

Authors

S. Robin Cohen, Richard Sawatzky, Lara B. Russell, Javad Shahidi, Daren K. Heyland, Anne M. Gadermann

Correspondence: robin.cohen@mcgill.ca

Title

Measuring the quality of life of people at the end of life: The McGill Quality of Life Questionnaire - Revised (MQOL-R)

Keywords quality of life; palliative care; end of life care; psychometrics; emotional adjustment; existentialism; spirituality; family relations; chronic disease

Abstract

Background

The McGill Quality of Life Questionnaire (MQOL) has been widely used with people with life-threatening illnesses without modification since its publication in 1996. With use, areas for improvement have emerged, therefore various minor modifications were tested over time.

Aim

To revise MQOL [MQOL-Revised (MQOL-R)] while maintaining or improving its psychometric properties and length, keeping it as close as possible to MQOL to enable reasonable comparison with existing MQOL literature.

Design

Datasets from 8 studies were used (4 studies originally used to develop MQOL, 2 to develop new MQOL versions, and 2 with unrelated purposes). MQOL-R was developed using analyses of measurement invariance, confirmatory factor analysis (CFA), and calculation of correlations with MQOL's global QOL item.

Setting/Participants

Data were from 1702 people with life-threatening illnesses recruited from acute and palliative care units, palliative home care services, and oncology and HIV/AIDS outpatient clinics.

Results

MQOL-R consists of 14 items (plus the global QOL item). A new Physical subscale was created combining physical symptoms and physical wellbeing and a new item on physical functioning. The Existential subscale was reduced to 4 items. The revised Support subscale, renamed Social, focuses more on relationships. The Psychological subscale remains unchanged. CFA results provide support for the measurement structure of MQOL-R. The overall scale has good internal consistency reliability ($\alpha = .94$).

Conclusions

MQOL-R improves on and can replace MQOL since it contains improved wording, a somewhat expanded repertoire of concepts with fewer items, and a single subscale for the physical domain, while retaining good psychometric properties.

Key statements

What is already known about this topic?

- The primary goal of palliative care is to improve QOL.
- The McGill Quality of Life Questionnaire (MQOL) is widely used in palliative and end-of-life care studies. It has not been updated since its publication in 1996.

What this paper adds?

- Use of MQOL over 2 decades suggests some aspects of MQOL could be improved.
- The development of MQOL-Revised (MQOL-R) and its psychometric properties is reported.
- MQOL-R is a 14-item self-report questionnaire with good psychometric properties that measures the physical, psychological, existential, and social domains.
- MQOL-R can replace MQOL.

Implications for practice, theory or policy?

- MQOL-R can be used in studies to describe QOL. Since it is based on MQOL it is likely to also be useful for evaluating the impact of interventions and services aimed at improving the quality of life of people with life-limiting illnesses, although further testing is required.
- MQOL-R's usefulness in other populations and in clinical practice remains to be tested.

Background

Since optimizing quality of life (QOL) through high-quality care is the primary goal of palliative care,¹ psychometrically sound measures of QOL are important. Many QOL instruments developed for a general population or for people with specific diseases such as cancer or heart failure are too long for those who are in a weakened state, contain items that are not relevant to, or inappropriate for, many people who are at the end of life (e.g., items about being as healthy as ever, ability to work or engage in vigorous exercise), and/or do not include domains that are important contributors to the QOL of people at the end of life (e.g., existential or spiritual domains). Nonetheless, such QOL instruments continue to be used in studies of people at the end of life (e.g. the EORTC-QLQ-C30²).³ Because of their limitations, several QOL instruments have been developed specifically for the end-of-life population. Two were developed based on instruments widely used in oncology. The EORTC-QLQ-15 PAL (based on the EORTC-QLQ-C30) is a 15-item instrument derived from the 30 items in the parent instrument.⁴ While at 15 items it is brief, like the parent instrument it is mainly a measure of the physical domain. The Functional Assessment of Chronic Illness Therapy – Palliative care instrument⁵ (FACIT-PAL; based on the Functional Assessment of Cancer Therapy-General⁶) has more comprehensive and balanced coverage of domains relevant to an end-of-life population but, with 46 items, it is long. A preliminary study to develop a 14-item version has recently been published (FACIT-PAL-14) but its psychometric properties have not yet been examined.⁷ Another set of measures, the Palliative Outcome Scale (POS) family of instruments⁸⁻¹⁰, measures a broad range of outcomes important to palliative care. They are scored either as individual items or a summary score. However, there are no subscales covering different domains. The McGill Quality of Life Questionnaire (MQOL) was developed specifically to measure the QOL of people with a life-threatening illness.¹¹⁻¹³ Reviews of QOL instruments for palliative care have frequently recommended MQOL over the years and, after the EORTC-QLQ-C30, it is the most widely used measure to assess QOL in palliative care.^{3, 14, 15}

MQOL, first published in 1996, is not derived from a parent instrument developed primarily for earlier stage serious illnesses but rather was developed expressly to measure the QOL of people at the end of life. It was designed to address the following specific concerns with existing QOL measures at that time, which: 1) were too long for many palliative care patients; 2) did not measure existential or spiritual wellbeing; and 3) focused exclusively on negative contributors to QOL, whereas QOL is influenced by both positive and negative contributors.^{11-13, 16} Long lists of physical symptoms in these existing questionnaires also meant that the signal from those that were most important to an individual was lost in the noise created by many items measuring less relevant symptoms. It was intended that MQOL users with a particular interest in physical symptoms would add a brief symptom measure such as the Edmonton Symptom Assessment System.^{17, 18} Further rationale for the development of MQOL can be found in the original publications.^{11-13, 16} MQOL has been used in over 120 scientific articles and translated into 20 languages. It has been used with various end-of-life populations (e.g. end-stage renal disease, ALS) as well as other populations (e.g. seniors).

The original MQOL contains 16 items measuring 4 domains: physical (symptoms and general physical wellbeing); psychological; existential; and support. It also includes the MQOL Single-Item Scale (SIS) measuring global (overall) QOL. MQOL studies have demonstrated the importance of the existential domain to the quality of life (QOL) of people at the end of life.^{12, 13, 16, 19} Furthermore, MQOL studies have shown that while physical aspects of QOL are worse in those with more advanced disease, other aspects such as support and existential wellbeing remain positive contributors to QOL.^{16, 19} QOL can change rapidly at the very end of life; therefore MQOL has a timeframe of two days (in contrast to the more typical 1 week or 1 month) so that respondents do not need to average across a number of days where QOL may be very different. All MQOL items use a 0-10 response scale with verbal anchors at each end.¹²

This study is part of ongoing efforts to further evaluate and improve MQOL. The aim was to create a revised version of MQOL (MQOL-R) using several datasets with experimental versions of MQOL containing reworded or new items and/or with some items deleted. The goal was to address issues that have arisen during use of the MQOL over the years, while maintaining or improving its psychometric properties and length (feasibility) and keeping MQOL-R as close as possible to the original instrument in order to enable reasonable comparison with the body of literature using MQOL. We specifically considered the following possibilities:

1. Rewording items
2. Replacing items
3. Adding new items
4. Removing redundant items
5. Collapsing subscales

In developing a self-report instrument intended for people who have little stamina, since some psychometric properties (e.g., reliability) tend to improve with more items, the design that is psychometrically ideal must be balanced against length.²⁰

Some domains that are relevant to the QOL of people with life-threatening illness are not included in MQOL.²¹⁻²³ However, we did not consider adding domains in MQOL-R since this would make the instrument less comparable to the existing MQOL literature and increase MQOL's length. Those domains will be considered for inclusion in an extended version of MQOL that we are also developing (MQOL-Expanded).

Method

This study is a secondary analysis of data collected with end-of-life populations in 6 studies, plus 2 others collected to develop new versions of MQOL (Datasets G and H, Table 1). These studies used both French and English versions of MQOL (originally developed simultaneously). Reworded and new items were also developed and tested simultaneously in both languages. Combining the data from multiple studies allowed us to develop MQOL-R using a large dataset, but also added complexity to the data in two ways. First, in the various studies, variations of MQOL were used in an attempt to address distributional challenges or a lack of clarity in the wording of some items, and to improve representation of MQOL domains. As a result, not all MQOL items were the same across studies. Second, the studies used different end-of-life populations. To address this complexity, we pooled all the datasets into a common dataset that included all samples and original, reworded and new items. This allowed us to evaluate the potential impact of reworded or new items in the context of the original items within the pooled dataset (as is further described below).

Datasets

Data from 4 different samples of people with life-threatening illnesses in Canadian healthcare settings were used (“new datasets”) and compared to data from 4 studies used for the original MQOL validation (“original datasets”). Details of the samples are given in Tables 1 and 2. In all studies except E, patients had the choice of completing the questionnaire on their own or having it read aloud to them. Because dataset H included only reworded MQOL items and new items, it was used only for examining objective 1 (see Table 3). All studies were approved by local ethics committees (see list in Acknowledgements). Participants in all studies except E provided written informed consent. Study E provided potential participants with an information sheet and a questionnaire which was completed anonymously if they chose to participate. The dataset can be found at <http://figshare.com/s/30fa34907e4611e5a45006ec4b8d1f61>.

Analysis

Various statistical approaches with corresponding decision rules were used to iteratively examine each of our analytical objectives (see Table 3). To investigate the construct validity of the MQOL-R, we examined its measurement structure and relationship to global QOL. Specifically, we used confirmatory factor analysis (CFA) to compare the fit of competing interim models using different items in the new datasets to the fit of the original MQOL. Then we examined the correlations of potential new and revised subscales with global QOL as measured by the MQOL SIS (included in MQOL-R). Through this process, a candidate MQOL-R was constructed, which was subsequently confirmed using CFA across and within the original and new datasets. All models were identified by fixing one of the factor loadings for each subscale at 1.²⁴ Acceptable model fit was indicated by a RMSEA of < .06, a CFI of $\geq .95$, and theoretical considerations.²⁵ Internal consistency reliability of the subscales was assessed by Cronbach’s alpha (α). Alpha > 0.7 is generally considered acceptable.²⁶

Multiple imputation (MI) and full information maximum likelihood (FIML) methods were used for different aspects of the analyses to account for missing data across the combined datasets.^{27,28} MI was applied only to the original MQOL items to allow for analyses of new datasets where these items were not included. New items that were not part of the original MQOL were not imputed. This allowed for the comparison of models that included or excluded new items against models based on the original MQOL items. The new and original datasets used in the final MQOL-R analyses had 17% and 22% missing data, respectively (excluding 33 cases with no responses to any MQOL-R items).

Results

Developing the new MQOL-R

Comparisons of each reworded item to the original item (see objective 1 in Table 3) revealed that the reworded items did not lead to substantial psychometric improvements (model fit and factor loadings). Thus the original item wordings were retained, except for the item that measures control over one's life. Although there was no statistically significant difference in the standardized factor loading for the reworded control item relative to the original wording, the revised wording was retained because the original wording assumes that people want to have control over their lives, which may not be true for everyone. The new item instead measures *how problematic* the amount of control is.

The wording of MQOL-R items is shown in Figure 1. The results pertaining to the remaining analytical objectives (2-4) are reported in Table 4 and organized by subscale.

Social subscale replaces Support subscale. The original MQOL Support subscale includes the item "The world is caring and responsive to my needs." Because the interpretation of "the world" has at times been interpreted as the whole world rather than the world that interacts with the respondent (unpublished data), it was not included in the MQOL used in studies E-H. Because it would therefore need to be almost wholly imputed in the new datasets, it was not included in further analyses. At the same time, interpersonal relationships, which are important to QOL,^{21, 29} are less well represented in MQOL. The Support subscale was replaced with a Social subscale that retains one Support item, excludes the "world is caring" item, and adds two new items that focus on relationships. A CFA of the measurement model including this subscale, tested using datasets E-G, revealed acceptable fit with the data (RMSEA = 0.044; CFI = .95). It was therefore retained for MQOL-R and used in place of the Support subscale in all subsequent analyses.

Physical subscale replaces Physical Symptoms and Physical Wellbeing subscales. For the Physical subscale, the CFA revealed acceptable fit for the measurement model of correlated latent factors (subscales) when replacing the three physical symptom items with a single item measuring how problematic all symptoms are when considered together (RMSEA = .044, CFI = .95). In addition, whereas previously there were two subscales for the physical domain, a single Physical subscale was created that includes the new item concerning the impact of all physical symptoms combined, the original physical wellbeing item, and a new physical functioning item, which resulted in acceptable model fit (RMSEA = .040, CFI = .94).

Reduced Existential subscale. Two items were removed from the Existential subscale, reducing it to four items, without substantial loss in model fit based on CFA results (see Table 4).

Psychological subscale unchanged. There was no attempt to reduce the 4-item Psychological subscale.

Confirmatory factor analysis of the MQOL-R

Once the revisions above had been explored, the final MQOL-R with four correlated latent factors was fit both using all datasets and separately using new datasets E-G, which resulted in acceptable fit (see Table 5). Factor loadings with the new datasets E-G ranged from 0.53 to 0.89 and included 13 factor loadings at 0.63 or higher. The latent factor correlations ranged from 0.28 to 0.63. In addition, a second-order factor model was specified (see Figure 1) to support deriving a MQOL-R total score.

Reliability and correlations with global QOL

Internal consistency reliability (Cronbach's alpha) is good for the overall scale ($\alpha = .94$) and acceptable ($\alpha > .7$) for most subscales (Psychological $\alpha = .85$; Existential $\alpha = .78$; Social $\alpha = .87$). However, the physical subscale is slightly less reliable ($\alpha = .66$). This is not surprising considering that this subscale combines different aspects of the physical domain (physical symptoms, wellness, and functioning) that may not necessarily correlate in a consistent fashion. Nonetheless, the Physical subscale (measured as a latent factor) has the strongest correlation with global QOL ($r = .65$; as measured using the single item). The other subscales were correlated with global QOL at .43 (Psychological), .52 (Existential), and .29 (Social), and the overall MQOL-R second-order factor had a correlation of .67.

Discussion

The MQOL-R covers a somewhat expanded repertoire of concepts (adding physical functioning, communication and relationships with people you care about) within the same general QOL domains as the original MQOL (physical, psychological, existential, social) with a reduction in the number of items from 16 to 14 (plus the global QOL item). Our results provide evidence for the construct validity of the MQOL-R by demonstrating a well-fitting measurement structure (tested with CFA) and expected correlations between each MQOL-R subscale and the item measuring global QOL. Internal consistency estimates of the subscales also remain similar to the original with the exception of the Existential Wellbeing subscale, where α is reduced, but which nevertheless remains acceptable ($\alpha = 0.78$). This is likely because alpha tends to be lower with fewer items. Alpha for the original Physical Symptoms subscale was always somewhat low at approximately 0.66; the new Physical subscale that replaces it has a similar α .

The MQOL-R improves on MQOL in several ways. First, the MQOL-R combines the 3 items from the physical domain into a single subscale; these cover physical wellbeing in general, physical symptoms, and physical functioning. This is an improvement over the MQOL because: 1) the physical domain is represented by a single subscale; 2) physical symptoms are captured in a single item encompassing all symptoms, thereby avoiding the issues related to the respondent having fewer or more than 3 physical symptoms that can arise with the MQOL; 3) the impact of reduced physical functioning is included. Second, the wording of the item about control over one's life is conceptually clearer. Third, the social domain more explicitly focuses on relationships. Finally, all of this has been achieved with a reduction by 2 items.

The MQOL-R has some advantages compared to other QOL instruments or general outcome measures designed for use in people with a life-threatening illness.^{4, 7-9, 30, 31} While the MQOL-R has validated subscales for the physical, psychological, existential, and social domains, the EORTC-QLQ-C15-PAL, FACIT-PAL-14, and POS instruments do not. The FACIT-PAL includes the FACT-G (which has physical, psychological, social, and functional subscales), but treats the palliative care items as a single subscale that includes items from several domains. A summary score is sometimes required; the EORTC-QLQ-C15-PAL does not have one. MQOL-R's summary score differs from that of the FACIT-PAL and POS instruments: MQOL-R's weights the four domains equally by calculating the mean of the subscale scores; those of the FACIT-PAL and POS instruments reflect the unequal number of items they include for each domain, meaning that the instrument developer has imposed different weights for each domain.^{4, 7, 32-34} In addition, it is often of interest to also measure QOL with a single global item that is context-free, in order to capture the respondent's *own* understanding of overall QOL, rather than what is selected to be measured or not measured in the rest of the QOL instrument. We therefore believe

that the global item should be presented before and separately from any other items, as it is in MQOL-R. In the EORTC-QLQ-C15-PAL the global item is at the end of the instrument; in the FACIT and FACIT-PAL-14, it is in the middle (the POS instruments do not have a global item). In terms of length, at 14 items plus one global item the MQOL-R is about the same length as the POS (11 items), EORTC-QLQ-C15-PAL (15 items) and FACIT-PAL-14 (14 items), but shorter than the Integrated-POS being developed (20 items if each of the physical symptoms is considered a separate item) and the FACIT-PAL (46 items).

The strength of our conclusions must be tempered by some of the limitations of the dataset, which was a combination of datasets collected for several different primary purposes. Different populations, and a different timeframe in dataset E, made it difficult to determine whether differences between studies are due to different items being used or different populations or timeframes being studied. However, we addressed this challenge by using methods of analyzing measurement invariance to account for differences in study populations when examining the measurement structure with items that were identical across study populations. In addition, because not all of the items were used in all studies, MI and FIML, well-established procedures for dealing with missing data, were used to minimize the effect of missing data.

We feel that since MQOL-R is built on a large and heterogeneous new dataset as well as the original MQOL, it is ready for use. We recommend MQOL-R be used instead of MQOL. However, there may be some situations where investigators might prefer MQOL, therefore it will remain available. Further studies of the psychometric properties of MQOL-R are warranted, including but not limited to conducting a CFA in a new population and evaluating the test-retest reliability and sensitivity to change. Other studies will examine the feasibility of developing short and long forms of MQOL-R with acceptable psychometric properties.

Acknowledgements

We are grateful to Geneviève Taylor and Stephanie Carpentier who did the cognitive interviewing and its analysis that informed some of the wording changes incorporated in MQOL-R.

The following ethics committees and hospitals in Canada approved the original studies that collected the data used in this report.

Dataset A: Royal Victoria Hospital, Hôpital Notre Dame, University of British Columbia

Dataset B: Royal Victoria Hospital, Montreal, Quebec

Dataset C: Royal Victoria Hospital

Dataset D: Royal Victoria Hospital

Dataset E: Jewish General Hospital, Montreal, Quebec #09-084

Dataset F: Queen's University DMED-1029-07, McMaster University, Dalhousie University, University of Alberta, University of British Columbia, Hôpital Maisonneuve-Rosemont

Dataset G: McGill University #A-08-B32-01B, Jewish General Hospital, McGill University Health Centre #97-004, SCO Health Service, Ottawa Hospital 2001503-01H, Queen's University ONGY-251-04, University of British Columbia and Kelowna General Hospital B01-0668, University of Saskatchewan BSC #01-226

Dataset H: McGill University A05-E04-97, Montreal General Hospital #97-004, Royal Victoria Hospital 399-779, Jewish General Hospital A06-B07-97, St. Mary's Hospital Centre #97-13, CLSC N.D.G./Montréal Ouest, Mount Sinai Hospital Centre, Centre Hospitalier de l'Université de Montréal #97.80, Hôpital Notre Dame de la Merci, Maison Michel Sarrazin, University of Toronto #2838, Providence Centre (Scarborough, Ontario), University of Saskatchewan, University of British Columbia #B97-0301, Victorian Order of Nurses for Canada

Funding

The initial project to revise MQOL and collection of dataset H was funded by the Canadian Cancer Society Research Institute (CCSRI) (#8057) with funds donated to the Canadian Cancer Society, as was data collection for dataset H (#012201). Dataset E was collected with no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Funding for the other datasets is acknowledged in the primary publications from those grants.^{12, 13, 16, 35, 36}

Conflict of interest

The Authors declare that there are no conflicts of interest.

REFERENCES

1. WHO. W.H.O. Definition of Palliative Care. World Health Organization, 2015.
2. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*. 1993; 85: 365.
3. Stiel S, Pastrana T, Balzer C, Elsner F, Ostgathe C and Radbruch L. Outcome assessment instruments in palliative and hospice care—a review of the literature. *Supportive Care in Cancer*. 2012; 20: 2879-93.
4. Groenvold M, Petersen MA, Aaronson NK, et al. The development of the EORTC QLQ-C15-PAL: A shortened questionnaire for cancer patients in palliative care. *Eur J Cancer*. 2006; 42: 55-64.
5. Cella D. FACIT manual. *Manual of the Functional Assessment of Chronic Illness Therapy (FACIT) Scales Version*. 1997; 4.
6. Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *J Clin Oncol*. 1993; 11: 570-9.
7. Lyons KD, Bakitas M, Hegel MT, Hanscom B, Hull J and Ahles TA. Reliability and Validity of the Functional Assessment of Chronic Illness Therapy-Palliative Care (FACIT-Pal) Scale. *J Pain Symptom Manage*. 2009; 37: 23-32.
8. Hearn J and Higginson IJ. Development and validation of a core outcome measure for palliative care: the palliative care outcome scale. Palliative Care Core Audit Project Advisory Group. *Qual Health Care*. 1999; 8: 219-27.
9. Cicely Saunders Institute. Palliative care Outcome Scale, <http://pos-pal.org/maix/> (2015, Accessed July 5, 2016).
10. Hearn J, Higginson IJ and Projec PCCA. Development and validation of a core outcome measure for palliative care: the palliative care outcome scale. *Qual Health Care*. 1999; 8: 219-27.
11. Cohen SR, Mount BM, Strobel MG and Bui F. The McGill Quality of Life Questionnaire: a measure of quality of life appropriate for people with advanced disease. A preliminary study of validity and acceptability. *Palliat Med*. 1995; 9: 207-19.
12. Cohen SR, Mount BM, Tomas JJ and Mount LF. Existential well-being is an important determinant of quality of life. Evidence from the McGill Quality of Life Questionnaire. *Cancer*. 1996; 77: 576-86.
13. Cohen SR, Mount BM, Bruera E, Provost M, Rowe J and Tong K. Validity of the McGill Quality of Life Questionnaire in the palliative care setting: A multi-centre Canadian study demonstrating the importance of the existential domain. *Palliat Med*. 1997; 11: 3-20.
14. Bruley DK. Beyond reliability and validity: analysis of selected quality-of-life instruments for use in palliative care. *J Palliat Med*. 1999; 2: 299-309.
15. Albers G, Ehteld MA, de Vet HC, Onwuteaka-Philipsen BD, van der Linden MH and Deliens L. Evaluation of quality-of-life measures for use in palliative care: a systematic review. *Palliat Med*. 2010; 24: 17-37.
16. Cohen SR, Hassan SA, Lapointe BJ and Mount BM. Quality of life in HIV disease as measured by the McGill Quality of Life Questionnaire. *AIDS*. 1996; 10: 1421-7.
17. Bruera E, Kuehn N, Miller MJ, Selmser P and Macmillan K. The Edmonton Symptom Assessment System (ESAS): A simple method for the assessment of palliative care patients. *J Palliat Care*. 1991; 7: 6-9.
18. Watanabe SM, Nekolaichuk CL and Beaumont C. The Edmonton Symptom Assessment System, a proposed tool for distress screening in cancer patients: development and refinement. *Psychooncology*. 2012; 21: 977-85.

19. Cohen SR and Mount BM. Living with cancer: "good" days and "bad" days--what produces them? Can the McGill quality of life questionnaire distinguish between them? *Cancer*. 2000; 89: 1854-65.
20. Evans CJ, Benalia H, Preston NJ, et al. The Selection and Use of Outcome Measures in Palliative and End-of-Life Care Research: The MORECare International Consensus Workshop. *J Pain Symptom Manage*. 2013; 46: 925-37.
21. Cohen SR and Leis A. What determines the quality of life of terminally ill cancer patients from their own perspective? *J Palliat Care*. 2002; 18: 48-58.
22. Padilla GV, Ferrell B, Grant MM and Rhiner M. Defining the content domain of quality of life for cancer patients with pain. *Cancer Nurs*. 1990; 13: 108-15.
23. Waldron D, O'Boyle CA, Kearney M, Moriarty M and Carney D. Quality-of-life measurement in advanced cancer: assessing the individual. *Journal of Clinical Oncology*. 1999; 17: 3603-11.
24. Byrne BM. *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge, 2013.
25. Vandenberg RJ and Lance CE. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*. 2000; 3: 4-69.
26. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Quality of Life Research*. 2013; 22: 1889-905.
27. Enders CK. Analyzing structural equation models with missing data. In: Hancock G and Mueller R, (eds.). *Structural equation modeling: A second course*. Greenwich, UK: Information Age Publishing, 2006, p. 313-42.
28. Peyre H, Leplège A and Coste J. Missing data methods for dealing with missing items in quality of life questionnaires. A comparison by simulation of personal mean score, full information maximum likelihood, multiple imputation, and hot deck techniques applied to the SF-36 in the French 2003 decennial health survey. *Quality of Life Research*. 2011; 20: 287-300.
29. Shahidi J, Bernier N and Cohen SR. Quality of life in terminally ill cancer patients: contributors and content validity of instruments. *J Palliat Care*. 2010; 26: 88-93.
30. Greisinger AJ, Lorimor RJ, Aday LA, Winn RJ and Baile WF. Terminally ill cancer patients. Their most important concerns. *Cancer practice*. 1997; 5: 147-54.
31. Zeng L, Bedard G, Cella D, et al. Preliminary Results of the Generation of a Shortened Quality-of-Life Assessment for Patients with Advanced Cancer: The FACIT-Pal-14. *J Palliat Med*. 2013; 16: 509-15.
32. Harding R, Selman L, Simms VM, et al. How to Analyze Palliative Care Outcome Data for Patients in Sub-Saharan Africa: An International, Multicenter, Factor Analytic Examination of the APCA African POS. *J Pain Symptom Manage*. 2013; 45: 746-52.
33. Siegert RJ, Gao W, Walkey FH and Higginson IJ. Psychological Well-Being and Quality of Care: A Factor-Analytic Examination of the Palliative Care Outcome Scale. *J Pain Symptom Manage*. 2010; 40: 67-74.
34. Siegert R, Selman L, Higginson IJ, et al. A Psychometric Evaluation of the Functional Assessment of Chronic Illness Therapy-Palliative Care (FACIT-Pal) Scale With Palliative Care Samples in Three African Countries. *J Pain Symptom Manage*. 2014; 48: 983-91.
35. Heyland DK, Cook DJ, Rocker GM, et al. The development and validation of a novel questionnaire to measure patient and family satisfaction with end-of-life care: the Canadian Health Care Evaluation Project (CANHELP) Questionnaire. *Palliat Med*. 2010; 24: 682-95.
36. Cohen SR and Mount BM. Living with cancer: "Good" days and "bad" days - What produces them? Can the McGill Quality of Life Questionnaire distinguish between them? *Cancer*. 2000; 89: 1854-65.

Table 1: Datasets

Dataset	Sample size (n)	Disease(s)	Setting(s)
<i>Original datasets (N = 595)</i>			
A ¹³	143	cancer	75% palliative care units 25% palliative home care
B ¹⁶	107	HIV/AIDS before the existence of anti-retrovirals	outpatient clinics
C ¹²	247	cancer	outpatient clinics
D ^{a36}	98	cancer	53% oncology clinics 47% palliative home care
<i>New datasets (N = 1107)</i>			
E	219	cancer	outpatient clinics
F ^{b35}	368	various end-stage diseases	71% acute care units 22% home care 7% palliative care units
G ^c	216	cancer	71% palliative care units 27% palliative home care
H ^c	204	cancer	69% palliative care units 30% palliative home care

^a Each participant completed MQOL multiple times, date not noted. One questionnaire randomly selected for each participant.

^b Timeframe of 1 month rather than 2 days.

^c Longitudinal study. The first questionnaire completed by each participant was selected to provide the largest n (*i.e.* before dropout).

Table 2. Demographics

Variables	Original samples				New samples			
	A	B	C	D	E	F	G	H
Gender (% female)	61	6	58	59	64	48	51	54
Age (mean)	65	N/A	51	59	58	77	66	“Over 65” ^b
Highest level of education (%) ^a								
Did not complete high school	48	N/A	8	33	10	46	32	30
High school completed but no post-secondary	22	N/A	35	22	34	22	20	26
Some post-secondary	26	N/A	46	45	56	30	42	42

^aFor dataset E, participants reported the highest level of education attended. For other datasets, participants reported the highest level of education completed.

^bAge range only recorded, median is reported.

Table 3: Analysis objectives and plan

Analytical objectives	Analyses conducted	Decision rules
1. Determine whether a reworded item is an improvement over the original	Each reworded item was compared to the original item by conducting measurement invariance analyses across the different data sources	A reworded item was retained if: (a) The standardized loading was greater than 0.4, (b) The specified invariance of its factor loading resulted in a statistically significant difference in model fit, and (c) The difference in standardized loadings with the non-invariance model was greater than the maximal difference observed for non-reworded items (thereby accounting for differences in the samples).
2. Determine whether to add a new item to an existing domain or replace an existing item	• Confirmatory factor analyses (CFAs) • Correlations with the global QOL item	A new item was retained if: (a) The standardized loading was greater than 0.4, (b) There was no loss in model fit, (c) Including the new item in the subscale improved the subscale's correlation with the global QOL item by a magnitude greater than a small effect size (<i>i.e.</i> $\Delta r > .1$).
3. Reconstruct the support and physical subscales ^a	• CFAs • Correlations with the global QOL item	The reconstructed subscale was retained if: (a) The standardized loadings were greater than 0.4, (b) There was no loss in model fit, (c) There was no reduction in the correlation of the reconstructed domain with the global QOL item of a magnitude greater than a small effect size (<i>i.e.</i> $\Delta r > .1$)
4. Reduce the number of items	• CFAs • Correlations with the global QOL item	For conceptually similar items, the item with the greatest factor loading was retained. In all other cases, an item was removed if: (a) There was no loss in model fit, and (b) There was no reduction in the correlation of the reconstructed domain with the global QOL item of a magnitude that is greater than a small effect size (<i>i.e.</i> $\Delta r > .1$).

Note.

^a Rationale provided in Table 4

Table 4: MQOL subscale reconstruction process

Item tested (paraphrased)	Modifications considered	Results and decision
Support/Social		
I felt my relationships with the people I care about were more distant than I would like	Remove the item “The world is caring and responsive to my needs” and add one or more items that focus on relationships. <u>Rationale:</u> Cognitive interviewing indicated that the original MQOL item “The world is caring and responsive to my needs” was interpreted by some respondents to include such a broad definition of “the world” that it did not reflect the intended meaning of the world that interacts with the respondent. Relationships are very important to QOL ⁷ but are not explicitly represented in MQOL.	<u>Results^a</u> A CFA of the measurement model replacing the original Support subscale with the new 3-item relationship subscale resulted in acceptable fit (RMSEA = .044, CFI = .95). <u>Decision</u> Retaining these items together with one of the original Support items resulted in a psychometrically sound Social subscale.
Communication with the people I care about was difficult		
Physical		
My physical symptoms were problematic	Could a single item replace the three identical physical symptom items? <u>Rationale:</u> In MQOL, respondents are asked to list their three most problematic symptoms and indicate how much of a problem they are; if they have < 3 symptoms, they are to write “none” for extra items and the item is given a score of 0. When some of these three items are left blank, it cannot be determined whether there are < 3 symptoms and they forgot to write “none” or if the item was skipped. In addition, a single item allows the symptom score to reflect the respondent’s direct rating of the combined impact of all physical symptoms together, rather than arbitrarily using the mean of three.	<u>Results</u> Comparison of the single physical symptom item versus the combined three physical symptom items revealed no substantial difference in fit of CFAs of the corresponding MQOL measurement models (RMSEA = .044, CFI = .95 versus RMSEA = .046, CFI = .94). <u>Decision</u> The three physical symptom items were replaced with the single item.
Being unable to do the things I wanted was problematic	Creation of a physical subscale that includes physical wellbeing, the single physical symptom item, and a new item on physical functioning. <u>Rationale:</u> Physical functioning is widely acknowledged to be important to QOL and was found to be a contributor to the QOL of palliative care patients in a qualitative study. ⁷	<u>Results</u> The MQOL model including the new subscale fit well (RMSEA = .040, CFI = .94). Standardized factor loadings range from .40 to .84. <u>Decision</u> Accept the new subscale.

Existential		
When I thought about my life, I felt that my life to this point has been: completely worthless - very worthwhile	Consider removing this original item because it is highly correlated with the item on progress in achieving life goals. <u>Rationale:</u> Reduce number of items.	<u>Results</u> There was no substantial difference in fit when both items were removed (RMSEA = .041, CFI= .96 for 4 items vs RMSEA = .046, CFI= .94 for 6 items).
The past two days were: a burden - a gift	Burden and gift are not antonyms, making this a good candidate for removal. Also, cognitive interviewing found that “a gift” could imply the existence of a higher being, limiting its appropriateness for people who do not believe in a higher being. <u>Rationale:</u> Reduce number of items and remove a potentially conceptually problematic item.	<u>Decision</u> The items were removed for the purpose of reducing the length of the instrument.

Note.

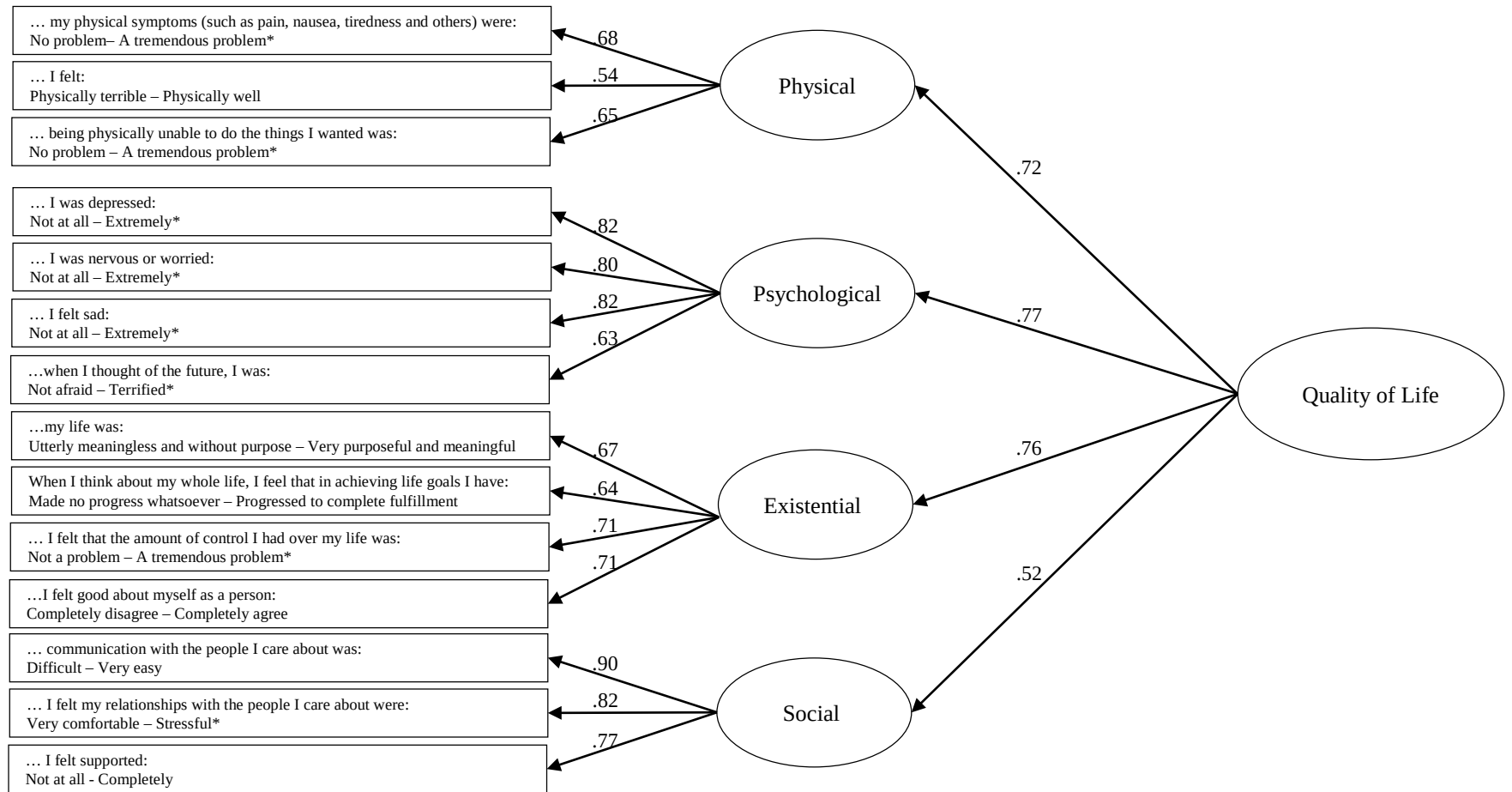
^a All analyses are based on samples E-G (n = 770, after cases missing all MQOL data are excluded).

Source H was excluded because this data source was used to test reworded items and did not include any of the original items. It was only used to examine the potential impact of rewording.

Table 5: MQOL-R model fit within datasets

	N	CFI	RMSEA	Chi-square (Df)
1) Correlated factor model				
a) New datasets (E-G)	770	0.96	0.041	164.71 (71)
b) All datasets (Excluding H)	1365	0.95	0.043	252.18 (71)
2) Second-order factor model				
a) New datasets (E-G)	770	0.95	0.046	191.51 (73)
b) All datasets (Excluding H)	1365	0.94	0.047	288.57 (73)

Figure 1: MQOL-R items and CFA for first (subscale) and second order (overall QOL) latent factors



Notes.

- Model fit based on full information maximum likelihood and using the new data sources, excluding dataset H which included revised items that were not selected in the final model (N = 770, excluding 33 missing cases): Root mean square error of approximation = .046. Comparative fit index = .95.
- Factor loadings are standardized.
- * = reverse scored items
- ... = Over the past two days (48 hours)

