# Toward Ensuring Health Equity: Readability and Cultural Equivalence of OMERACT Patient-reported Outcome Measures

Jennifer Petkovic, Jonathan Epstein, Rachelle Buchbinder, Vivian Welch, Tamara Rader, Anne Lyddiatt, Rosemary Clerehan, Robin Christensen, Annelies Boonen, Niti Goel, Lara J. Maxwell, Karine Toupin-April, Maarten De Wit, Jennifer Barton, Caroline Flurey, Janet Jull, Cheryl Barnabe, Christoph Pohl, Mehmet Tuncay Duruöz, Jasvinder A. Singh, Peter S. Tugwell, and Francis Guillemin

***ABSTRACT. Objective.*** The goal of the Outcome Measures in Rheumatology (OMERACT) 12 (2014) equity working group was to determine whether and how comprehensibility of patient-reported outcome measures (PROM) should be assessed, to ensure suitability for people with low literacy and differing cultures.

***Methods.*** The English, Dutch, French, and Turkish Health Assessment Questionnaires and English and French Osteoarthritis Knee and Hip Quality of Life questionnaires were evaluated by applying 3 readability formulas: Flesch Reading Ease, Flesch-Kincaid grade level, and Simple Measure of Gobbledygook; and a new tool, the Evaluative Linguistic Framework for Questionnaires, developed to assess text quality of questionnaires. We also considered a study assessing cross-cultural adaptation with/without back-translation and/or expert committee. The results of this preconference work were presented to the equity working group participants to gain their perspectives on the importance of comprehensibility and cross-cultural adaptation for PROM.

***Results.*** Thirty-one OMERACT delegates attended the equity session. Twenty-six participants agreed that PROM should be assessed for comprehensibility and for use of suitable methods (4 abstained, 1 no). Twenty-two participants agreed that cultural equivalency of PROM should be assessed and suitable methods used (7 abstained, 2 no). Special interest group participants identified challenges with cross-cultural adaptation including resources required, and suggested patient involvement for improving translation and adaptation.

***Conclusion.*** Future work will include consensus exercises on what methods are required to ensure PROM are appropriate for people with low literacy and different cultures. (J Rheumatol First Release xxxx; doi:10.3899/jrheum.141168)

*Key Indexing Terms:*
HEALTH EQUITY MUSCULOSKELETAL CONDITIONS RHEUMATOLOGY
LITERACY CULTURAL ADAPTATION

*From the Centre for Global Health, University of Ottawa, Ottawa, Ontario, Canada; University of Lorraine, EA 4360 Apemac, Nancy, France; Cabrini Institute and Monash University, Melbourne, Australia; Bruyère Research Institute, and the Institute of Population Health, University of Ottawa, Ottawa, Ontario, Canada; Faculty of Medicine, Nursing and Health Sciences, Monash University, Australia; Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Copenhagen University Hospitals, Bispebjerg and Frederiksberg, Denmark; Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Center and Caphri Research Institute, Maastricht University, Maastricht, The Netherlands; Quintiles Inc.; Division of Rheumatology, Department of Medicine, Duke University School of Medicine, Durham, North Carolina, USA; Children's Hospital of Eastern Ontario Research Institute, Department of Pediatrics, University of Ottawa, Ottawa, Ontario, Canada; VU Medical Centre, Amsterdam, The Netherlands; University of California at San Francisco, San Francisco, California, USA; Faculty of Health and Applied Sciences, University of the West of England, Bristol, UK; Departments of Medicine and Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; Department of Internal Medicine II Rheumatology, Clinical Immunology, Osteology, Physical Therapy and Sports Medicine, Schlosspark-Klinik, Teaching Hospital of the Charité, University Medicine Berlin, Berlin, Germany; University of Marmara, Faculty of Medicine, Department of Physical and Rehabilitation Medicine, Rheumatology Clinics, Istanbul, Turkey; Birmingham Veterans Affairs Medical Center and University of Alabama at Birmingham, Birmingham, Alabama, USA; Ottawa Hospital Research Institute, Ottawa, Ontario, Canada.*

*J. Petkovic, MSc, Centre for Global Health, University of Ottawa; J. Epstein, MD, MSc, University of Lorraine, EA 4360 Apemac; R. Buchbinder, MBBS (Hons), PhD, Cabrini Institute and Monash University; V. Welch, PhD, Bruyère Research Institute, University of Ottawa; T. Rader, MLIS, Institute of Population Health, University of Ottawa; A. Lyddiatt, Patient Research Partner, Ingersoll, Ontario, Canada; R. Clerehan, PhD, Faculty of Medicine, Nursing and Health Sciences, Monash University; R. Christensen, MSc, PhD, Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Copenhagen University Hospitals; A. Boonen, MD, PhD, Professor of Rheumatology, Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Center and Caphri Research Institute, Maastricht University; N. Goel, MD, Senior Medical Director, Quintiles Inc., Adjunct Assistant Professor, Division of Rheumatology, Department of Medicine, Duke University School of Medicine, Patient Research Partner; L.J. Maxwell, MSc, Centre for*

*Global Health, University of Ottawa; K. Toupin-April, Children's Hospital of Eastern Ontario Research Institute, Department of Pediatrics, University of Ottawa; M. De Wit, PhD, Patient Research Partner, VU Medical Centre; J. Barton, MD, University of California at San Francisco; C. Flurey, PhD, Faculty of Health and Applied Sciences, University of the West of England; J. Jull, OT, PhD(c), Institute of Population Health, Centre for Global Health, University of Ottawa; C. Barnabe, MD, MSc, FRCPC, departments of Medicine and Community Health Sciences, Cumming School of Medicine, University of Calgary; C. Pohl, MD, Department of Internal Medicine II Rheumatology, Clinical Immunology, Osteology, Physical Therapy and Sports Medicine, Schlosspark-Klinik, Teaching Hospital of the Charité, University Medicine Berlin; T. Duruöz, MD, University of Marmara, Faculty of Medicine, Department of Physical and Rehabilitation Medicine, Rheumatology Clinics; J. Singh, MBBS, MPH, Associate Professor of Medicine, Birmingham Veterans Affairs Medical Center and University of Alabama at Birmingham; P.S. Tugwell, MD, University of Ottawa, Department of Medicine, Faculty of Medicine, Ottawa Hospital Research Institute, Clinical Epidemiology Program, University of Ottawa, Department of Epidemiology and Community Medicine, Faculty of Medicine, Institute of Population Health, University of Ottawa; F. Guillemin, MD, PhD, University of Lorraine, EA 4360 Apemac.*

*As an OMERACT Fellow, J. Epstein received a bursary from the European League Against Rheumatism and support from OMERACT to attend the meeting; F. Guillemin was supported by La Mission Recherche de la Direction de la recherche, des études, de l'évaluation et des statistiques as part of financial support provided to IReSP in the domain of handicap and loss of autonomy; R. Buchbinder is supported in part by an Australian National Health and Medical Research Council Practitioner Fellowship; conference attendance for A. Lyddiatt was made possible by The Arthritis Society, Canada; J. Barton received support from the American College of Rheumatology, Rheumatology Research Foundation, to attend the meeting; C. Flurey is funded by an Arthritis Research UK Fellowship and received support from OMERACT to attend the meeting; C. Barnabe receives salary support from the Canadian Rheumatology Association and The Arthritis Society Clinician Investigator Award, and is a Canadian Institutes of Health Research New Investigator (Community-Based Primary Health Care); R. Christensen is supported by grants from the Oak Foundation; J.A. Singh is supported by grants from the Agency for Health Quality and Research Center for Education and Research on Therapeutics U19 HS021110, US National Institute of Arthritis, Musculoskeletal and Skin Diseases P50 AR060772 and U34 AR062891, National Institute on Aging U01 AG018947, National Cancer Institute U10 CA149950, the resources and the use of facilities at the VA Medical Center at Birmingham, Alabama, and research contract CE-1304-6631 from the Patient Centered Outcomes Research Institute.*

*Address correspondence to J. Petkovic, University of Ottawa Centre for Global Health, 43 Bruyere St., Annex E, Room 302, Ottawa, Ontario K1N 5C8, Canada. E-mail: jennifer.petkovic@uottawa.ca*

Inequities in health refer to differences in health outcomes that are avoidable and unfair[1]. Patients' low literacy and the lack of adequate cross-cultural adaptation of patient-reported outcome measures (PROM) used in clinical trials can contribute to inequities.

At the Outcome Measures in Rheumatology (OMERACT) 12 meeting (2014), the goal of the equity special interest group (SIG) session was to determine whether and how PROM should be assessed for readability and comprehensibility for patients with different literacy levels and cultures.

The new OMERACT Filter 2.0 outlines the importance of contextual factors as variables in trials[2]. Both the patient's ability to understand the questions being asked and the appropriateness for the patient's culture are contextual factors that need consideration. Filter 2.0 includes a checklist for developing core outcome measurement sets. The equity SIG's longterm goal is to add checklist items that remind developers to consider equity, in terms of the readability of PROM and the potential need for cross-cultural adaptation.

## Background

The distinction needs to be made between literacy and health literacy[3]: this article confines itself to the former. There is no internationally agreed-upon definition of literacy, but most definitions include skills for reading, writing, and numeracy[4]. Globally, the percentage of the population aged 15 years and older who can read, write, and understand simple statements is reportedly 84%[5]. However, this varies by country, region, and population group. Low literacy is associated with lack of health knowledge and preventive behaviors, increased hospitalizations, and poorer self-management of chronic disease[6,7].

Literacy is also an important equity issue for clinical trials. Despite the attention of PROM developers, many PROM include high-level language and complex sentences, which may make understanding and completing them difficult for people with low literacy skills[8]. Because of poor understanding, people with low literacy skills may be less likely to be recruited into trials[9]. This presents ethical and equity issues, because all intended users of a health intervention have the right to participate in research[10], and their exclusion may lead to a risk of selection bias. In addition, when they are included, they may answer questions less accurately or less completely.

Comprehensibility remains an important consideration when a questionnaire is adapted for use in another culture. Cross-cultural adaptation is a related but separate equity issue that is equally essential. Achieving equivalence between the original and adapted versions of a questionnaire refers to the extent to which an instrument is interpreted similarly in 2 or more cultures[11]. While translation has been a focus for PROM developers, cultural adaptation has received less attention[12]. In addition to translation, an adaptation process is required to ensure that items remain equivalent in content when applied in different cultural contexts. This ensures that a construct is measured the same way across cultures, supports fidelity of the culturally adapted tool, and allows for valid comparisons of trial results across countries.

A recent literature review identified 31 guidelines for cross-cultural adaptation, with no consensus on the best method[13]. A summary of those guidelines is provided in Table 1. Most guidelines recommend that questionnaires be translated, back-translated, and then reviewed by a committee to ensure equivalence to the original; however, empirical data are lacking for these recommendations. The role of the committee is to ensure that each item is functionally equivalent in the different setting, and that the translation will be understandable and elicit the same answers.

Failing to consider the role of readability, comprehensibility, and cultural differences in the development/implementation of PROM may lead to measurement errors. This can

*Table 1*. Characteristics of cross-cultural adaptation guidelines. From: Epstein, et al. J Clin Epidemiol 2015;68:435-41[13]; adapted with permission.

| Study, (reference*) | Forward Translation | Synthesis I / Reconciliation | Back Translation | Synthesis II / Back Translation Review | Expert Committee | Pilot Testing |
|---|---|---|---|---|---|---|
| Brislin 1970 (1) | One bilingual translator, familiar with content involved in the instrument | | Second bilingual expert without access to original language version | | Bilingual raters check for errors in original and target versions. English raters examine original and back-translated and check for errors | Bilingual subjects, some who see the source version, some who see the target version, and some who see both. When no meaning errors are found, pretest the translated materials on target language-speaking people |
| Hunt 1986 (2), Hunt et al 1991 (3) | Panel of 8–12 experts (at least 2 bilingual) | Consensus of panel | English teachers review conceptual equivalence based on lay samples | | Experts representing community medicine, sociology, health research and medical specialties | Panel of lay people (2–8) and test/retest with bilinguals (in France) |
| EuroQol Group 1990 (4), Rabin & Charro 2001 (5) | Forward translation of the English source version by 2 qualified translators, native in target language | Consensus of translators and de project manager | Qualified translators, native in English and fluent in target language | | EuroQol Group members cognitive debriefing exercise | Lay sample of target population in a |
| Guillemin et al 1993 (6), Beaton et al 2000 (7) | Two bilingual translators native in the target language, 1 aware of concepts examined in questionnaire, the other naive | Two translators and a recording observer. | Two translators native in source language, neither aware nor informed of concepts explored, preferably without medical background | No | Minimum composition: methodologists, health professionals, language professionals, and translators (forward and back translators). Original developers of the questionnaire in close contact with the expert committee | Ideally, between 30 and 40 individuals, followed by interviews |
| Geisinger 1994 (8) | Translator fluent in both languages, knowledge about both cultures and expert in both characteristics and content measures on the instrument | | Second translator not familiar with instrument | Expert committee | Group of individuals with same criteria as forward translator | Cognitive interviews with a small sample of individuals comparable to target population |
| Sperber et al 1994 (9) | One experienced translator, bilingual physician | | One back translation by a company specialized in medical and scientific translation | | 29 individuals (students and faculty members of school of public health) rated back translated version in comparison with original instrument | Bilingual physicians (test translated and back translated versions), native English monolinguals (test original version), native in target language (translated version) |
| Hambleton 1994 (10), Van de Vijver & Hambleton 1996 (11), Hambleton & Patsula 1999 (12), Hambleton 2001 (13) | Multiple translators (team of translators working independently, additional team of translators to review results) with knowledge of cultures and general knowledge of subject matter | When translators cannot agree on the best translation/ adaptation, both variations should be field tested | Might be useful but not as only design | No | Committees of translators and judges —functionally bilingual, academic background in psychology or education, technical experts in field of translation | Pilot test in small sample of individuals representative of target population. It should consist of administering the test as well as interviewing the individuals to obtain their criticisms of test itself, instructions, etc. |
| Mathias, Fifer & Patrick 1994 (14) | Parallel translation of original instrument by 10 certified translators (for each target language) of Language Institute, but not necessarily native speakers | | 10 different translators | The authors of the instrument | Review by principal investigator from each target country | In the USA, 3–5 bilingual respondents who are native speakers from target country test the translated version. Retest with original and translated version. No debriefing |

*Table 1*. Continued.

| Study, (reference*) | Forward Translation | Synthesis I / Reconciliation | Back Translation | Synthesis II / Back Translation Review | Expert Committee | Pilot Testing |
|---|---|---|---|---|---|---|
| Ware et al 1995 (15), Bullinger et al 1998 (16) | At least 2 translators native in target language, who have experience in questionnaire translation but are not familiar with the measure | Rating of difficulty by original translators, then rating of quality by 2 other independent translators (native English speakers) | National principal investigator and the original translators | Two other translators, native English speakers | Translators, national principal investigator, IQOLA team | Focus groups of up to 50 respondents (patients, healthy members of the community) |
| Bonomi et al 1996 (17) | Two professional translators native speakers of target language, 1 residing in USA and the other in Europe | A third translator, native in target language | Fourth translator native in English | | 3–4 bilingual health professionals residing in native countries | 15–16 patients in each country |
| Erkut et al 1999 (18), Erkut et al 2010 (19) | Bilingual/bicultural research team including researchers indigenous to cultures being studied | | No | No | Bilingual/bicultural research team composed by experts both in the subject matter and the cultures being studied | Successive focus groups of bilingual and monolingual informants provide discussion and consensus |
| Banville et al 2000 (20) | Two bilingual translators (i.e., of statements rather than word-to-word translation) | Two translators | Two other bilingual translators | Expert committee | Three to 5 persons (including translators and researchers involved in the project) | Representative sample of the target population (number of participants is not an important factor because no statistical tests are performed at this time). They are invited to write down their suggestions directly next to the statements |
| Sumathipala & Murray 2000 (21) | Independent translations by members of an expert panel (9 bilingual individuals with some medical background) | Each translation is rated by each member. Inappropriate translations are excluded. The translations are ranked and voted on by the group. Most appropriate translation selected | No | No | Translators and the researcher | Not informed |
| Jones et al 2001 (22) | Two or more independent bilingual translators (from different regions where language is used) | No | Two different bilingual translators blind to the original version of the instrument | Four bilingual experts review the back translations and adapt a new target version that is back translated by two more bilingual experts. Then review the new back translations. The process continues until the team agrees on the cultural equivalence of the source and target versions of instrument | Bilingual and bicultural experts representing a variety of dialects | No pilot testing. Field testing for validation in a sample of bilingual participants to answers both versions (source and target) of instrument |
| Lohr 2002 (23) | At least 2 forward translations from source language; no information about characteristics of translators | A pooled forward translation | At least 1, preferably more | A pooled back translation | Lay and expert panel, not specified | Field tests, not specified |

*Table 1*. Continued.

| Study, (reference*) | Forward Translation | Synthesis I / Reconciliation | Back Translation | Synthesis II / Back Translation Review | Expert Committee | Pilot Testing |
|---|---|---|---|---|---|---|
| Rahman et al 2003 (24) | At least 2 translators, clinically experienced health professionals, native in target language and fluent in original language | Translators, based on key informants (teachers, health professionals, laypersons) interviews | Third translator who has excellent command of both languages, can be a non-technical person | Original translators compare the back translation with the original version | Not specified | Focus group of 8–10 subjects representing study group, paired in 4–5 groups, cognitive debriefing—1 translator acts as moderator and the other takes notes |
| Swaine-Verdier 2004 (25), McKenna et al 2005 (26) | Translation panel: 5 or 7 translators with varied profiles to work as a team; may be informed of model underlying questionnaire, at least 1 native speaker of source language. Work under experienced bilingual coordinator | The translation panel | Not recommended | No | Review under experienced coordinator who has command of both languages | Lay panel: 5–7 people with varied occupational and social backgrounds; individuals with disease covered by questionnaire should be omitted; discussion as a group |
| Wild et al 2005 (27), Wild et al 2009 (28) | At least 2 independent professional translators, native speakers of target language, with prior experience in translation of PRO measures | Key in-country person compares forward translations and discuss with project manager or by independent translator | Native speakers of original language; project manager establishes whether literal or conceptual back translation is required | Project manager and agreement with key in-country person. | Key in-country consultants (native speakers of target language, fluent in source language) with medical/health/psychology or social background, who have experience in translating PRO measures; translators; project manager | Group of 5–8 respondents representing the target population; cognitive debriefing carried out by in-country consultants |
| Eremenco, Cella & Arnold 2005 (29) | Two independent translators, at least 1 professional translator; 1 should live in USA and the other, if possible, in the target country | One native speaker of target language who was not involved with forward translation | One native English speaker, fluent in target language, who was not involved in previous steps of the translation process | A member of the FACIT staff, representing the test developer | Modified Delphi approach: 3–4 bilingual experts, including health professionals and if possible, developer of instrument. After reviews are completed, evaluation by FACIT staff and language coordinator | A wide range of the target population; retrospective debriefing and cognitive debriefing interview |
| Wang, Lee & Fetzer 2006 (30) | Two bilingual translators, native in target language | Peer group (high school students and 2 PhD researchers native in target language and who had lived in the USA) | Graduate student, native in target language and majoring in language arts, who was not familiar with the instrument | A language arts university professor | No information | Monolingual testing: 2 native English speakers compared source and back-translated versions. Bilingual testing: 32 bilingual high school students answered the source instrument and 1 week later the target version |
| Koller et al 2007 (31), Dewolf et al 2009 (32), Kulis et al 2011 (33) | Two professional translators, native speakers of target language, fluent in English, and familiar with medical terminology and patient language | Project coordinator (clinician or methodologist, native speaker of target language), consensus with 2 translators. A third translator may be consulted | Two translators native speakers of English | Expert committee | Translation Committee (EORTC), project coordinator, and translators | Minimum of 10–15 patients belonging to the target population; cognitive interviewing |

*Table 1*. Continued.

| Study, (reference*) | Forward Translation | Synthesis I / Reconciliation | Back Translation | Synthesis II / Back Translation Review | Expert Committee | Pilot Testing |
|---|---|---|---|---|---|---|
| Aquadro et al 2008 (34), MAPI Research Institute 2012 (35) | Two independent, professional translators, native in target language | Two "forward" translators and local project manager | One independent translator | Local team and discussion with MAPI Institute | Not specified (MAPI Institute) | Interviews with patients/healthy subjects to test interpretation of the translation |
| WHO 2008 (36) | One translator, preferably health professional, familiar with terminology of area covered by instrument, native in target language, with knowledge of the English-speaking culture | Agreement with expert panel | Independent translator native in English, who has no knowledge of the original questionnaire | Expert committee | Bilingual (in English and in target language) experts, including original translator, health professionals, and experts with experience in instrument development and translation | Minimum of 10 individuals representative of population, males and females, different socioeconomic groups, followed by cognitive interviews |
| Martinez et al 2008 (37) | Two independent translators native speakers of the target language, including those who speak different dialects if necessary | Group of bilingual people who are similar to the target population | Two different translators | Translators and group of bilingual people | Not specified | Cognitive interviews |
| Price et al 2008 (38) | Single translation by bilingual clinical expert, native speaker of target language | No | Single back translation by a qualified professional translator, naive to the objectives of the translation | Expert committee | Experts in the disease and experts in the measure's development | 4 cognitive debriefers to debrief at least 10 children and 10 parents (children less than 7 years) |
| Gjersing, Caplehorn & Clausen 2010 (39) | Two independent translators fluent in target language, good understanding of original language | Third independent translator, fluent in target language, good understanding of original language | Two translators fluent in original language, good understanding of target language | A third back translator, fluent in original language, good understanding of target language | Methodologists, health professionals, language professionals, and the translators (forward and back translators) | Between 30 and 40 individuals, followed by interviews |
| Baker et al 2010 (40) | Community members, certified in language translation | Community team leader and researcher | Simultaneous to forward translation by community members, certified in language translation | Team leader and researcher | Community team and researchers | Cognitive interviews with representative members across socioeconomic spectrum of community. Final version is field tested with a sample of 10 bilingual community members |
| Two et al 2010 (41) | One forward translation per target country produced by the in-country investigator who is a native speaker of target language, fluent in English with background in medicine or linguistics | Teleconference to produce a single reconciled translation | Two back translations | Project manager/Teleconference | A clinician in each country and the instrument developer | 5 patients in each country |
| King et al 2011 (42) | Multi-language versions of the study questionnaires were developed by an international company, which uses accredited translators and guarantees accurate translations | | No | | Healthcare reviewers fluent in English and in language of interest examine study materials to ensure that "clinical meaning" was appropriate. Lay reviewers offered suggestions to achieve conceptual equivalence. A key informant of each language group reviewed all comment. | Not performed |

*Table 1*. Continued.

| Study, (reference*) | Forward Translation | Synthesis I / Reconciliation | Back Translation | Synthesis II / Back Translation Review | Expert Committee | Pilot Testing |
|---|---|---|---|---|---|---|
| Sousa & Rojjanasrirat 2011 (43) | At least 2 independent translators, preferably certified, native in the target language, fluent the original language, preferably bicultural, one of them with knowledge about health terminology and the other, naive | Two forward translators with participation of a in third bilingual translator | Two other independent translators, native in source language, fluent in target language, blind to original version of instrument, one of them with knowledge about health terminology and the other, naive | Expert Committee | At least 1 methodologist, 1 health care professional who is familiar with the content areas of the construct of the instrument, and all involved translators | Sample size of 10–40 individuals. Instructions, response format and items found to be unclear by at least 20% of sample must be re-evaluated by expert panel |
| Hou et al 2013 (44) | One translation by researcher and nursing expert with knowledge of both English language and nursing terminology (Translation 1) | A second round of nursing experts validate the preferred terms categorizing their level of agreement using a 5 level-scale (disagreement to complete agreement) | No | No | Nursing experts with clinical and academic experience | Not informedI |

QOLA: International Quality of Life Assessment; FACIT: Functional Assessment of Chronic Illness Therapy; EORTC: European Organisation for Research and Treatment of Cancer.


*Table References

1. Brislin RW. Back-translation for cross-cultural research. J Cross Cult Psychol 1970;1:185–216.
2. Hunt SM. Cross-cultural issues in the use of socio-medical indicators. Health Policy 1986;6:149–58.
3. Hunt SM, Alonso J, Bucquet D, Niero M, Wiklund I, McKenna S. Cross-cultural adaptation of health measures. European Group for Health Management and Quality of Life Assessment. Health Policy Amst Neth 1991;19:33–44.
4. The EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy Amst Neth 1990;16:199–208.
5. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001;33:337–43.
6. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417–32.
7. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000;25:3186–91.
8. Geisinger KF. Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. Psychol Assess 1994;6:304–12.
9. Sperber AD, Devellis RF, Boehlecke B. Cross-cultural translation methodology and validation. J Cross Cult Psychol 1994;25:501–24.
10. Hambleton RK. Guidelines for adapting educational and psychological tests: a progress report. Eur J Psychol Assess 1994;10:229–44.
11. Van de Vijver F, Hambleton RK. Translating tests: Some practical guidelines. Eur Psychol 1996;1:89–99.
12. Hambleton RK, Patsula L. Increasing the validity of adapted tests: myths to be avoided and guidelines for improving test adaptation practices. J Appl Test Technol 1999;1:1–30.
13. Hambleton RK. The next generation of the ITC Test Translation and Adaptation Guidelines. Eur J Psychol Assess 2001;17:164–72.
14. Mathias SD, Fifer SK, Patrick DL. Rapid translation of quality of life measures for international clinical trials: avoiding errors in the minimalist approach. Qual Life Res 1994;3:403–12.
15. Ware JE Jr, Keller SD, Gandek B, Brazier JE, Sullivan M. Evaluating translations of health status questionnaires. Methods from the IQOLA project. Int J Technol Assess Health Care 1995;11:525–51.
16. Bullinger M, Alonso J, Apolone G, Leplège A, Sullivan M, Wood-Dauphinee S, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Assessment. J Clin Epidemiol 1998;51:913–23.
17. Bonomi AE, Cella DF, Hahn EA, Bjordal K, Sperner-Unterweger B, Gangeri L, et al. Multilingual translation of the Functional Assessment of Cancer Therapy (FACT) quality of life measurement system. Qual Life Res 1996;5:309–20.
18. Erkut S, Alarcón O, García Coll C, Tropp LR, García HA. The dual-focus approach to creating bilingual measures. J Cross Cult Psychol 1999;30:206–18.
19. Erkut S. Developing multiple language versions of instruments for intercultural research. Child Dev Perspect 2010;4:19–24.
20. Banville D, Desrosiers P, Genet-Volet Y. Translating questionnaires and inventories using a cross-cultural translation technique. J Teach Phys Educ 2000;19:374–87.
21. Sumathipala A, Murray J. New approach to translating instruments for cross-cultural research: a combined qualitative and quantitative approach for translation and consensus generation. Int J Methods Psychiatr Res 2000;9:87–95.
22. Jones PS, Lee JW, Phillips LR, Zhang XE, Jaceldo KB. An adaptation of Brislin's translation model for cross-cultural research. Nurs Res 2001;50:300–4.

*Table References (continued)
23. Lohr KN. Assessing health status and quality-of-life instruments: Attributes and review criteria. Qual Life Res 2002;11:193–205.
24. Rahman A, Iqbal Z, Waheed W, Hussain N. Translation and cultural adaptation of health questionnaires. J Pak Med Assoc 2003;53:142–7.
25. Swaine-Verdier A, Doward LC, Hagell P, Thorsen H, McKenna SP. Adapting quality of life instruments. Value Health 2004;7:S27–30.
26. McKenna SP, Doward LC. The translation and cultural adaptation of patient-reported outcome measures. Value Health 2005;8:89–91.
27. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health 2005;8:94–104.
28. Wild D, Eremenco S, Mear I, Martin M, Houchin C, Gawlicki M, et al. Multinational trials-recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force report. Value Health 2009;12:430–40.
29. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. Eval Health Prof 2005;28:212–32.
30. Wang W-L, Lee H-L, Fetzer SJ. Challenges and strategies of instrument translation. West J Nurs Res 2006;28:310–21.
31. Koller M, Aaronson NK, Blazeby J, Bottomley A, Dewolf L, Fayers P, et al. Translation procedures for standardised quality of life questionnaires: The European Organisation for Research and Treatment of Cancer (EORTC) approach. Eur J Cancer 2007;43:1810–20.
32. Dewolf L, Koller M, Velikova G, Johnson C, Scott N, Bottomley A, on behalf of the EORTC Quality of Life Group. EORTC quality of life group translation procedure, 3rd ed. Brussels, Belgium: EORTC; 2009. [Internet. Accessed May 11, 2015.] Available from: http://groups.eortc.be/qol/sites/default/files/archives/translation_manual_2009.pdf
33. Kuliś D, Arnott M, Greimel ER, Bottomley A, Koller M, European Organisation for Research and Treatment of Cancer Quality of Life Group. Trends in translation requests and arising issues regarding cultural adaptation. Expert Rev Pharmacoecon Outcomes Res 2011;11:307–14.
34. Acquadro C, Conway K, Hareendran A, Aaronson N, European Regulatory Issues and Quality of Life Assessment (ERIQA) Group. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. Value Health 2008;11:509–21.
35. MAPI Research Institute, Lyon, France; 2012. [Internet. Accessed May 11, 2015.] Available from: www.mapi-institute.com/linguistic-validation/methodology
36. WHO. Process of translation and adaptation of instruments. World Health Organisation. [Internet. Accessed May 11, 2015.] Available from: www.who.int/substance_abuse/research_tools/translation/en/
37. Martinez SM, Ainsworth BE, Elder JP. A review of physical activity measures used among US Latinos: guidelines for developing culturally appropriate measures. Ann Behav Med 2008;36:195-207.
38. Price VE, Klaassen RJ, Bolton-Maggs PH, Grainger JD, Curtis C, Wakefield C, et al. Measuring disease-specific quality of life in rare populations: a practical approach to cross-cultural translation. Health Qual Life Outcomes 2009;7:92.
39. Gjersing L, Caplehorn JR, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. BMC Med Res Methodol 2010;10:1–10.
40. Baker DL, Melnikow J, Ly MY, Shoultz J, Niederhauser V, Diaz-Escamilla R. Translation of health surveys using mixed methods. J Nurs Scholarsh 2010;42:430–8.
41. Two R, Verjee-Lorenz A, Clayson D, Dalal M, Grotzinger K, Younossi ZM. A methodology for successfully producing global translations of patient reported outcome measures for use in multiple countries. Value Health 2010;13:128–31.
42. King KM, Khan N, Leblanc P, Quan H. Examining and establishing translational and conceptual equivalence of survey questionnaires for a multi-ethnic, multi-language study. J Adv Nurs 2011; 67:2267–74.
43. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract 2011;17:268–74.
44. Hou I-C, Chang P, Chan H-Y, Dykes PC. A modified Delphi translation strategy and challenges of International Classification for Nursing Practice (ICNP®). Int J Med Inf 2013;82:418–26.
45. Schuster C, Hahn S, Ettlin T. Objectively-assessed outcome measures: a translation and cross-cultural adaptation procedure applied to the Chedoke McMaster Arm and Hand Activity Inventory (CAHAI). BMC Med Res Methodol 2010;10:106.

affect our ability to accurately evaluate the effect of interventions across all populations with rheumatic diseases, including disadvantaged groups, and may contribute to inequities. Special attention must be given to the equity aspect of PROM, particularly as clinical trials often use these instruments as primary outcome measures.

## MATERIALS AND METHODS

*Literature review.* In preparation for the OMERACT 12 meeting, we conducted a literature review to identify methods for measuring comprehensibility and cultural adaptation of questionnaires and instruments used in health research in Medline. A sensitive search strategy was designed to retrieve systematic reviews describing the methods of measuring or validating the comprehensibility and cultural adaptation of questionnaires and instruments in health research. An electronic search strategy (see Appendix 1) was developed for OVID MEDLINE (1946–December 11, 2013) by a librarian (TR) and refined after expert review of a selection of sample citations retrieved by the OMERACT equity group. Experts in the field of comprehensibility of written materials and cultural adaptation were also consulted to identify relevant papers.

*Comprehensibility and overall quality of questionnaires.* Two questionnaires available in at least 2 languages, the Health Assessment Questionnaire (HAQ) and the Osteoarthritis Knee and Hip Quality of Life (OAKHQOL), were considered. The HAQ, developed in English, is widely used and has been adapted into over 60 languages[14,15]. The HAQ was assessed in English, Dutch, French, Spanish, and Turkish; and the OAKHQOL in English and French. The OAKHQOL is a recently developed disease-specific questionnaire[16].

Two methods for assessing the comprehensibility of written text were applied: (1) Readability formulas: We used 3 standard readability formulas: the Flesch Reading Ease (FRE), the Flesch-Kincaid grade level (FK), and the Simple Measure of Gobbledygook (SMOG)[17,18,19]. These assess text complexity using sentence length and syllables per word (Table 2). To apply the formulas in this study, when a question stem had multiple responses, the text was modified to create complete sentences to allow assessment. (2) Evaluative Linguistic Framework for Questionnaires (ELF-Q): This tool, developed from the ELF[20,21] is based upon systemic functional linguistics and provides a more meaningful assessment of the likely comprehensibility of written materials and how they can be improved[22]. The ELF-Q assesses

*Table 2*. Standard readability formulas.

| Test | Formula |
|---|---|
| Flesch Reading Ease* | FRE = 206.835 – (1.015 × average sentence length*) – (84.6 × average number of syllables per word) |
| Flesch-Kincaid** | FK = (0.39 × average sentence length⁺) + (11.8 × average number of syllables per word) – 15.59 |
| SMOG# | SMOG = 3 + √(polysyllable count)† |

*FRE scores range from 0–100, with 0 representing "practically unreadable" text and 100 representing text that is "easy for any literate person." **FK readability test is intended to indicate the US grade level, or number of years of education needed to comprehend the written text[30]. ⁺Average sentence length: average number of words per sentence. #Information should be written at a grade 6 SMOG level to ensure that material is understandable for patients with low literacy[23]. Higher scores correlate with higher US grade levels. †To calculate the polysyllable count: count every word with > 3 syllables in each group of 10 sentences (1 group from the beginning, from the middle, and the end of the text, respectively)[18]. SMOG: Simple Measure of Gobbledygook.

*Table 3*. Readability of the HAQ and OAKHQOL according to readability formulas.

| | Language | FK* | FRE** | SMOG⁺ |
|---|---|---|---|---|
| HAQ | | | | |
| | English | 7 | 79.45 "Fairly easy" | 7.33 |
| | French | 6.85 | 48.00 "Difficult" | 9.17 |
| | Spanish | 10 | 41.38 "Difficult" | 8.40 |
| | Dutch | 9.86 | 45.52 "Difficult" | 11.76 |
| OAKHQOL | | | | |
| | English | 5.94 | 68.45 "Standard" | 9.26 |
| | French | 5.87 | 65.00 "Standard" | 8.66 |

* FK: range 0–12, although higher scores are possible with complicated text. ** FRE: range 0-100 (0 = unreadable, 100 = easy)[19]. ⁺SMOG: range 1–19 and above (grades 13–16 indicate need for college education; 17–18, need for postgraduate training; and 19+, need for higher professional qualification)[18]. HAQ: Health Assessment Questionnaire; OAKHQOL: Osteoarthritis Knee and Hip Quality of Life; FRE: Flesch Reading Ease; FK: Flesch-Kincaid grade level; SMOG: Simple Measure of Gobbledygook.

characteristics related to the overall organizational or generic structure of the text, metadiscourse, headings, rhetorical elements, relationship between writer and reader, technicality of vocabulary, lexical density, context appropriateness, and format. An overall judgment can then be made about overall text quality.

*Cultural adaptation*. The results were presented of a study initiated at OMERACT 10 (2010) that compared various methods of cross-cultural adaptation with/without back-translation and with/without expert committee in an experimental design[23]. The expert committees each had 6 people: translator, linguist, clinician, health education theory expert, patient, and methodologist.

*SIG session*. Participants were introduced to the concept of readability and comprehensibility. The limitations of standard readability formulas were acknowledged and the ELF-Q was described. Results of the readability assessments and cross-cultural adaptation study were presented and discussed. The participants were asked to discuss the potential for these concepts to be included in the Filter 2.0 checklist for PROM developers.

## RESULTS

*Literature review*. Our MEDLINE search identified 177 review articles. After duplicates were removed, 166 were screened for relevance by title and abstract. We obtained 19 for full-text review and 5 were included in the literature review. Of these, 4 reviews assessed patient questionnaires using readability formulas. Two studies used the Flesch-Kincaid grade level test and 2 studies compared a range of readability formulas including Windows-based software, Reading Calculations, FORCAST, Flesch Reading Ease, and Gunning FOG formulas[22,24,25,26]. However, no other frameworks to assess the readability and overall quality of the questionnaires were identified. In addition, we identified 94 articles that discussed approaches and methods for cross-cultural adaptation and 31 different guidelines, but no best practices for cross-cultural adaptation of surveys and questionnaires were identified, although some methods were used regularly in the literature[27,28,29,30,31].

*Comprehensibility and overall quality of questionnaires (Table 3)*. The reading levels of all HAQ versions were above the recommended Grade 6 level using the FK and SMOG tests[32]. The FRE assessed the English HAQ as "fairly easy," but the French, Spanish, and Dutch versions were "difficult." The French OAKHQOL was rated as marginally easier than the English version according to the FK and SMOG, but both were "standard" using the FRE.

All versions of the HAQ and OAKHQOL were considered acceptable according to the ELF-Q; however, minor changes could be made to ensure optimal questionnaire comprehensibility. All versions used vocabulary that was considered difficult or rare according to lists of most commonly used words in the different languages[33,34,35,36]. Improvements could include explicitly stating the purpose of the questionnaire and simplifying the word choices. In terms of context appropriateness, items in the 2 questionnaires were considered generalizable to respondents in all social strata, social/national groups in the society, and they appeared clear and unambiguous. The response options were also clear and unambiguous.

*Cultural adaptation*. The study results indicated that, among 4074 patients and 15 bilingual people, back-translation had a moderate effect, but expert committees were more effective in ensuring accurate content when adapting a questionnaire. The adaptations made with a back-translation step were not considered better or worse than the others, whereas the adaptations using an expert committee were considered to have better face and content validity. The effects of back-translation and expert committees on other psychometric properties were not significant[23].

*SIG session*. Thirty-one OMERACT 12 delegates from Australia, Europe, and North America attended the equity SIG, including 6 patient research partners.

Participants discussed the challenges of ensuring target patients are considered when assessing PROM for text

comprehensibility and cross-cultural adaptation. Distinguishing comprehensibility of PROM and cross-cultural adaptation of PROM as separate but related concepts was considered important. Comprehensibility is an issue for PROM intended for use within 1 culture, but is also an issue for cross-cultural adaptation of PROM for use across different cultures. SIG participants agreed that back-translation does not guarantee accuracy. Challenges discussed included the resources required to complete translation and cross-cultural adaptation successfully, and that these may be barriers to using the committee approach. Patient involvement in the cross-cultural adaptation of questionnaires was discussed as a way of improving the process.

Twenty-six participants agreed that PROM should be assessed for comprehensibility and that suitable methods should be used (4 abstained, 1 no). Twenty-two participants agreed that the cultural equivalency of PROM in different cultures should be assessed and that suitable methods are currently being applied (7 abstained, 2 no).

## DISCUSSION

For the first time, both comprehensibility and cross-cultural adaptation of PROM have been considered together at OMERACT and this was found to be a fruitful initiative.

The 2 concepts presented in this article, comprehensibility and cultural appropriateness of PROM, are important considerations to ensure equity in trials. Despite their wide use, readability formulas, which only consider text complexity, take no account of important discourse features or nontextual dimensions such as context and cultural differences. They do not measure "top-down" factors involved in reading comprehension such as recognizing the structure and organization of a text, or "bottom-up" factors such as the density of information and appropriateness of the language. Thus, they cannot provide useful information on text comprehensibility and therefore, their utility as assessment tools for PROM is questionable. In contrast, the ELF-Q considers the overall structure and organization of a text, the clarity of function, the language and vocabulary used, as well as the content, layout, and cultural appropriateness. These considerations are well known among linguists for being important in determining a person's ability to comprehend text; and patient information that has been generated using linguistic considerations included in the ELF has been found by patients to be clearer and more effective in communicating information compared to information that has not[21]. The ELF-Q could help PROM developers ensure that their instrument is understandable and suitable for lower literacy groups by identifying aspects that could be improved. The ELF-Q could also be used during cross-cultural adaptation to increase the quality of the adaptation.

The readability assessments of the HAQ and OAKHQOL in the other languages demonstrate that the differences between languages make it difficult to use a standard readability formula to compare different versions of the same questionnaire. The readability tests we used are intended to assess English text and may not provide accurate assessments of the non-English text complexity.

Despite the existence of many different guidelines addressing the process of cross-cultural adaptation of questionnaires, there are currently no definite methods for cultural equivalence other than the ones included in the guidelines (e.g., use of an expert committee and/or a focus group of patients).

Assessing the comprehensibility of questionnaires and culturally adapting them for the intended audience requires the development of separate methodologies. Although only 6 of the 31 participants were patient research partners and none were representative of patients with low literacy, overall, equity SIG participants agreed that the literacy skills of the target population and the comprehensibility and cross-cultural adaptation of PROM are important considerations for PROM developers.

The equity SIG's longterm goal is to include comprehensibility and cross-cultural adaptation as items in the OMERACT checklist for developing core outcome measurement sets in the new Filter 2.0 handbook. This goal was considered premature for Filter 2.0 but should be considered as an option for developers of PROM in OMERACT core sets. Developers should be encouraged to think through the contextual factors of their setting and target audience, including literacy levels, populations at risk for disadvantage, and/or different cultures.

For the assessment of comprehensibility of questionnaires, we will conduct a consensus exercise on the methods required to ensure appropriateness of instruments for groups at risk for disadvantage, especially those with lower literacy levels.

Future work of the equity SIG will also include an investigation into cross-cultural adaptation methods. This will include a consensus exercise on what constitutes adequate cross-cultural adaptation for OMERACT Filter 2.0.

## ACKNOWLEDGMENT

The authors thank the OMERACT 2014 delegates who attended the SIG and contributed to the discussion. We also recognize additional members of the equity special interest working group: Laurence Carton, Kanta Kumar, Richard Osborne, and Jordi Pardo.

## REFERENCES

1. Whitehead M. The concepts and principles of equity and health. Int J Health Serv 1992;22:429-45.
2. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014;67:745-53.
3. Elder C, Barber M, Staples M, Osborne RH, Clerehan R, Buchbinder R. Assessing health literacy: A new domain for collaboration between language testers and health professionals. Lang Assess Q 2012;9:205-24.
4. UNESCO. Education for all: literacy for life. Paris, France: United Nations Educational, Scientific and Cultural Organization. Paris:

UNESCO; 2005.
5. Institute for Statistics, UNESCO. Adult and Youth Literacy: National, regional and global trends, 1985-2015. Montreal: Institute for Statistics, UNESCO; 2013.
6. Trifiletti LB, Shields WC, McDonald EM, Walker AR, Gielen AC. Development of injury prevention materials for people with low literacy skills. Patient Educ Couns 2006;64:119-27.
7. Williams MV, Baker DW, Honig EG, Lee TM, Nowlan A. Inadequate literacy is a barrier to asthma knowledge and self-care. Chest 1998;114:1008-15.
8. Adams J, Chapman J, Bradley S, Ryan SJ. Literacy levels required to complete routinely used patient-reported outcome measures in rheumatology. Rheumatology 2013;52:460-4.
9. Lowe W, Ballinger C, Protheroe J, Lueddeke J, Nutbeam D, Armstrong R, et al. Effectiveness of musculoskeletal education interventions in people with low literacy levels: a systematic review. Arthritis Care Res 2013;65:1976-85.
10. Canadian Institutes of Health Research, National Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada. Tri-council policy statement: Ethical conduct for research involving humans. CIHR; 2010.
11. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. Qual Life Res 1998;7:323-35.
12. Bowden A, Fox-Rushby JA. A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. Soc Sci Med 2003;57:1289-306.
13. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. J Clin Epidemiol 2015;68:435-41.
14. Ramey DR, Raynauld JP, Fries JF. The Health Assessment Questionnaire 1992: status and review. Arthritis Care Res 1992;5:119-29.
15. Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). Clin Exp Rheumatol 2005;23:S14-8.
16. Rat AC, Coste J, Pouchot J, Baumann M, Spitz E, Retel-Rude N, et al. OAKHQOL: a new instrument to measure quality of life in knee and hip osteoarthritis. J Clin Epidemiol 2005;58:47-55.
17. Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. Millington, Tennessee: Naval Air Station Memphis; 1975.
18. Mclaughlin GH. SMOG grading — New readability formula. J Reading 1969;12:639-46.
19. Flesch R. A new readability yardstick. J Appl Psychol 1948; 32:221-33.
20. Clerehan R, Buchbinder R, Moodie J. A linguistic framework for assessing the quality of written patient information: its use in assessing methotrexate information for rheumatoid arthritis. Health Educ Res 2005;20:334-44.
21. Hirsh D, Clerehan R, Staples M, Osborne RH, Buchbinder R. Patient assessment of medication information leaflets and validation of the Evaluative Linguistic Framework (ELF). Patient Educ Couns 2009;77:248-54.
22. Alas AN, Bergman J, Dunivan GC, Rashid R, Morrisroe SN, Rogers RG, et al. Readability of common health-related quality-of-life instruments in female pelvic medicine. Female Pelvic Med Reconstr Surg 2013;19:293-7.
23. Epstein J, Osborne RH, Elsworth GR, Beaton DE, Guillemin F. Cross-cultural adaptation of the Health Education Impact Questionnaire: experimental study showed expert committee, not back-translation, added value. J Clin Epidemiol 2015;68:360-9.
24. Zraick RI, Atcherson SR. Readability of patient-reported outcome questionnaires for use with persons with dysphonia. J Voice 2012;26:635-41.
25. Bergman J, Gore JL, Singer JS, Anger JT, Litwin MS. Readability of health related quality of life instruments in urology. J Urol 2010;183:1977-81.
26. Pace CC, Atcherson SR, Zraick RI. A computer-based readability analysis of patient-reported outcome questionnaires related to oral health quality of life. Patient Educ Couns 2012;89:76-81.
27. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000;25:3186-91.
28. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417-32.
29. Kulis D, Arnott M, Greimel ER, Bottomley A, Koller M; European Organisation for Research and Treatment of Cancer Quality of Life Group. Trends in translation requests and arising issues regarding cultural adaptation. Expert Rev Pharmacoecon Outcomes Res 2011;11:307-14.
30. Leplège A, Verdier A. The adaptation of health status measures: methodological aspects of the translation procedure. The international assessment of health-related quality of life: theory, translation, measurement and analysis. Oxford, UK: Rapid Communication of Oxford; 1995:95-101.
31. McKenna SP, Doward LC. The translation and cultural adaptation of patient-reported outcome measures. Value Health 2005;8:89-91.
32. Hedman AS. Using the SMOG formula to revise a health-related document. Am J Health Educ 2008;39:61-4.
33. Browne C, Culligan B, Phillips J. The new general service list. 2013. [Internet. Accessed April 30, 2015.] Available from: www.newgeneralservicelist.org
34. Wiktionnaire: Liste de 1750 mots français les plus courants. [Internet. Updated July 13, 2012. Accessed April 30, 2015.] Available from: http://fr.wiktionary.org/wiki/Wiktionnaire:Liste_de_1750_mots_fran%C3%A7ais_les_plus_courants.
35. WikiWoordenboek: Lijst met 1000 basiswoorden. [Internet. Updated Dec 1, 2012; accessed April 30, 2015.] Available from: http://nl.wiktionary.org/wiki/WikiWoordenboek:Lijst_met_1000_basiswoorden.
36. Wiktionary: Frequency lists/Turkish wordlist. [Internet. Updated August 31, 2011; accessed April 30, 2015.] Available from: http://en.wiktionary.org/wiki/Wiktionary:Frequency_lists/Turkish_wordlist.

**APPENDIX 1.** Literature review search strategy and results summary.

| | |
|---|---|
| Methods terms | |
| 1 | *"Outcome Assessment (Health Care)"/ (20924) |
| 2 | *Clinical Trials as Topic/mt (5031) |
| 3 | Psychometrics/mt (4322) |
| 4 | Validation.tw. (106397) |
| 5 | Validation Studies as Topic/ (1554) |
| 6 | Face validity.tw. (1680) |
| 7 | Content validity.tw. (3520) |
| 8 | Construct validity.tw. (10814) |
| 9 | Concurrent validity.tw. (3908) |
| 10 | Convergent validity.tw. (3441) |
| 11 | Discriminant validity.tw. (3381) |
| 12 | or/1-11 (151626) |
| Questionnaires/Instrument terms | |
| 13 | *Questionnaires/ (29828) |
| 14 | tool.mp. or toolkit.tw. (281574) |
| 15 | Checklist.tw. (17906) |
| 16 | instrument.tw. (78808) |
| 17 | survey.tw. (329220) |
| 18 | *Evaluation Studies as Topic/ (6233) |
| 19 | or/13-18 (704783) |
| Readability/Cultural Adaptation/Literacy terms | |
| 20 | *Information Literacy/ (71) |
| 21 | *Health Literacy/ (1098) |
| 22 | *Cultural Characteristics/ (4999) |
| 23 | Quality of Life/ (120720) |
| 24 | *Cultural Diversity/ (5147) |
| 25 | *Language/ (15378) |
| 26 | Patient Education as Topic/mt [Methods] (13508) |
| 27 | *Reading/ (10294) |
| 28 | *Writing/ (8064) |
| 29 | *Linguistics/ (3192) |
| 30 | (print$ adj2 (question$ or information or instruction$ or advice or advise$ or educat$)).tw. (680) |
| 31 | (written adj2 (information or question$ or instruction$ or advice or advise$ or educat$)).tw. (4444) |
| 32 | *Comprehension/ (3913) |
| 33 | readability.tw. (1566) |
| 34 | (cultur$ adj3 adaptat$).tw. (2033) |
| 35 | or/20-34 (188422) |
| Systematic review filter | |
| 36 | meta analysis.mp,pt. (82839) |
| 37 | review.pt. (1925475) |
| 38 | search:.tw. (257076) |
| 39 | or/36-38 (2125371) |
| All concepts combined (no study design filter): | |
| 40 | 12 and 19 and 35 (5753) |
| With McMaster systematic review filter: | |
| 41 | 39 and 40 (394) |