Selection of bolt-ons after factor analysis identification: are linear regression models a useful technique?

Aureliano P Finch†, John E Brazier, Clara Mukuria
The University of Sheffield, Sheffield, UK

† Email: APFinch1@sheffield.ac.uk

1. Background and scope

It is now recognized that the EQ-5D may miss dimensions important for some conditions. When this happens, a possible solution is adding bolt-ons to expand its descriptive system.

Previous bolt-on studies have identified potential candidates using information on validity in specific areas such as vision (1). Although this is a useful approach for identifying individual bolt-ons, it does not help in identifying what other dimensions may be missing from the EQ-5D.

Factor analysis has been seen to be a potential approach for bolt-on identification. This techniques pinpoints to a list of factors, and items loading on them, that are not related to the EQ-5D latent constructs (2). These can be adapted / developed into bolt-ons.

However, not all bolt-ons can be added to the EQ-5D simultaneously, as this would affect the measure's acceptability and feasibility. Hence, methods to select bolt-ons from the identified list are needed.

This study investigates the possibility of using linear regression models for the selection of bolt-ons after factor analytic identification.

2. Methods

The Multi Instrument Comparison Database was used for the analysis. Six factors: energy/vitality, satisfaction, relationships, hearing, vision and speech, and 37 items loading on them, drawn from previous factor analysis, were used (2). Two tests were performed.

TEST 1: linear regressions were fitted to determine whether different factors and items helped explain variations of self reported health as measured by the VAS health scale. Bolt-on relevance was judged comparing the strength, direction and statistical significance of unadjusted β coefficient. Regression were considered an appropriate technique if they managed to discriminate between bolt-on.

TEST 2: linear regressions were fitted to further investigate whether factors and items helped explain the negative effect of six chronic conditions on self reported health. A reduction in the coefficients for the chronic conditions dummies meant that the factor or item detected the effect of the condition. Regression were considered an appropriate technique if they managed to discriminate bolt-ons impact based on expectations derived from previous research (e.g. hearing factor having an impact on hearing problems).

3. Results

Test 1

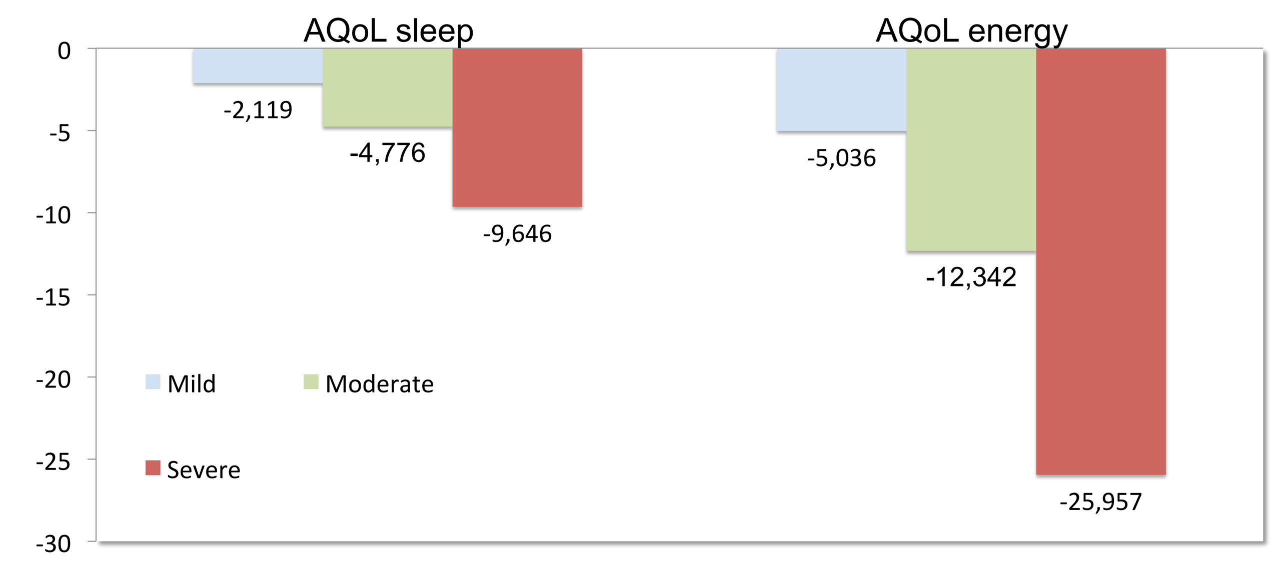
β coefficients and standard errors for latent factor regressions

Factors	*β coefficients*	*Standard Errors*
Satisfaction	-4,323**	,112
Relationships	-5,298**	,235
Hearing	-1,209**	,353
Speech/cognition	-2,269**	,287
Vision	-2,185**	,257
Energy/ vitality	-7,648**	,217

Note: ** p≤0.01

Table 1. All factors were able to explain variations in self reported health. Energy/vitality, relationships and satisfaction reported substantially larger coefficients than the other three factors.

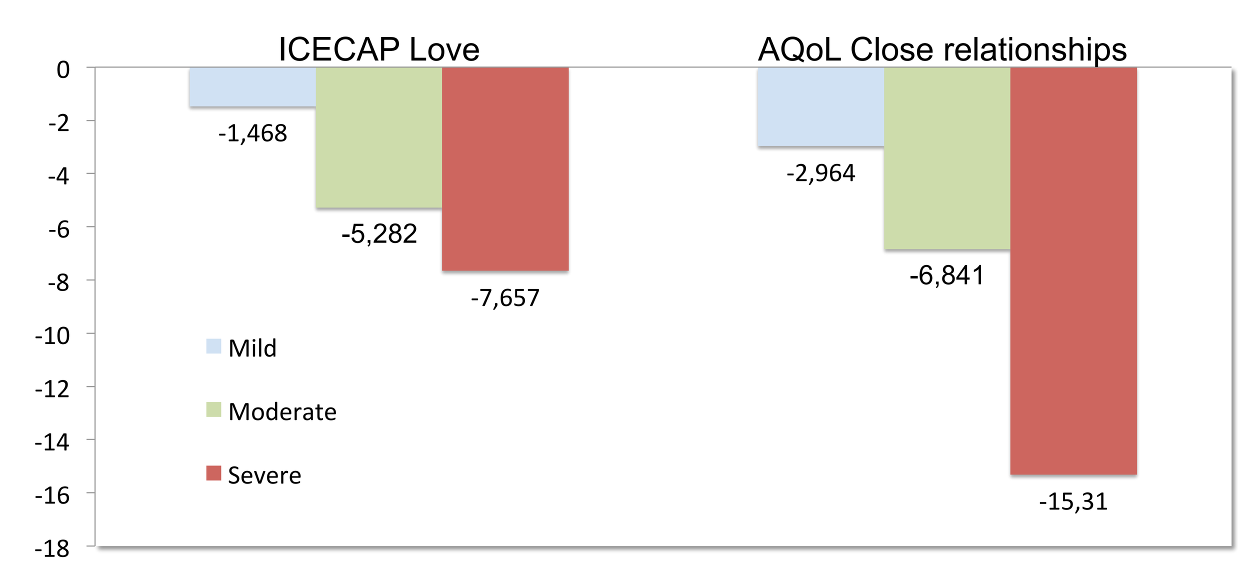
β coefficients of two items loading on energy/ vitality



Note: All dummy variables were statistically significant at p≤0.01

Figure 1. Both items were able to explain variations in self reported health. Coefficients for AQoL energy are substantially larger than coefficients for AQoL sleep, for all levels of severity.

β Coefficients of two items loading on relationships



Note: All dummy variables were statistically significant at p≤0.01

Figure 2. Both items were able to explain variations in self reported health. Coefficients for AQoL relationships are larger than coefficients for ICECAP love.

β Coefficients of two items loading on speech/ cognition

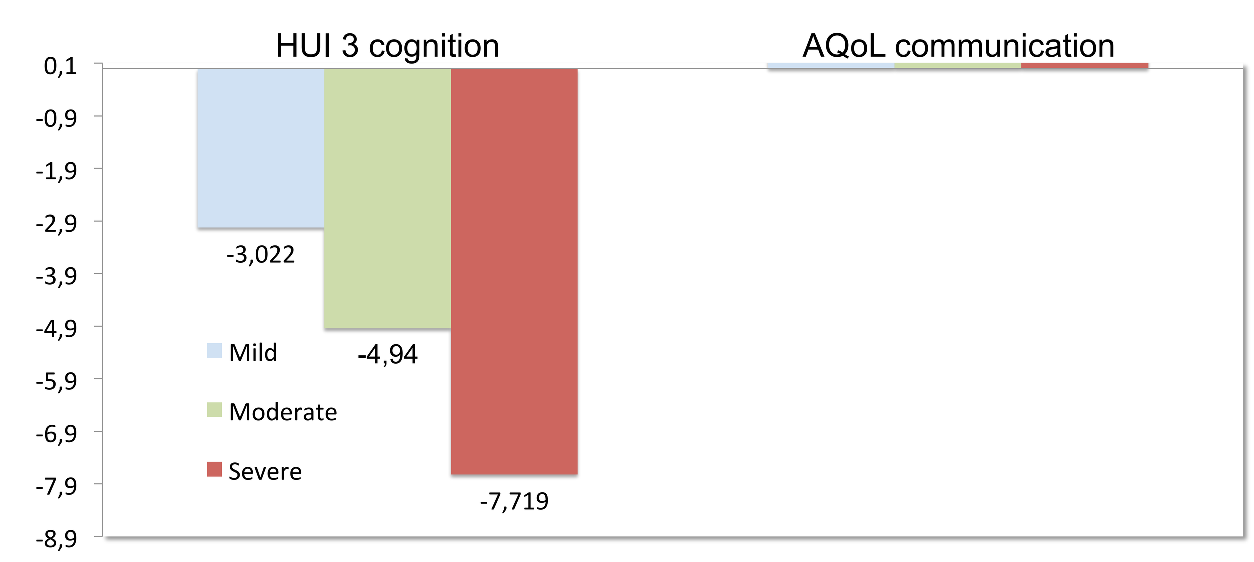


Figure 3. Only cognition was able to explain variations in self reported health. None of the coefficients for AQoL communication were statistically significant.

Linear regressions were able to differentiate factors' ability to detect changes in self reported health. Results for factors and items regressions are generally concordant, as coefficients were generally greater for items loading on factors with larger coefficients in the latent regressions. Items regressions appear useful in discriminating between different wordings of the same constructs, or to select the most relevant concept in multi-concept factors e.g. cognition preferred to communication.

Test 2

Decrements of disease dummies for factors regressions

	Cancer	*Asthma*	*COPD*	*Depression*	*Diabetes*	*Hearing problems*	*Arthritis*	*Heart diseases*	*Stroke*
Satisfaction	0,045	0,019	-0,617	0,037	0,062	-0,014	0,003	0,047	1,735
Relationships	0,067	0,048	0,294	0,053	0,053	0,031	0,039	0,063	0,010
Hearing	-0,069	-0,04	-0,093	-0,055	-0,074	-0,009	0,026	-0,069	-0,589
Speech	-0,07	-0,042	-0,115	0,055	-0,071	-0,014	-0,029	-0,069	-0,580
Vision	-0,397	-0,206	-0,436	-0,863	-0,503	-0,274	-0,06	-0,322	-2,623
Energy	0,005	-0,003	0,638	-0,022	-0,050	0,027	0,010	-0,030	-1,561

Note: Red represent coefficient decrements larger than -0,3.

Table 2. All dummy variables coefficients remained statistically significant after the addition of the factors. Decrements were not consistent with expectations. For example, hearing did not reduce the impact of hearing problems. Similarly, vision was seen to have an impact on COPD.

The second test did not detect β coefficient decrements that were consistent with expectations when testing both factors and individual items.

4. Conclusion

The first test appears a useful technique for bolt-on selection after factor analysis. The second test does not appear appropriate and further assessment of the results is required. Datasets with a large number of questionnaire items alongside the EQ-5D are required to undertake this type of analysis.

Sources and acknowledgments:
(1) Longworth L, Yang Y, Young T, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. Southampton (UK): NIHR Journals Library; 2014 Feb. (Health Technology Assessment, No. 18.9.) Available from: https://www.ncbi.nlm.nih.gov/books/NBK261616/ doi: 10.3310/hta18090
(2) Finch AP, Brazier JE, Mukuria C et al. An exploratory study on using prinicipal component analysis, confirmatory factor analysis and structural equation modelling to identify bolt-ons. The EQ-5D case study. 2016. 33° EuroQol conference. Berlin