Editorial

WOMAC: A 20-Year Experiential Review of a Patient-Centered Self-Reported Health Status Questionnaire

A key element in clinical research and clinical practice in musculoskeletal medicine is the evaluation of the therapeutic benefit of interventions used either singularly or in combination. In both clinical research and clinical practice environments, reliability, validity, and responsiveness are essential attributes of health status measurement tools, and in the latter brevity, simplicity, and ease of scoring are regarded with high importance[1,2].

Prior to 1981, measurement procedures for quantifying pain, stiffness, and physical disability in hip and knee osteoarthritis (OA) in rheumatology were diverse and lacked standardization in content, format, and scaling[3]. Further, health status questionnaires were available in very few languages, most often having been developed in English and translated into a few European languages.

The challenge in 1981 was to build a standardized disease-specific patient-relevant self-reported health status questionnaire for hip and knee OA. In 1982, I had the opportunity in the course of completing an MSc thesis to describe the development of a health status questionnaire termed the Western Ontario and McMaster (WOMAC) Osteoarthritis Index[4]. Twenty years later, the WOMAC Index has been extensively validated and has been translated and linguistically validated in over 60 alternative-language forms. In the majority of alternative-language forms it is available in both Likert (LK) and visual analog (VA) scaling formats. There are several hundred citations (full manuscripts, abstracts, reviews) to the use of WOMAC in validation studies, comparative studies against other health status measures, and in its application in various clinical research and clinical practice settings[5].

The idea for the WOMAC index evolved from a brief discussion with Professor Watson Buchanan, a conversation in which I sought his advice in selecting a thesis topic that would address an unmet need in clinical measurement. While development of the idea took only 12 months, the validation and implementation was to consume much of the next 15 years. Between 1996 and 1999 the Index underwent significant refinement, a process that has been consolidated between 1999 and the present, and has resulted in the 3.1 series of WOMAC questionnaires. The WOMAC LK3.1 and WOMAC VA3.1 versions of the Index are now extensively used, particularly in assessing efficacy in pharmaceutical and biotechnology environments.

The success of the WOMAC index is in large part related to 6 factors: (1) Extensive patient involvement in the development of the item inventory[6]. This is perhaps the most important since it is an approach that reduces the potential influence of paternalism, and anchors the item content into aspects of the disease experience that are relevant to patients, and to which they can therefore relate. (2) The conduct of numerous studies evaluating different clinimetric properties of the Index, including analyses evaluating validity, reliability, and responsiveness, comparative studies assessing LK versus VA scaling, blind versus informed presentation, tracking signal items versus complete index usage, parametric versus non-parametric analyses and time frame variations[5]. (3) The development and linguistic validation of numerous alternative-language forms of WOMAC VA3.1 and WOMAC LK3.1 using a standard operating procedure based on tandem forward and backward translation processes and subsequent linguistic validation[5]. (4) Continued research and development into content and administration issues including the application of WOMAC in telephone interviews[7] as well as mouse driven cursor and touch screen electronic data capture formats[8,9]. (5) The incorporation of WOMAC into Osteoarthritis Research Society International (OARSI) clinical trials guidelines as an index relevant to outcome measurement in OA[10]; and (6) the provision of the WOMAC Index, in the required scaling format, alternative-language form, and administration format for academic, commercial, and clinical applications, and ongoing user support.

The development of WOMAC has not been without its challenges. Trans-cultural adaptation of the WOMAC 3.1 Index has been a complex process for which Health Outcomes Group, Palo Alto, California, USA, have taken primary responsibility and in which they have applied their standard operating procedures to develop linguistically valid alternative-language forms of extremely high quality. The preponderance of instruments developed in either North America or Europe might be viewed with concern given the global nature of OA and diversity of lifestyles. It is gratifying, therefore, that the performance of the WOMAC Index has been maintained in its global applications. Thus, while potentially reflecting a restricted view of global diversity, the Index nevertheless appears to tap into the commonalities that exist in the

dimensionality of the symptomatology of OA. That notwithstanding, it is clear that the impact of environmental challenges involved in, for example, stair climbing and transportation are different in different parts of the world, and bathing and toileting habits are quite varied. The index constructor therefore is faced with the dilemma of whether to modify the item content and risk comparing apples with oranges or making minor accommodations in order to maintain a standard question battery. In the case of the WOMAC Index, the latter strategy has been followed, and allows a small degree of flexibility in interpretation. In terms of large multinational clinical trials I believe this is the preferred solution. However, it is possible that at a national level, for clinical practice applications, modification of the item inventory either at an individual question or group of questions (module) level might provide additional advantage.

We have assessed the performance of items self-selected by individual patients, a so-called signal strategy, but have been concerned by the inconsistency with which patients adhere to the selected signal with the passage of time. Our recommendation at the present time, therefore, is to use the entire Index, rather than the signal form.

Scaling format selection is a challenge for any instrument developer, trade-offs often being involved. Likert and VA scales are both commonly used in health status questionnaires. Likert scaling provides a simple and easing scoring system, while the more demanding VA scale may be slightly more sensitive. For this reason we have created parallel forms of the WOMAC 3.1, making available both LK and VA formats for most language forms. In the alternative-language forms, it has been interesting to note that even for the standard scales, word usage is different in different countries. For example, words such as "moderate" and "extreme" may be deemed appropriate in one context, but not in another. As a result the equivalent words may be "average" or "very severe," respectively, in some cultures.

I have been interested to note over the last several years that in some cases the WOMAC Index appears to have passed from one user to another and occasionally in that process the instrument has been altered in a variety of ways. Sometimes the modifications seem quite minor, such as crowding the questions on to one or 2 pages. On other occasions, more radical alterations of the Index have been made such as rescaling the instrument using Health Assessment Questionnaire-style scaling or using a 5 centimeter instead of 10 centimeter visual analog scale on a paper version of the instrument. From time to time I have been sent versions of the instrument that are incomplete, usually the result of the provider not having photocopied the entire instrument when passing it on to a friend or colleague. I have also encountered versions in which additional questions have been added but for which there is no apparent evidence of subsequent revalidation. The concern here is that some modifications may degrade instrument performance, or at the very least erode the level of standardization previously achieved. For this reason, and because the Index, even in English, exists in a number of different forms having different applications, I prefer to provide the most appropriate form of the Index directly to end users in order to better meet their specific measurement needs.

In comparative analyses against other disease-specific and generic health status measures, the WOMAC Index has frequently been superior in performance[11-14]. Two Rasch analyses using an item response theory approach to index construction seem to generally uphold the current structure, although this now popular approach might suggest some modification. However, the consequence of such modification on responsiveness has yet to be determined[15,16]. Recommendations both for shortening the Index[17] and for lengthening the Index[18] have been made, the former to reduce responder burden, the latter to encompass other, potentially younger and more athletic, individuals in orthopedic environments. A role for the WOMAC Index in predicting future health status[19] and health resource utilization[20] has been suggested, but remains to be clarified. Similarly, an application of the WOMAC Index in the assessment of lower limb involvement in rheumatoid arthritis has been suggested, but remains to be verified[21].

It is important to consider whether the development of the WOMAC Index is static or dynamic. The answer is most certainly that it is and remains distinctly dynamic. The developmental form of the WOMAC had 5 subscales (pain, stiffness, physical function, social function, emotional function), the first 3 of which were retained in the original form of the WOMAC and probed the symptom experience of OA in the "hips/and or knees." The WOMAC 3.0 focused on an investigator selected study joint. During that phase of development we also experimented with strategic variations such as using separate WOMAC indices for the study knee and the contralateral knee, and using separate WOMAC pain and stiffness subscales for the left and right knees but a common WOMAC physical function subscale. We have experimented with setting the time frame at 24 h, 48 h (WOMAC 3.1), past 7 days (WOMAC 3.1W), and past month (WOMAC 3.1M), and have created alternative-language forms and a signal version (WOMAC 3.1S). The development of the alternative-language translations has resulted in enhancements to the instructions to patients, the subscale introductory comments, the question stems, and to the WOMAC User Guide. We have looked at short-forming the Index (WOMAC 3.1SF), initial analyses suggesting the preferred short form may be in part dependent on clinical setting, geographic environment, and analytic strategy[22]. Opportunities for electronic data capture by computer-assisted technology have resulted in programs looking at alternatives to patient in-office self-completion on paper[9]. We are currently engaged in an initiative to assess the added value, from an effectiveness and cost-effectiveness standpoint, of providing quantitative WOMAC data to practitioners in a routine clinical care setting. We are also examining an expansion of the current WOMAC inventory

(WOMAC 3.1ER) to accommodate some potential opportunities that may exist in the study of purported structure-modifying OA drugs. WOMAC data have been used in developing a definition of minimum perceptible clinical improvement[23], and together with data from other instruments in developing the OARSI Responder Criteria[24]. We are further evaluating a weighting and aggregation system for the WOMAC Index using a device called the Patient Assessment of the Relative Importance of Symptoms (PARIS) Sectogram[5], and examining the relationship between WOMAC scores and scores from several generic health-related quality of life measures in patients with and without comorbidity. We are currently redeveloping the WOMAC website at www.womac.org to enhance information flow with new and established WOMAC users. An additional consequence of the WOMAC development has been the advantage provided by that experience, in the rapid development of a comparable index, termed the Australian/Canadian (AUSCAN 3.0) Index[25,26], for OA hand studies, details of which can be located at www.auscan.org. All these activities are indicative of a dynamic longterm commitment to advance and refine patient-centered outcome measurement in OA, for application in clinical research and clinical practice environments.

The last 20 years' development of the WOMAC has not been simply the application of classical measurement theory to symptom quantification. It has also involved an extensive collaboration with colleagues in musculoskeletal medicine and other health disciplines, and the interest and commitment of many patients with knee and/or hip OA. I am most grateful to all those who have given their time and resources to support this international initiative. The principal challenges now are to make a good measure even better, to maintain its relevancy in a changing multicultural world, to broaden its application in clinical practice environments, particularly considering issues such as individual response, shared goal setting, and personal and environmental modulators of outcome, to meet emerging needs in structure modifying environments, and to take advantage of emerging technological opportunities.

NICHOLAS BELLAMY, MD, MSc, MBA,
Professor and Director,
Centre of National Research on Disability
and Rehabilitation Medicine (CONROD),
The University of Queensland,
Mayne Medical School, Level 3,
Herston Road, Brisbane, Queensland 4006, Australia.

Address reprint requests to Prof. Bellamy.

REFERENCES

1. Bellamy N, Kaloni S, Pope J, Coulter K, Campbell J. Quantitative rheumatology: A survey of outcome measurement procedures in routine rheumatology outpatient practice in Canada. J Rheumatol 1998;25:852-8.
2. Bellamy N, Muirden KD, Brooks PM, Barraclough D, Tellus MM, Campbell J. Quantitative rheumatology: A survey of outcome measurement procedures in routine rheumatology outpatient practice in Australia. J Rheumatol 1999;26:1593-9.
3. Bellamy N, Buchanan WW. Outcome measurement in osteoarthritis clinical trials: The case for standardisation. Clin Rheumatol 1984;3:293-303.
4. Bellamy N. Osteoarthritis — an evaluative index for clinical trials [MSc thesis]. McMaster University, Hamilton, Ontario, Canada; 1982.
5. Bellamy N. WOMAC Osteoarthritis Index User Guide. Version V. Brisbane: Australia; 2002.
6. Bellamy N, Buchanan WW. A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. Clin Rheumatol 1986;5:231-41.
7. Bellamy N, Campbell J, Hill J. A comparative study of telephone vs on-site completion of the WOMAC 3.0 Osteoarthritis Index. J Rheumatol 2002;29:783-6.
8. Bellamy N, Campbell J, Stevens J, Pilch L, Stewart C, Mahmood Z. Validation study of a computerized version of the Western Ontario and McMaster Universities VA3.0 Osteoarthritis Index. J Rheumatol 1997;24:2413-5.
9. Theiler R, Speilberger J, Bischoff HA, Bellamy N, Huber J, Kroesen S. Clinical evaluation of the WOMAC 3.0 OA Index in numeric rating scale format using a computerised touch screen version. Osteoarthritis Cart 2002;10:479-81.
10. Osteoarthritis Research Society (OARS) Task Force Report. Design and conduct of clinical trials of patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. Osteoarthritis Cart 1996;4:217-43.
11. Hawker G, Melfi C, Paul J, Green R, Bombardier C. Comparison of a generic (SF-36) and a disease-specific (WOMAC) instrument in the measurement of outcomes after knee replacement surgery. J Rheumatol 1995;22:1193-6.
12. March LM, Oh E-S, Cross M, et al. A comparison of WOMAC and MOS SF-36 in OA patients undergoing joint replacement. 8th APLAR Congress of Rheumatology. April 21-26, 1996, Melbourne, Australia. Programme and Abstracts. Abstract #439.
13. Davies GM, Watson DJ, Bellamy N. Comparison of the responsiveness and relative effect size of the WOMAC and SF-36 in a randomized clinical trial in patients with osteoarthritis. Arthritis Care Res 1999;12:172-9.
14. Stucki G, Sangha O, Stucki S, et al. Comparison of the WOMAC (Western Ontario and McMaster Universities) Osteoarthritis Index and a self-report format of the self-administered Lequesne-Algofunctional Index in patients with knee and hip osteoarthritis. Osteoarthritis Cart 1998;6:79-86.
15. Wolfe F, Kong SX. Rasch analysis of the Western Ontario McMaster questionnaire (WOMAC) in 2205 patients with osteoarthritis, rheumatoid arthritis, and fibromyalgia. Ann Rheum Dis 1999;58:563-8.
16. Ryser L, Wright BD, Aeschlimann A, Mariacher-Gehler S, Stucki G. A new look at the Western Ontario and McMaster Universities Osteoarthritis Index using Rasch analysis. Arthritis Care Res 1999;12:331-5.
17. Whitehouse SL, Lingard EA, Learmonth ID, and Kinemax Outcomes Group. A reduced WOMAC function scale — derivation and validation [abstract]. Arthritis Rheum 2000;43 Suppl:S393.
18. Roos EM, Roos HP, Lohmander LS. WOMAC Osteoarthritis Index — additional dimensions for use in subjects with post-traumatic osteoarthritis of the knee. Western Ontario and McMaster Universities. Osteoarthritis Cart 1999;7:216-21.
19. Whitehouse SI, Lingard EA, Learmouth ID. Reduced WOMAC Function Scale validation and derivation [abstract]. Arthritis Rheum 2000;43 Suppl:S393.
20. Ethgen O, Kahler KH, Kong SX, Reginster Y-Y, Wolfe F. Are health-related quality of life scores useful in predicting the use of

health care services? Exploration in patients with arthritis [abstract]. Arthritis Rheum 2000;43 Suppl:S163.
21. Hobby KJ. The effects of aerobic walk-based exercise on women with rheumatoid arthritis [MSc thesis]. London: University of Western Ontario; 1994:1-239.
22. Bellamy N, Bolognese J, Barlas S, et al. Evaluation of factors influencing the item content of the Short Form WOMAC VA 3.1 (SFWOMAC VA3.1) Osteoarthritis Index [abstract]. Arthritis Rheum 2002;46 Suppl:S114.
23. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol 2000;27:2635-41.
24. Dougados M, LeClaire P, van der Heijde D, Bloch DA, Bellamy N, Altman RD. Special Article: Response criteria for clinical trials on osteoarthritis of the knee and hip: A report of the Osteoarthritis Research Society International Standing Committee for Clinical Trials Response Criteria Initiative. Osteoarthritis Cart 2000; 8:395-403.
25. Bellamy N, Campbell J, Haraoui B, et al. Dimensionality and clinical importance of pain and disability in hand osteoarthritis: development of the Australian/Canadian (AUSCAN) Osteoarthritis Hand Index. Osteoarthritis Cart 2002; (In Press).
26. Bellamy N, Campbell J, Haraoui B, et al. Clinimetric properties of the AUSCAN Osteoarthritis Hand Index: an evaluation of reliability, validity and responsiveness. Osteoarthritis Cart 2002; (In Press).