years by different preference-based instruments. Med Care 2003;41:791-801.
6. Harrison MJ, Davies LM, Bansback NJ, Ingram M, Anis AH, Symmons DP. The validity and responsiveness of generic utility measures in rheumatoid arthritis: a review. J Rheumatol 2008;35:592-602.
7. Chiou CF, Sherbourne CD, Cornelio I, et al. Development and validation of the revised Cedars-Sinai Health-related Quality of Life for Rheumatoid Arthritis instrument. Arthritis Rheum 2006;55:856-63.
8. Fries JF, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. Clin Exp Rheumatol 2005;23 Suppl 39:S53-7.
9. Kopec JA, Badii M, McKenna M, Lima VD, Sayre EC, Dvorak M. Computerized adaptive testing in back pain: validation of the CAT-5D-QOL. Spine 2008;33:1384-90.
10. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol 1997;50:239-46.
11. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum 2007;57:1358-62.