18
Testing construct validity of EQ-5D by Confirmatory Factor Analysis (CFA) and Structural Equation Modeling (SEM)

(addendum to EuroQol 2007 proceedings)

Clouth J[1], Brähler E[2], Schmidt P[3], Kohlmann T[4]

18.1 ABSTRACT

Introduction

The objective of this study was to evaluate the construct validity of the 5 items of the EQ-5D by CFA and SEM. At least to our knowledge CFA and SEM has not often been applied for analyzing psychometric properties of the EQ-5D.

Methodology

Data came from a representative population survey conducted in Germany in 2002. 2022 subjects between 16 and 93 years of age were randomly selected. In addition to a standard one-factor model for the 5 EQ-5D items, models with correlations between error terms and direct effects between observed variables were estimated. AMOS (Ver. 5.0) was used for parameter estimation and assessment of fit.

Results

Goodness of fit (GFI = 0,976) and adjusted goodness of fit (AGFI = 0,928) for the standard one-factor model were acceptable. Pclose and chi-square statistics were sub-optimal (pclose < 0.001, chi-square = 132.2 / df = 5). Inclusion of error correlations and direct effects between items resulted in a relevant improvement of fit statistics. In all models, factor loadings were acceptable except for the dimension anxiety / depression (factor loading: 0,42).

1. Lilly, Germany
2. University of Leitpzig, Leitpzig, Germany
3. University of Gießen, Gießen, Germany
4. Universität Greifswald, Germany (Member of the Euro Qol Group)

Scalone L and Mantovani L, editors.
25th Scientific Plenary Meeting of the EuroQol Group - Proceedings:355-363

Discussion and Conclusions

Although the EQ-5D instrument was not constructed using CFA and SEM models, our results indicate that most of the fit measures are acceptable. Closer inspection of correlations between error terms and direct effects on the level of observed variables can give a more detailed account on the internal structure of the EQ-5D.

18.2 INTRODUCTION

Background

The EQ-5D is a generic instrument to measure health related quality of life (HRQOL) and consists of three sections[1]. The first part describes HRQOL in 5 health related dimensions: morbidity, self care, usual activities, pain/discomfort and anxiety/ depression. Every item is presented in the same order and each one has three degrees of severity such as no problem=1, some/moderate problems=2 and many problems=3. The individual indicates the degree to which he experiences problems for each health status. Consequently a measurement error of zero has to be assumed because for each heath status only one item for measurement is used. The EQ-5D represents a model with five items and one latent variable which express HRQOL and can be illustrated as in Figure 18.1.

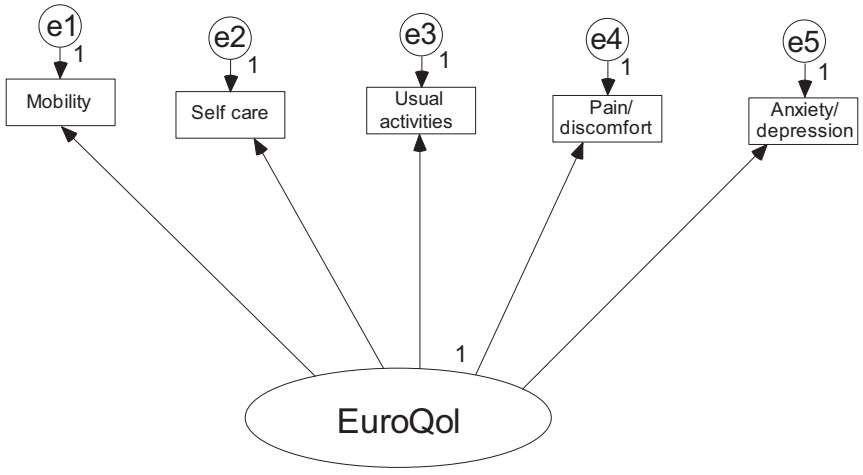

Figure 18.1.

The EQ-5D is translated in many languages and one of the mostly used instruments worldwide to investigate HRQOL.

The second part of the EQ-5D is a visual analog scale (VAS). Individuals can rate their health status through the VAS. Alternatively the rating of health status can be performed by either standard gamble of time trade off methodology in combination with multiple regression.

The third part is the result of the activity describe in part two and ends in an index describing the HRQOL[2].

According to our opinion it seems evident that the validity of the index depends on the validity of the model shown by Figure 18.1.

The confirmatory factor analysis (CFA) is a method that supply validity data for both descriptive and inferential statistics. CFA is a widely applied tool to test the internal construct validity of instruments to measure HRQOL[3,4,5]. A literature search with the key words "Structural equation modeling", SEM, Confirmatory factor analysis, CFA, LISREL, EQ-5D, EQ5D and EuroQol in Medline and Embase resulted in 1 hit.

```
(Database: EMBASE (emed), Ovid MEDLINE(R) (mesz)
Search Strategy:
--------------------------------------------------------------------------------
1     ("structural equation modeling" or "confirmatory factor analysis"
      or LISREL).af. (5605)

2     (EQ-5D or EQ5D or EuroQol).af. (1985)

3     1 and 2 (1)

4     from 3 keep 1 (1).
```

The research of Gasquet et al (2006)[7], evaluated by SEM the causal effects between treatment and HRQOL but not the internal construct validity.

The estimation of utilities in Germany through multiple regression achieved a determination coefficient of 0,51[8]. The means of multiple regression requires independent variables. In the case of HRQOL the question occurs if an improvement in one dimension implies an improvement for another dimension. This problem can be solved not by multiple regression but by structural equation modeling (SEM) which controls multicollinearity.

Objectives

The objective of present analysis is three fold. First the study described in this paper aimed to test the internal construct validity of the EQ-5D for a representative sample of the German population. The second objective consists in constructing a SEM which could be used for estimating utilities. The third objective is to deliver propos-

als for improving both the internal construct validity of the EQ-5D and to deliver an approach through SEM to estimate utilities.

The results of the present analysis are encouraging and imply concrete proposals to improve the internal construct validity of the EQ-5D and the estimation of utilities of the EQ-5D

18.3 METHODS

Data

In 2002 the university of Leipzig commissioned the market research institute USUMA in Berlin to collect a representative sample for the German population. In a first step households through random-route procedure have been recruited. The target persons of each household were also randomized, visited at home and interviewed by skilled personal. The sample size was 2022 people (age from 16 to 93) out of which 948 were male and 1074 female. All persons had to answer several questionnaires one of which was the EQ-5D.

Confirmatory factor analysis

Exploratory factor analysis (EFA) is designed for the situation where links between the observed and the latent variables are unknown or uncertain. In the case of the EQ-5D as shown in Figure 18.1 the relation of observed and latent variables is fixed. In contrast to EFA confirmatory factor analysis (CFA) is appropriately used when the researcher has knowledge of the underlying structure of observed and latent variables. CFA evaluates by statistical means the model to be tested to determine the adequacy of its goodness of fit to the sample data (for more detailed discussion of CFA see Bollen)[6]. It is important not only to compute the validity of the test but also the validity of single indicators. The test of indicators is also covered by CFA.

Structural equation modeling

Structural equation modeling (SEM) in its general form consists of two parts:

(i) the measurement model and
(ii) the structural equation model

The measurement model specifies how latent variables depend on or are indicated by the observed variables.

The structural equation model specifies the causal relationships between latent variables, describes causal effects, and assigns the explained and unexplained variance of

the dependent variable. SEM integrates multivariate techniques such as linear regression analysis, analysis of covariance, and factor analysis as special cases. By comparing the covariance matrix derived by the model and the estimated parameters with the observed matrix, the validity of the model can be evaluated. In this paper, graphical (rather than matrix equations) representation of SEM is used.

The hypothesis for the presented model is that impairments in the physical and or mental field carry over to other domains of HRQOL. Due to the fact as shown in the result section that the mental dimension has an unsufficient factor loading and only causal relations from the physical dimension have been taken into consideration. It must be pointed out that this model doesn't call for validity but just for illustrating a new approach for estimating utilities. With the given data it was not possible to model a relationship between "pain/discomfort" and "mobility".

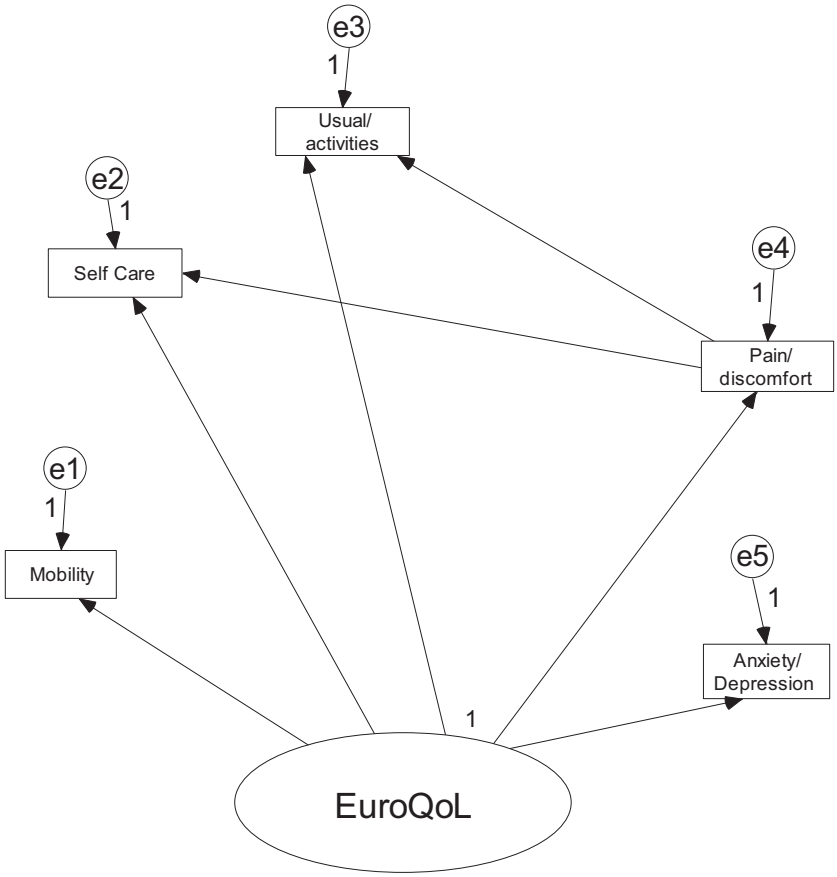

Figure 18.2.

18.4 RESULTS

The results for the internal construct validity are illustrated in Figure 18.3.

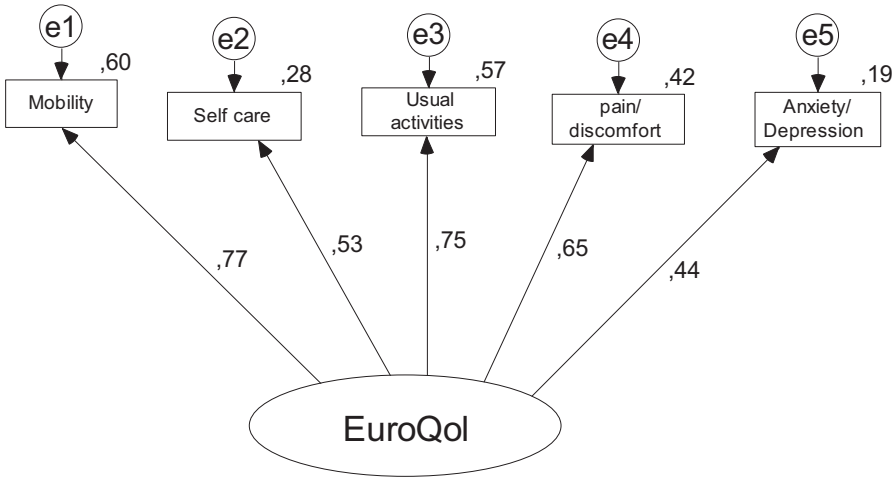

Figure 18.3. gfi=,976 agfi=,928 pclose=,000 chi-square=132,195 df=5 rmr=,006

The new index's goodness of fit was measured using the chi-squared test. The probability value associated with the chi-squared test represents the likelihood of obtaining a chi-squared value that exceeds the chi-squared value when the null hypothesis (H_0) is true. Thus, the higher the probability associated with the chi-squared value, the closer is the fit between the hypothesized model under H_o and the perfect fit The goodness-of-fit (GFI) value can fall between zero (0) and unity (1), where unity indicates a perfect fit. (See Table 18.1) Appendix for explanation of parameters used to test goodness of fit.)

Table 18.1. Key indices and fit measures

Descriptive measure	Range	Criterion	Results
GFI	0-1,0	>0,95	0,976
AGFI	0-1,0	>0,95	0,928
Chi^2/df	>0	3,0> Chi^2/df>1,0	26,439
Pclose	0-1,0	>0,5	0
RMSEA	0-1,0	<0,05	0,06

Factor scores are necessary to construct weighted indexes. Based on the mode of their calculation, factor scores produce the strength of relation between items (variables) and target factors with prediction of the latter parameters Scores for both first-order factors and factors of higher order were calculated with AMOS 5.0. The factor scores were further processed with the program SPSS 11.0 (SPPS Inc.; Chicago, Illinois, USA).

Factor loadings are standardized or unstandardized partialized regression coefficients from the latent variables to the indicators. (The term "partialized" means that other independent variables [factors] are held constant) Factor scores are the weight for predicting a latent variable from its observed indicators[5].

The explained variance for "mobility" with 0,6 and for "usual activities" seem acceptable. For "pain and comfort" with 0,42 the threshold of acceptance could be discussed. The determination coefficients for "self care" and" anxiety/depression" with 0,19 seems not acceptable.

The factor loading for "mobility, "usual activities" and "pain/discomfort" are sufficient whilst for "self care" a higher loading would be favorable. The factor loading for "anxiety depression is too low.

The GFI with 0,976 is in the expected range and the AGFI is near to this threshold. Chi^2/df with 26,439 is absolutely out of range. The RMSEA with 0,06 with the expected range. The results for the structural equation modeling is shown in Figure 18.4.

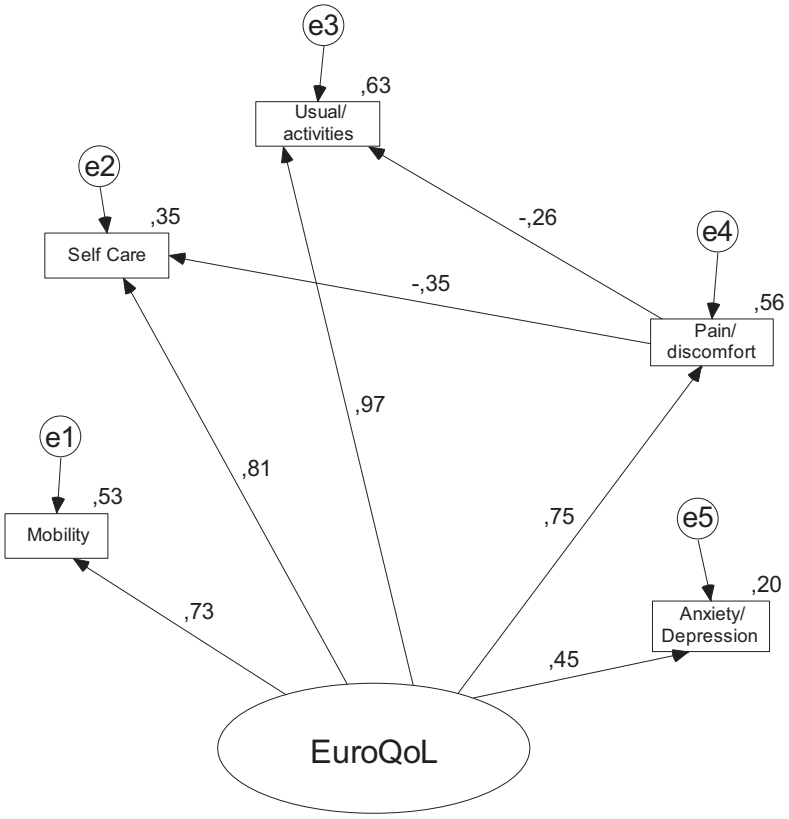

Figure 18.4. gfi=,95 agfi=,974 df=3 rmr=,002 chi-square=26,017 pclose=,173

The explained variance for “mobility”, “usual activities” and “pain/discomfort” are on an acceptable level whilst for “self care” and particularly “anxiety/depression” a higher explained variance is to be postulated. Factors loadings except for “depression/anxiety” are sufficient. The strengths of effects form pain discomfort with -0,26 to “usual activities” and -0,35 to “self care” are interesting and will be discussed later. The fit measures for the model are, excepted chi^2/df, in the accepted range.

18.5 DISCUSSION

The internal construct validity of the EQ-5D is of mixed evidence. It could be expected that due to the low explained variance and to the low factor loading the dimension “anxiety/depression” will have difficulties to differentiate. The reason for this underperformance of this item might be the fact that two meanings are within one item. A golden rule for constructing items is to have never more meanings than items. A rewording might offer a real contribution to improve the internal construct validity of the EQ-5D particularly because the rest of the model shows robust validity.

For the estimation of utilities multiple regressions are applied. This methodology requires that variables are independent. Particularly in HRQOL it might be that an improvement for e.g. “pain/discomfort” might have a positive effect on “usual activities”. Consequently the items are not independent and multiple regressions would necessarily skew its t-values. The present analysis could demonstrate that these relations respectively indirect effects can be estimated respectively controlled by means of structural equation modeling. An expected result of this approach might be a more precise estimation of utilities and a better explained variance.

These results are, according to my opinion, encouraging. The rewording could be easily done and a new version could be easily tested. Having such a more valid version of the EQ-5D a SEM including effects from the dimension “psyche” to other dimensions could be modeled and ongoing serve to estimate utilities.

There are, certainly, some limitations to the present work. The analysis is valid for these data and is consequently limited to the German population. It has still to be proved if these models are also valid for other populations.

The internal construct validity has been testes through CFA. Other methods such as Rasch Modeling has not been applied and could consequently be not compared to CFA. Only internal construct validity has been tested but not external validity
Further research might be fruitful.

18.6 REFERENCES

[1] Prieto L. et al, Psychometric validation of a genereic health-related quality of life measure (EQ-5D) in a sample of schizophrenic patients. Current Medical Research and opinions, Vol 20, No 8, 2004, 827-835.

[2] Hinz A. et al, Der Lebensqualitätsfragebogen EQ-5D: Modelle und Normwerte für die Allgemeinbevölkerung. Psychotherapie Psychosomatik Medizinische Psychologie, S. 56 (2): 42-48, Feb 2004.

[3] Byrne B. Structural Equation Modeling with AMOS. Lawrence Erlbaum Associates, Publishers, 2001.

[4] Bühler M. Einführung in die Test- und Fragebogenkonstruktion. Pearson Studium, 2003.

[5] Schmidt P. et al, Constructing an Index fort he Subjective Well-being Under Neuroleptics Scale (SWN) short form: Applying structural equation modeling for testing teliability and validity of the index. Quality of Life Research (2006), 15; 1191-1202.

[6] Bollen K.A. Structural equations with latent variables. New York: Wiley 1989

[7] Gasquet I. et al. Patient satisfaction with psychotropic drugs: sensitivity to change and relationship to clinical status, quality of life, compliance and effectiveness of treatment. Results from a nation-wide 6-month prospective study. European Psychiatry 21 (2006), 531-538.

[8] Schöffski et al. Gesundheitsökonomische Evaluationen, 2000.