

## Introduction to special section: quantitative methods

Dennis A. Revicki · Carolyn E. Schwartz

Published online: 18 December 2014  
© Springer International Publishing Switzerland 2014

For the past 40 years, quantitative methods and psychometric analyses have been an integral component for developing and evaluating the measurement properties of patient-reported and health-related quality of life (HRQL) outcome measures. Advances in the development of new quantitative methods and in the application of these new methods have increased our understanding of the relationship among physiologic, clinical, and HRQL outcomes and improved the development and evaluation of new health outcome instruments. In this section, we are pleased to provide a good cross section of the quantitative methods applied to HRQL and other patient-reported outcome measures in this issue of *Quality of Life Research*. The twelve papers included in this section cover a range in topics from various approaches to evaluating longitudinal HRQL data, evaluating theoretical models, and applying advanced psychometric methods to understanding conceptual equivalence across countries or in analyzing PRO item-level data.

First, there are several articles summarizing advanced quantitative methods of handling longitudinal data. For

example, Anota et al. [1] summarize the issues and definitions associated with time-to-deterioration type analyses, with illustrations from early breast cancer and metastatic pancreatic cancer samples. A number of different definitions for deterioration in HRQL outcomes can be specified, and these decisions have implications for the results of the time-to-deterioration analyses. The authors provide some guidance on the use of time-to-deterioration versus time-to-definitive-deterioration. In the article by de Bock et al. [2], Rasch analysis is used to handle informative intermittent missing data for longitudinal comparisons of PRO data. They developed several simulations with varying amounts of informative and non-informative missing data and applied longitudinal Rasch mixed models and linear mixed models. The two analysis methods were comparable when there was little missing data (<15 %), but the longitudinal Rasch mixed models performed better when there was greater missing data (>15 %).

Terrin et al. [3] evaluated prediction models for transplant-related mortality based on HRQL data in a study pediatric hematopoietic stem cell transplant patients. Joint models were used to analyze the longitudinal HRQL data and the time-to-mortality data within the same statistical analysis using a single likelihood function. They found that trajectories of HRQL outcomes predicted transplant-related mortality in pediatric hematopoietic stem cell transplant patients, even after adjusting survival for baseline demographic and clinical characteristics.

The next set of several articles evaluate theoretical models for understanding the relationship between clinical and HRQL measures. Mayo et al. [4] evaluated the Wilson–Cleary model [5] in patients recovering from a recent stroke. They apply structural equation models (SEM) to examine the relationship among biological variables, symptoms, functional outcomes, and health perceptions in 533 patients

---

D. A. Revicki (✉)  
Outcomes Research, Evidera, 7101 Wisconsin Ave., Suite 1400,  
Bethesda, MD 201814, USA  
e-mail: dennis.revicki@evidera.com

C. E. Schwartz  
DeltaQuest Foundation, Concord, MA, USA

C. E. Schwartz  
Departments of Medicine and Orthopaedic Surgery, Tufts  
University School of Medicine, Boston, MA, USA

C. E. Schwartz  
Faculty of Health Sciences, Oslo and Akershus University  
College of Applied Sciences, Oslo, Norway

during the initial 3 months of stroke recovery. The final model was able to explain 76 % of the variance in health perceptions scores. This article provides an excellent example of using theory-guided SEM to evaluate a number of predictive variables for health perceptions. Eilayyan et al. [6] also based their analyses on the Wilson–Cleary model [5]. They examined whether symptom status, physical activity, beliefs about medications, self-efficacy, emotional status, and healthcare utilization predict perceived asthma control over a period of 16 months in a sample of primary care asthma patients. Path analysis was used to evaluate these relationships over time.

The next set of articles applies a number of advanced psychometric methods including bifactor models and item response theory (IRT) analyses. Bifactor models and multidimensional IRT are increasing applied to HRQL data and have several advantages in understanding the structure of PRO instruments. Paap et al. [7] use multidimensional IRT, bifactor models, and Mokken scale analysis to evaluate the dimensionality of the St. Goerge’s Respiratory Questionnaire (SGRQ) in a sample of 444 Dutch patients with chronic obstructive respiratory disease. The findings of the Mokken and multidimensional IRT models demonstrated support for a unidimensional structure for the SGRQ items, although the bifactor analysis indicated a strong general factor and provided evidence for several unique factors. A reduced 31-item version of the SGRQ was developed, after removing poorly performing items; the IRT derived score correlated strongly with the original SGRQ total scores. Deng et al. [8] also used a bifactor model to examine the general group and specific factors underlying a large set of 42 fatigue-related items derived from several legacy measures. While there was good evidence supporting an overall factor labeled as “vitality,” specific factors covering energy and fatigue were identified.

Edelen et al. [9] quantify differential item functioning (DIF) as part of IRT analyses. They define two DIF metrics: (1) a weighted area between the expected score curves; and (2) a difference in expected a posteriori scores across item response categories. The two metrics were designed to identify problematic DIF and provide useful methods for differentiating statistically significant versus problematic DIF. An example, using a cancer stigma index was used to illustrate the DIF metrics.

Twiss and McKenna [10] used Rasch analysis to co-calibrate the Psoriasis Quality of Life Questionnaire and the Quality of Life in Atopic Dermatitis scale based on five common items across the two measures. The study provides an example of a method for co-calibrating two different disease-specific measures. Previous research has used IRT and Rasch analysis to co-calibrate and link

different measures of a specific domain (e.g., physical functioning, emotional distress, etc.; see [www.prosettas-tone.org](http://www.prosettas-tone.org)).

Several articles applied multi-group confirmatory factor analysis in their studies. Regnault and Herdman [11] apply a universalist model equivalence approach to evaluating cross-cultural equivalence across language translations of PRO measures. Six types of equivalence were identified, including conceptual equivalence, item equivalence, semantic equivalence, operational equivalence, measurement equivalence, and functional equivalence. Quantitative methods are summarized for evaluating different forms of equivalence. For example, multi-group confirmatory factor analysis can be used to examine conceptual equivalence across countries, and DIF analysis can be used to evaluate item equivalence across language translations.

Costa et al. [12] evaluated the configural, metric, and scalar equivalence of the EORTC Quality of Life Questionnaire-C30 across different cancers using multi-group confirmatory factor analysis. Multi-group confirmatory factor analysis can be used to fit a measurement model for multiple groups, in this case 7 cancer diagnoses. By constraining parameters in different ways, they evaluate different hypotheses about configural, metric, and scalar equivalence. Model fit is assessed using goodness-of-fit statistics.

Reese et al. [13] used latent class analysis to compare different cultural settings on measures of symptoms, function, and supportive care needs in cancer patients from the USA, Canada, and Japan. Latent class analysis models the relationship among discrete observed variables and a categorical unobserved latent variable. This type of analysis uses patterns in the data to assign persons to the classes and allows the different measured domains to have different relationships with the derived classes. The latent class analysis methods provide an alternative way to understand similarities and differences between different defined groups on PRO and other related measures.

Taken together, this set of 12 articles demonstrates the breadth of quantitative and psychometric analyses methods applied to PRO data. They represent a range of different statistical methods for examining the relationship among HRQL outcomes and other predictors, application of modern measurement methods, and approaches for evaluating measurement equivalence of HRQL measures across disease types and countries. In the future, more advanced psychometric methods will be increasingly applied to understanding item responses and measurement characteristics, and for examining, latent constructs over time [14]. These new methods will help health outcomes researchers improve the measurement of HRQL outcomes.

## References

1. Anota, A., Hamidou, Z., Paget-Bailly, S., Chibaudel, B., Bascoul-Mollevi, C., Auquier, P., Westeel, V., Fiteni, F., Borg, C., & Bonnetain, F. (2015). Time to health-related quality of life score deterioration as a modality of longitudinal analysis for health-related quality of life studies in oncology: Do we need RECIST for quality of life to achieve standardization? *Quality of Life Research*. doi:10.1007/s11136-013-0583-6.
2. De Bock, E., Hardouin, J.-B., Blanchin, M., Le Neel, T., Kubin, G., & Sebille, V. (2015). Assessment of score and Rasch-based methods for group comparison of longitudinal patient-reported outcomes with intermittent missing data (informative and non-informative). *Quality of Life Research*. doi:10.1007/s11136-014-0648-1.
3. Terrin, N., Rodday, A. M., & Parsons, S. K. (2015). Joint models for predicting transplant-related mortality from quality of life data. *Quality of Life Research*. doi:10.1007/s11136-013-0550-2.
4. Mayo, N. E., Scott, S. C., Bayler, M., Cheung, A., Garland, J., Jutai, J., & Wood-Dauphinee, S. (2015). Modeling health-related quality of life in people recovering from stroke. *Quality of Life Research*. doi:10.1007/s11136-013-0605-4.
5. Wilson, I., & Cleary, P. (1995). Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA*, 273(1), 59–65.
6. Eilayyan, O., Gogovo, A., Mayo, N., Ernst, P., & Ahmed, S. S. (2015). Predictors of perceived asthma control among patients managed in primary care clinics. *Quality of Life Research*. doi:10.1007/s11136-014-0700-1.
7. Paap, M. C. S., Brouwer, D., Glas, C. A. W., Monninkhof, E. M., Forstreuter, B., Pieterse, M. E., & Palen, J. V. D. (2015). The St. George's Respiratory Questionnaire revisited: A psychometric evaluation. *Quality of Life Research*. doi:10.1007/s11136-013-0570-y.
8. Deng, N., Guyer, R., & Ware, J. E. (2015). Energy, fatigue or both? A bifactor modeling approach to the conceptualization and measurement of vitality. *Quality of Life Research*. doi:10.1007/s11136-014-0839-9.
9. Edelen, M. O., Stucky, B. D., & Chandra, A. (2015). Quantifying 'problematic' DIF within an IRT framework: Application to a cancer stigma index. *Quality of Life Research*. doi:10.1007/s11136-013-0540-4.
10. Twiss, J., & McKenna, S. P. (2015) Comparing the impact of psoriasis and atopic dermatitis on quality of life: Co-calibration of the PSORIQoL and QoLIAD. *Quality of Life Research*. doi:10.1007/s11136-014-0630-y.
11. Regnault, A., & Herdman, M. (2015). Using quantitative methods within the universalist model framework to explore the cross cultural equivalence of patient-reported outcome instruments. *Quality of Life Research*. doi:10.1007/s11136-014-0722-8.
12. Costa, D. S. J., Aaronson, N. K., Fayers, P. M., Pallant, J. F., Velikova, G., & King, M. T. (2015). Testing the measurement invariance of the EORTC QLQ-C30 across primary cancer sites using multigroup confirmatory factor analysis. *Quality of Life Research*. doi:10.1007/s11136-014-0799-0.
13. Reese, J. B., Blackford, A., Sussman, J., Okuyama, T., Akechi, T., Bainbridge, D., Howell, D., & Snyder, C.F. (2015) Cancer patients' function, symptoms and supportive care needs: A latent class analysis across cultures. *Quality of Life Research*. doi:10.1007/s11136-014-0629-4.
14. Reise, S. P., & Revicki, D. A. (Eds.). (2015). *Handbook of item response theory modeling: Applications to typical performance assessment*. New York: Routledge/Taylor & Francis Group.