VALUE IN HEALTH 16 (2013) 211–214

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/jval

COMMENTARY

Some Comments on Mapping from Disease-Specific to Generic Health-Related Quality-of-Life Scales

*Mari Palta, PhD**

Department of Population Health Sciences, University of Wisconsin, Madison, WI, USA

ABSTRACT

An article by Lu et al. in this issue of *Value in Health* addresses the mapping of treatment or group differences in disease-specific measures (DSMs) of health-related quality of life onto differences in generic health-related quality-of-life scores, with special emphasis on how the mapping is affected by the reliability of the DSM. In the proposed mapping, a factor analytic model defines a conversion factor between the scores as the ratio of factor loadings. Hence, the mapping applies to convert true underlying scales and has desirable properties facilitating the alignment of instruments and understanding their relationship in a coherent manner. It is important to note, however, that when DSM means or differences in mean DSMs are estimated, their mapping is still of a measurement error–prone predictor, and the correct conversion coefficient is the true mapping multiplied by the reliability of the DSM in the relevant sample. In addition, the proposed strategy for estimating the factor analytic mapping in practice requires assumptions that may not hold. We discuss these assumptions and how they may be the reason we obtain disparate estimates of the mapping factor in an application of the proposed methods to groups of patients.

Keywords: cross-walk, HRQOL, mapping, reliability.

Copyright © 2013, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

Measuring health-related quality of life (HRQOL) has become standard in clinical trials of treatments and in studies assessing the impact of health conditions. There is little agreement, however, on the choice of instrument to achieve this. It is often felt that disease-specific measures (DSMs) of HRQOL are most sensitive to changes in HRQOL related to a specific condition or disease. For example, the Minnesota Living with Heart Failure (MLHF) questionnaire [1,2] may be chosen to assess the progress of patients with congestive heart failure entering intensive treatment, the 25-item National Eye Institute Visual Function Questionnaire (NEI-VFQ-25) [3] to assess the benefit of eye surgery, or the Beck Depression Inventory [4] to measure the impact of a treatment for depression. However, it is desirable to express treatment impact on HRQOL in terms that allow comparison to other health conditions, and preference scoring for application in cost-effectiveness analysis. When only a DSM is available from a trial, there may be a need to convert (map or cross-walk) the scores of the MLHF questionnaire to the EuroQol five-dimensional (EQ-5D) questionnaire or other generic HRQOL utility measure. In this issue of the journal, Lu et al. [5] address methods for mapping DSMs into generic scores.

The word "mapping" has taken on several meanings, at least two of which are relevant to Lu et al. [5] and this discussion. The first is anticipated in the Introduction of Lu et al. [5], in which "mapping" is defined as "treatment effects estimated on the disease-specific measure (DSM) are multiplied by a mapping coefficient to convert them into treatment effect estimates on the generic QOL scale." What the authors end up developing, however, is a second conceptual mapping applicable to the true, not estimated, mean score or treatment effect. This second factor loading–based mapping informs about the relationships between what the instruments truly attempt to measure. Because it pertains to a theoretical framework not mudded by measurement error in a particular setting, it has more desirable properties. As noted in Lu et al. [5], the framework has been used in educational testing to "align instruments," that is, to ascertain that they fundamentally measure the same underlying abilities. One needs to be clear about the distinction and how it affects the correct way to perform a conversion in practice. In this commentary, I attempt to clarify how the proposed methods may be put to use and also further discuss the assumptions needed for the specific methods proposed by Lu et al. [5].

Two different approaches to estimating the mapping between true underlying means are proposed. (It should be noted that models in Lu et al. [5] are actually not presented in terms of differences, but in terms of group means, later to be differenced.) Both approaches depend on the availability of data or results from a separate study that measured the DSM and obtained

* *Address correspondence to*: Mari Palta, Population Health Sciences, University of Wisconsin, 610 Walnut Street, Madison, WI 53726. E-mail: mpalta@wisc.edu.

1098-3015 /$36.00– see front matter Copyright © 2013, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.jval.2012.07.009

simultaneous scores on at least one generic HRQOL instrument. Briefly, the first method fits a linear regression of the generic score on the DSM and then applies corrections for measurement error through knowing the reliability of the DSM. The second method assumes that there are data on two generic scores together with the DSM and that the two generic scores share the aspect—and only the aspect—of health covered by the DSM. Linear correlations are then used to extract a conversion factor.

We illustrate the impact of the assumptions by applying the methods proposed in Lu et al. [5] to a study of the MLHF questionnaire in congestive heart failure clinics and the NEI-VFQ-25 in patients undergoing cataract extraction [6]. This study also obtained simultaneous data on several generic instruments, as well as repeated measures, allowing the extraction of reliability coefficients for the MLHF questionnaire and the NEI-VFQ-25.

Comment on Methods

The Role of Reliability in Mapping

The measurement error, reliability, and underlying factors of HRQOL instruments have received relatively little attention in the cost-effectiveness analysis and HRQOL literature, with some exceptions [e.g., 7–10], and Lu et al. [5] should be lauded for bringing renewed attention to these important measurement aspects. As pointed out, in an ordinary regression analysis of a generic HRQOL instrument on a DSM, the regression coefficient is affected by the reliability of the DSM in the particular sample. Specifically, the slope of the relationship is $b_x = \beta\rho_x$, where β is the regression coefficient that would be obtained had the true value of DSM been known and ρ_x is the reliability of the predictor X (i.e., DSM), obtainable, for example, from test/retest data. Nonetheless, the equation using b_x as the mapping slope is correct (unbiased) for the particular study when X is either a measured individual or an estimated mean score.

It is relevant to note that ρ_x is the ratio $\sigma_a^2/(\sigma_a^2 + \sigma_e^2)$, where σ_a^2 is the variance between individuals in the true x (DSM) and σ_e^2 is the measurement error variance. Then, the estimated coefficient for obtaining a generic score from a DSM may vary between studies with different underlying σ_a^2 (i.e., spread in DSM), and formulating the mapping as one between true values (i.e., in terms of β) is advantageous. However, when DSM values measured in practice are to be mapped, β still needs to be multiplied by a ρ_x appropriate to the sample under consideration. In fact, it may be desirable to build into a study the estimation of ρ_x, for example, via a test-retest component. When taking the mean of DSM across a treatment group, both variance components become averaged so that the reliability coefficient is

$$\left(\frac{\sigma_a^2}{n}\right) \Big/ \left(\frac{\sigma_a^2}{n} + \frac{\sigma_e^2}{n}\right) = \frac{\sigma_a^2}{(\sigma_a^2 + \sigma_e^2)}$$

which is identical to that of the initial measurement. Hence, $b_x = \beta\rho_x$ should be used in the mapping of estimated means, albeit ideally with a ρ_x modified to fit the specific study.

Role of Linearity of the Relationship

It should be noted that the estimation of the relationship between the generic score and the DSM via $b_x = \beta\rho_x$ depends on the regression of the observed generic score on DSM being linear. Given the strong ceiling effects in many generic scores, this assumption may not hold, and b_x across the range of X may depend on the HRQOL distribution of the sample. Several approaches to the linearity/ceiling effect problem have been applied and investigated in the literature, such as using splines [11], censored or latent class models [12], discretizing the scores [10], or fitting different models to different ranges of the DSM [5]. The mappings proposed [5] would need further work as to how reliability would be estimated and applied in such models. It should be noted that this problem is separate from the linearity assumed in the factor analytic relationship, which is between underlying unobserved constructs.

Assumptions in Using Two Generic Scores

Another proposed way to obtain the measurement error–corrected mapping of the DSM onto a generic instrument requires that the DSM be available in an auxiliary study together with two generic instruments [5]. A rather strict assumption is made that the components in the two instruments not captured by the DSM are uncorrelated. In other words, the component(s) the two generic instruments have in common are those and only those captured by the DSM. One may envision that such an assumption can be true if the DSM is comprehensive enough to capture all standard domains of HRQOL. On the other hand, even if the generic instruments capture all domains of the DSM affected by the disease under study, it is possible that individuals have other health condition, or even unexpected side effects of treatment captured by both generic instruments, but not by the DSM. In the latter case, the assumption would not hold. It has been shown that many generic instruments have much variance in common [9,10], requiring careful consideration before the assumption of nonshared extraneous-to-DSM domains can be made.

Equation 10 in Lu et al. [5] provides the estimator. For consistency, we use X to denote values of the DSM, Q1 to denote values of the generic instrument desired, and Q2 of the auxiliary generic instrument:

$$\beta = (\text{Sign of relationship})\left(\left|\frac{corr(Q1,Q2)}{corr(Q1,X)}\right|\right)\sqrt{\left(\frac{Var(Q1)}{Var(x)}\right)}$$

In this formula, the expression under the square-root is simply a conversion of the two instruments to the same scale, and the sign converts, for example, a DSM scale where lower value indicates better HRQOL to a generic scale where higher value indicates better HRQOL. The ratio of correlations, however, will capture the needed quantity only if corr(Q1,Q2) is not inflated by shared extraneous-to-DSM factors. The feasibility of the assumption likely depends on the level of health in the sample used, as generic instruments have been shown to capture underlying dimensions differentially at different levels of health [9]. Ceiling effects in Q1 can reduce the linear correlation between Y and Q1, possibly more so than the correlation between Q1 and Q2, where ceiling may be shared. It is difficult to assess the impact of the linearity assumption between Q1 and X, except to note that nonlinearity-based bias in the directly reliability adjusted and the two-instrument–based estimators will be in opposite directions, as the linear relationship is in the numerator of the former and in the denominator of the latter.

Results of Application

Problems arising with the assumptions discussed above can be summarized as 1) a linear relationship between the measured generic score and the DSM may not hold because of ceiling effects in most generic scores and 2) generic scores may share health dimensions not measured by the DSM, except for specific DSMs and generic scores. We now turn to examining the performance of the methods in a patient sample evaluated with a number of HRQOL measures [6].

The Clinical Outcomes and Measurement of Health Study [6] provides longitudinal data on the NEI-VFQ-25 [3] in a group of

Download English Version:

https://daneshyari.com/en/article/10485052

Download Persian Version:

https://daneshyari.com/article/10485052

Daneshyari.com