VALUE IN HEALTH ▮ (2017) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

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

Handling Data Quality Issues to Estimate the Spanish EQ-5D-5L Value Set Using a Hybrid Interval Regression Approach

Juan M. Ramos-Goñi, MSc[1,8,*], *Benjamin M. Craig, PhD*[2], *Mark Oppe, PhD*[1], *Yolanda Ramallo-Fariña, MSc*[3,8], *Jose Luis Pinto-Prades, PhD*[4,5], *Nan Luo, PhD*[6], *Oliver Rivero-Arias, PhD*[7,8]

[1]*EuroQol Research Foundation, Rotterdam, The Netherlands;* [2]*University of South Florida, Tampa, FL, USA;* [3]*HTA Unit of Canary Islands Health Service, Canary Islands, Spain;* [4]*University of Navarra, Navarra, Spain;* [5]*Glasgow Caledonian University, Glasgow, UK;* [6]*National University of Singapore, Singapore;* [7]*University of Oxford, Oxford, UK;* [8]*Research Network on Chronic Diseases (REDISSEC), Madrid, Spain*

ABSTRACT

Background: The Spanish five-level EuroQol five-dimensional questionnaire (EQ-5D-5L) valuation study was the first to use the EuroQol Valuation Technology protocol, including composite time trade-off (C-TTO) and discrete choice experiments (DCE). In this study, its investigators noticed that some interviewers did not fully explain the C-TTO task to respondents. Evidence from a follow-up study in 2014 confirmed that when interviewers followed the protocol, the distribution of C-TTO responses widened. **Objectives:** To handle the data quality issues in the C-TTO responses by estimating a hybrid interval regression model to produce a Spanish EQ-5D-5L value set. **Methods:** Four different models were tested. Model 0 integrated C-TTO and DCE responses in a hybrid model and models 1 to 3 altered the interpretation of the C-TTO responses: model 1 allowed for censoring of the C-TTO responses, whereas model 2 incorporated interval responses and model 3 included the interviewer-specific protocol violations. For external validation, the predictions of the four models were compared with those of the follow-up study using the Lin's concordance correlation coefficient. **Results:** This stepwise approach to modeling C-TTO and DCE responses improved the concordance between the valuation and follow-up studies (concordance correlation coefficient: 0.948 [model 0], 0.958 [model 1], 0.952 [model 2], and 0.989 [model 3]). We recommend the estimates from model 3, because its hybrid interval regression model addresses the data quality issues found in the valuation study. **Conclusions:** Protocol violations may occur in any valuation study; handling them in the analysis can improve external validity. The resulting EQ-5D-5L value set (model 3) can be applied to inform Spanish health technology assessments.

Keywords: economic, health status index, life valuation, quality of life.

Introduction

In 2012, the EuroQol Group developed a new standardized protocol (version 1.0) to perform country-specific valuation studies for the five-level EuroQol five-dimensional questionnaire (EQ-5D-5L) using EuroQol Valuation Technology (EQ-VT) [1]. The EQ-VT protocol was developed to elicit health preferences through face-to-face interviews using two valuation techniques, the composite time trade-off (C-TTO) [2,3] and a discrete choice experiment (DCE) [4]. Each respondent completed C-TTO tasks for 10 EQ-5D-5L health states and forced-choice pair comparisons for seven pairs of EQ-5D-5L health states without duration. The C-TTO was a modified version of the traditional TTO technique [5,6], which used the traditional TTO technique for health states considered to be better than immediate death (BTD) and a lead-time TTO technique [7–9] for states considered to be worse than immediate death (WTD).

The C-TTO task entailed a series of consecutive and adapted choices terminating when respondents stated indifference. Because of the complexity of the task, the EQ-VT protocol included an example of this task (being in a wheelchair), which was designed to facilitate and standardize interviewers' explanations. In a previous publication, we described the Spanish EQ-5D-5L valuation study [10]. During this initial analysis, interviewer effects were identified, which were attributed to protocol violations by specific interviewers. Some interviewers did not explain the WTD sections of the C-TTO task and respondents may not have been aware of these sections, leading to fewer WTD values. In fact, evidence from a follow-up study performed in Spain [11], which used an updated protocol version, showed that when interviewers properly explained the WTD sections of the C-TTO task, a higher proportion of negative numbers were observed [12], altering the distribution of the C-TTO responses. In addition, some interviewers did not properly explain the wheelchair

* *Address correspondence to:* Juan M. Ramos-Goñi, EuroQol Research Foundation, Executive Office, C/Calvario 271-B 1IZQ, Canary Islands 38350, Spain.

E-mail: jramos@euroqol.org

1098-3015$36.00 – see front matter

https://doi.org/10.1016/j.jval.2017.10.023

example or showed only a few steps from the iterative procedure to respondents [12]. These participants may have responded imprecisely either because they were not aware of the full iterative procedure or to avoid the time and effort needed to reach their accurate indifference points (i.e., satisficing) [13]. We hypothesized in this study that the C-TTO responses in the Spanish EQ-5D-5L valuation study, although not being as precise as we had expected, still contain valuable information about health preferences from the Spanish population. We demonstrate that such information can be retrieved by assessing individual's paths during the iterative procedure when completing each C-TTO task. At this time, we have no reason to believe that the DCE responses in the valuation study were affected by the protocol violations in the C-TTO tasks.

The primary objectives of this article were to introduce an analytical approach based on hybrid interval regression models (jointly incorporating C-TTO and DCE responses), which updates our previous work [10] to handle the data quality issues commented earlier, and to produce an EQ-5D-5L value set for health technology assessments in Spain. Furthermore, we assessed the external validity of the resulting value set by comparing its estimates with those of a follow-up study.

Methods

Data

The Spanish EQ-5D-5L valuation study has been previously reported in the literature [10,12], and therefore we describe it only briefly here. The valuation study included 1000 face-to-face interviews conducted in 2012 following the EQ-VT protocol version 1 [1]. After applying exclusions, the analytical sample included 9730 C-TTO responses on 86 health states and 7000 DCE responses on 196 pairs of health states. The sample was representative of the Spanish general population with respect to age and sex.

We used C-TTO and DCE responses from a follow-up study conducted also in Spain in 2014 to assess the external validity of the models described later [11]. This follow-up study was performed in only one Spanish region (Canary Islands), and therefore it was not representative of the Spanish population. Nevertheless, it included the quality control process currently recommended by the EuroQol Group to improve data quality. The original aim of the follow-up study was to test the effect of adding a ranking task to the protocol and its results showed that this addition had no significant effect. Therefore, the data from all study arms of the follow-up study were used for external validation.

The C-TTO Iterative Procedure

The C-TTO task used an iterative procedure (Fig. 1) composed of a series of consecutive and adapted choices terminating when respondents stated indifference. Across its four sections, boxes indicate the possible C-TTO responses (i.e., values) and the arrows represent steps from one value to another. Each C-TTO task started (Start box) by asking whether the respondent preferred 10 years in full health or 10 years in the EQ-5D-5L state. If the respondent preferred 10 years in the EQ-5D-5L state (double arrow up from 1 to 1), the same question was asked again to confirm the extreme value. If the respondent preferred 10 years in full health over 10 years in the EQ-5D-5L state (i.e., double arrow down from 1 to 0), the next question was whether the respondent preferred 0 years in full health (i.e., die immediately) or 10 years in the EQ-5D-5L state.

In the iterative procedure (Fig. 1), the "immediate death" question separated the BTD and WTD scenarios (0 at center left). If the respondent preferred 10 years in the EQ-5D-5L state (i.e., BTD state; double arrow up from 0 to 0.5), the next question was whether the respondent preferred 5 years in full health or 10 years in the EQ-5D-5L state. If the respondent preferred to die immediately over 10 years in the EQ-5D-5L state (i.e., WTD state; double-dash arrow from 0 to 0 on the left), the next question was a confirmation of the response but in a lead-time TTO scenario, that is, 10 years in full health versus 10 years in full health followed by 10 years in the EQ-5D-5L state. If the respondent preferred 10 years in full health (double arrow down from 0 to −0.5), the next question was whether the respondent preferred 5 years in full health or 10 years in full health followed by 10 years in the EQ-5D-5L state. If the respondent preferred 10 years in full health followed by 10 years in the EQ-5D-5L state (double arrow horizontal from 0 to 0.05), the iterative procedure changed back to the BTD scenario and the next question asked whether the respondent preferred 0.05 years in full health or 10 years in the EQ-5D-5L state. After these initial steps (double arrows to −0.5, 0.05, 0.5, and 1), the iterative procedure imposed 1-year increments/decrements (i.e., single arrows) followed by half-year corrections (i.e., single-dash arrows) depending on the respondent's preferences. Respondents who visited the BTD scenario after the three initial steps and switched later to the WTD scenario, that is, preferred to die immediately over 10 years in the EQ-5D-5L state (double-dash arrow from 0 to 0 on the right), also had to complete the WTD confirmatory question. This was, however, only once per state.

Although respondents were allowed to go from −0.05 to 0 (immediate death), they were not allowed to go from 0 to −0.05 because of a survey programming error (elbow arrows from 0 to −0.5).

Analysis

Modeling

In a previous publication, we developed and estimated a hybrid model using C-TTO and DCE responses [10]. This initial hybrid model (model 0) assumed normality, homoscedasticity, and that respondents completed the C-TTO tasks accurately. In this study, we followed an analytical approach that relaxed the initial assumption about the accuracy of the C-TTO responses. Specifically, we reconsidered censoring, respondent uncertainty, and protocol violations on the C-TTO tasks [12] as follows.

Censoring of C-TTO responses at −1. The C-TTO task had a minimum TTO value bounded at −1 by design and produced responses in the range [−1, 1]. Nevertheless, feedback from interviewers suggested that some respondents would have responded beyond −1 if allowed, which corroborates the findings of Attema et al. [14]. Because values may be in the range $(-\infty,1]$, we relaxed this lower bound assumption and considered responses at the lower bound (−1) to be censored, similar to the open intervals produced by DCE responses ($A > B$) [15].

Inaccuracy of C-TTO responses. The EQ-VT recorded the full path in the C-TTO iterative procedure for each state presented. Using these paths, we built intervals for each state for each respondent. Instead of considering only the final indifference point, this interval assessment used all path information in a conservative manner. Specifically, we observed four response patterns (see examples of each in Supplemental Materials 1 found at http://dx.doi.org/10.1016/j.jval.2017.10.023):

1. *Straight-lining:* This refers to an uninterrupted path, only up or only down, that leads to extreme values of a section, namely, 1, 0.95, 0.05, 0, −0.05, −0.95, and −1, using the minimum number of steps. We refer to this response behavior as straight-lining