Using Discrete Choice Experiments (DCE) based approaches to produce anchored values: comparative results from application to EQ-5D-Y-3L

Tianxin Pan[1,2], Nancy Devlin[1], Brendan Mulhern[3], Richard Norman[2]

1. Health Economics Unit, Centre for Health Policy, Melbourne School of Population and Global Health, The University of Melbourne, 207 Bouverie Street, Carlton, VIC, 3053, Melbourne, Australia.
2. School of Population Health, Curtin University, Perth, WA, Australia.
3. Centre for Health Economics Research and Evaluation, University of Technology Sydney, Sydney, NSW, Australia.

Abstract

Objective: In the international protocol for valuing EQ-5D-Y, both DCE and cTTO are included. However, the role of DCE is to determine the relative importance of dimensions and severity levels. Little methodological work has been done to explore the feasibility and appropriateness of using DCE to generate values and convert to QALY scale. The Australian EQ-5D-Y-3L valuation study followed the international protocol in collecting both latent scale DCE data and cTTO data. Alongside the latent scale DCE, we also included DCE tasks containing either duration or dead. The aim of this paper is to explore and compare different DCE-based approaches to valuing EQ-5D-Y, and anchoring values onto the utility scale.

Methods: The choice experiment comprised three components: (1) latent scale DCE, following the Y protocol using 10 blocks and 15 choice sets per block, with each choice set including two health state options; (2) DCE + duration, using 15 blocks and 3 choice set per block, with each choice set including two EQ-5D-Y-3L health profile options, and a duration attribute (1, 4, 7 or 10 years); (3) DCE + dead – an unblocked design of 32 pairs, with each choice set involving comparisons of a 'severe' health state (all dimensions at least level 2) to being dead. Respondents were recruited by Survey Engine, and each answered 21 choice sets (15+3+3).

We explored four approaches (A-D). (A) uses DCE + duration as a 'stand-alone' approach, estimated using 1) an 11 parameter main effects model (i.e. duration and the 10 interactions between duration and level 2 and 3 of each of the five dimensions); and 2) a model introducing an additional interaction captured by N3 term; using the wtp Stata command to estimate QALY weights. We also analysed DCE latent scale data (using a mixed logit model allowing for unobservable random preference heterogeneity) and used three approaches to rescale these values onto a QALY scale: (B) a linear mapping model between latent scale coefficients and the utility decrements produced from Approach A; (C) anchoring on the pits state (33333) value obtained from Approach A; and (D) anchoring on dead, through a logit model to identify the relative position of health states and dead, and setting to zero the value of the health state with a 50% chance of being preferred to dead.

Results: A representative sample of 1002 adults completed DCE online between Dec 2021 and Jan 2022. DCE + duration data suggested PD was the most important dimension (with the largest overall decrement), followed by AD, MO, UA and SC. The N3 coefficient was negative but not statistically significant at the 10% level; the Likelihood-ratio test rejected the hypothesis that the interaction model outperformed the main effect model, supporting the conclusion of AIC and BIC. The estimated utilities from the main effect model range from -0.332 to 1. Results from the mixed logit model on latent DCE suggested different dimension ranking: PD, AD, UA, SC and MO. Utilities have different ranges when anchoring using approaches B (-0.319 to 1) and C (-0.332 to 1). Under approach D, the level decrements in all dimensions were smaller compared to previous approaches, resulting in a narrower range of utilities (0 to 1).

Conclusion: DCEs are feasible as a stand-alone approach to producing interpretable value sets for the EQ-5D-Y. However, different DCE approaches produce different results both with respect to the relative importance of dimensions and level decrements, and different value ranges. It is not clear what criteria should be used to choose between the alternative approaches.