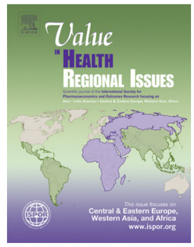
ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

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



Mapping the Nottingham Health Profile onto the Preference-Based EuroQol-5D Instrument for Patients with Diabetes

Zoltán Vokó[1,2,], Renáta Németh[3], László Nagyjánosi[2], György Jermendy[4], Gábor Winkler[5,6], Tibor Hídvégi[7], Zoltán Kalotai[8], Zoltán Kaló[1,2]*

[1]*Faculty of Social Sciences, Department of Health Policy & Health Economics, Institute of Economics, Eötvös Loránd University;* [2]*Syreon Research Institute;* [3]*Faculty of Social Sciences, Department of Statistics, Eötvös Loránd University;* [4]*3rd Department of Internal Medicine, Bajcsy-Zsilinszky Hospital;* [5]*2nd Department of Internal Medicine – Diabetology, Saint John Hospital, Budapest, Hungary;* [6]*Faculty of Health Care, Institute of Theoretical Health Sciences, University of Miskolc, Miskolc, Hungary;* [7]*Department of Metabolic Disorders, Endocrinology and Diabetology, Petz Aladár County Teaching Hospital, Győr, Hungary;* [8]*Novartis Ltd., Budapest, Hungary*

ABSTRACT

Objective: The aim of this study was to derive a function that can map the Nottingham Health Profile (NHP) questionnaire onto a utility measure, the EuroQol five-dimensional (EQ-5D) questionnaire index, for diabetic patients. **Methods:** A cross-sectional study was performed on diabetic patients in Hungary with different complications in which quality of life was measured by using both the NHP questionnaire and the EQ-5D questionnaire. Ordinary stepwise-backward least-squares regression was used to develop a mapping function. Adjusted R^2, Akaike's information criterion, and root mean square error were used to assess the performance of the model. The robustness of the models was tested using 10-fold cross-validation and bootstrapping. **Results:** The best-fitting models were those that contained all the NHP statements as predictors and a stepwise reduced version that contained only 19 statements. The latter model, however, showed considerable variability in the selection of predictors. The adjusted R^2 of the former model was 0.68, the root mean square error was 174, and the Akaike's information criterion was −559.9. **Conclusions:** The expected value of the EQ-5D questionnaire can be reasonably predicted on the basis of results of the NHP in patients with diabetes mellitus. The mapping function of the NHP onto the EQ-5D questionnaire is capable of estimating the expected EQ-5D questionnaire utility values in a group of patients with diabetes. The function's applicability for individual-level predictions, however, is limited. Further research is needed to find out whether mapping functions developed in Central-Eastern European countries are transferable to Western European countries.

Keywords: diabetes mellitus, economic models, linear models, quality of life, health profile.

Introduction

The burden of diabetes mellitus, which is already large in most countries, is steadily increasing with the epidemic of obesity, inactivity, population aging, and greater longevity in patients with diabetes [1]. An important aspect of this burden is the loss of quality of life (QOL) among patients who suffer from complications of diabetes. Therefore, many studies regularly measure the QOL of diabetic patients as an important outcome [2,3]. In these studies, QOL is usually measured by using a generic or a disease-specific instrument that does not provide utility measures. To support an efficient allocation of resources to both prevent and control diabetes, the cost-effectiveness of alternative strategies needs to be considered. Cost-effectiveness analysis requires, however, preference-based utility measures of QOL. Therefore, it would be desirable to have mapping functions that could translate the results obtained by non–preference-based instruments to utility values. Accordingly, the aim of this study was to develop a function that can map the Nottingham Health Profile (NHP) questionnaire onto the EuroQol five-dimensional (EQ-5D) questionnaire utility index.

Methods

Participants, Setting

The study was performed in 15 centers in Hungary specializing in the care of diabetic patients. Study groups were defined by the type of diabetes (type 1 or type 2), type of treatment (oral antidiabetic or insulin treatments), and the presence of complications (visual impairment, nephropathy, neuropathy, coronary

Conflict of interest: This study was financially supported by Novartis Hungary and Sandoz Hungary.

* *Address correspondence to:* Zoltán Vokó, Faculty of Social Sciences, Department of Health Policy and Health Economics, Institute of Economics, Eötvös Loránd University, Pázmány Péter sétány 1/a, 1117 Budapest, Hungary.

E-mail: voko@tatk.elte.hu.

http://dx.doi.org/10.1016/j.vhri.2014.06.002

heart disease, and/or cerebrovascular disease). Quotas were given to the centers according to these modalities, and the centers enrolled the patients consecutively. The primary objective of the study was to estimate the QOL among diabetic patients with different complications [4]. This report presents the results related to the secondary objective, which was mapping the NHP questionnaire onto the EQ-5D questionnaire. The study was approved by the institutional ethical boards and by the National Ethical Board for Medical Research (8-239/2009-1018EKU). All persons gave their informed consent before their inclusion in the study.

Instruments

All patients were asked to fill in the NHP questionnaire and the EQ-5D questionnaire [5,6]. The questionnaires were self-administered, although help was provided by the study personnel on request. The EQ-5D questionnaire consists of the EQ-5D descriptive system and the EQ visual analog scale. We used the former for this study. The EQ-5D descriptive system comprises five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has three levels: no problems, some problems, extreme problems. The respondent indicates which statement in each of the five dimensions describes his state most appropriately. A total of 243 possible health states can be defined this way. Utility values (EQ-5D questionnaire index values) were attached to these states in our analysis by using the time trade-off valuation technique from a UK study [7]. The EQ-5D questionnaire has been validated in Hungary [8].

The 38 statements of the NHP fall into six sections: energy (nine items), pain (eight items), emotional reactions (three items), sleep disturbances (five items), social isolation (five items), and physical mobility (eight items). The respondents need to consider whether each statement applies to him or her in general. For each section, a score is calculated according to the NHP scoring guidelines [9]. Individual items are scored 1 for a "yes" response and 0 for a "no" response. The score for each section represents the summation of the item scores expressed as a percentage. As such, the scores range from 0 to 100, with a higher number representing greater distress. A patient-completed index of distress from the Nottingham Health Profile (NHP-D) can be calculated as the sum of 24 items out of the total 38 statements [10]. The NHP was one of the first QOL instruments validated in Hungary [11]. As in the Hungarian validation study NHP-D showed great reliability and sensitivity, it proved to be a valuable outcome measure for trials for which carefully developed disease-specific QOL measures are unavailable.

Analysis

OLS regression models

The most common method for establishing a functional relationship between two QOL measures is ordinary least squares (OLS) regression [12–17]. There are many possible ways to obtain EQ-5D questionnaire indices from a non–preference-based instrument such as the NHP. Having a data set in which patients have responded to both instruments, empirical matching is the most straightforward method. The essence of this approach is to explore the relationship between the two instruments by using regression analyses. Because the range of the EQ-5D questionnaire is bounded, the use of OLS regression models risks predicted values outside the range [12]. Alternatives would be to use generalized linear models, tobit regression, or censored least absolute deviations to allow for censoring. A recent review, however, showed that the standard method for mapping to the EQ-5D questionnaire was OLS regression [17]. The review of 90 studies found that 80% of the studies used OLS regression. There are arguments supporting OLS. Tsuchiya et al. [15] showed that in mapping problems such as ours, performance of the generalized linear model was comparable to that of the OLS regression, but not better. They concluded that "the associated benefits of generalized linear model do not seem to outweigh its costs." In another recent study, where OLS, tobit, and censored least absolute deviation models were compared, OLS turned out to be the best-performing model because it produced the lowest RMSE and mean absolute error and the highest R^2 [18]. Taking into account the above considerations and the fact that the range of the predicted EQ-5D questionnaire values of our OLS models hardly exceeded the limits (see Table 4), we decided not to use the generalized linear model.

In our analysis, the dependent variable was the overall EQ-5D questionnaire index, while the explanatory variables were the 38 NHP items used in the different aggregation levels. The study did not attempt to predict separately the individual EQ-5D questionnaire dimensions because previous studies showed it to be no more efficient in terms of prediction [12,14,19].

First, we regressed the overall EQ-5D questionnaire index on the total NHP-D score as a start-up. Next, we refined this basic model by first using the six NHP scores and later using all 38 separate statements as dependent variables.

In each model, we looked for nonlinear dependence between the independent and the explanatory variables as a signal to put either squared or otherwise transformed terms into the model. In addition, we studied whether two-way interactions could lead to better-fitting statistics, which are commensurate with the higher complexity of a given model.

If the full models fitted reasonably well, then we looked for more parsimonious models by running stepwise backward regression. We started with the full model and set the removal criterion to $P = 0.1$ and the reentry criterion to $P = 0.05$. This means that if a predictor was not significant at the level of 0.1, it was removed from the model but it could reenter if its significance reached 0.05 after removing other predictors.

OLS regression assumptions were examined by using the following methods:

1. The variance inflation factor (VIF) index was used to test collinearity. Some of the problematic (VIF index >10) predictors were removed.
2. Nonlinearity in any of the predictors was checked with the help of augmented partial residual plots.
3. Normality assumption for regression residuals was checked by plotting the quantiles of the regression residuals against the quantiles of standard normal distribution (Q-Q plot).
4. The assumption of the homoscedasticity of the residuals was visually checked by plotting predicted values against standardized residuals because known statistical tests for homoscedasticity are very sensitive to the violation of the normality assumption and cannot be used if the normality assumption fails.

Assessment of goodness of fit

Goodness of fit and predictive power were measured with the root mean square error (RMSE), the adjusted R^2 indices, and Akaike's information criterion (AIC). AIC is an information-theoretical model selection criteria with the advantage of applicability to non-nested models. Lower AIC values indicate a better model. The range of the predicted EQ-5D questionnaire values is also reported because OLS models struggle to produce EQ-5D questionnaire indices that are negative or equal to 1.

Internal validity

Judging the internal validity of the results is of primary importance when predictive models are built. Because no external data set is available, within-sample validation was carried out with the help of replication techniques. Two tests were conducted. First, stability of model coefficients is of interest because the relatively large number of predictors may lead to an overparametrized model. Stability,