EQ-5D-5L valuation project for the Spanish population

Ramos-Goñi JM [1,2], JM Cabasés[3], M Errea[3], O Rivero-Arias[2,4],JL Pinto-Prades[5]

[1] HTA Unit of Canary Island Health Care Service (FUNCIS)

[2] REDISSEC

[3] Universidad Pública de Navarra

[4] University of Oxford

[5] Glasgow Caledonian University

XX Encuentro de Economía Pública
Sevilla, 31 de enero y 1 de febrero de 2013

Corresponding authors:

Juan M. Ramos-Goñi
HTA Unit of Canary Island Health Care Service and CIBERESP
jramos@sescs.es

Juan M Cabasés
Departamento de Economía, Universidad Pública de Navarra
jmcabases@unavarra.es

Abstract

The latest EQ-5D instrument, EQ-5D-5L, needs new country-specific valuation studies to obtain a value set adapted to the characteristics of the updated instrument. Eight countries from Europe, North and South America and Asia have participated on pilot exercises to develop a final protocol which will be commonly used to perform the valuation studies in each country. Spain is the first country where this protocol has been introduced as part of the Valuation Project for the Spanish Population. This discussion paper reports a descriptive overview of EQ-5D-5L valuation results in the Spanish population and its preliminary results.

The survey has a two-stage sample plan. The first stage concentrates on the selection of Spanish regions. The 50 Spanish regions were ordered by population size, and the first 20 regions covering 80% of the total Spanish population were selected. In the second stage, a simple random sampling strategy on each of the selected regions was conducted. The sample size on each region was calculated multiplying the total sample size (1,000) by the percentage of the population on the region respect the total population of the select 20 regions. Data will be collected between 21st May and 15th June 2012 using the final agreed protocol by the EuroQol group. Primary data collection will be conducted by a specialist survey company with a second company conducting a strict quality control process to ensure interviews and data collection of highest quality.

The final survey has three blocks of questions. The first block includes patient characteristics (age, gender, socioeconomic status and so on), and a respondent valuation of own health using the EQ-5D-5L. The second block contains 10 composite time trade-offs (TTO) questions, for states better than death classic TTO is used and the "Lead Time" TTO is used for those health states considered worse than death. In this case the lead time (period in full health) is 10 years and the time in the disease is another 10 years to be comparable with the classical TTO were there are 10 years in full health and 10 years in the disease. The last block contains 7 discrete choice experiments questions, where the participant has to choose between two states. Finally some questions about the difficulties of the survey are also included.

Descriptive statistics of the final sample are reported. A detailed overview of summary statistics for the health state valuations is included. Different models were explored; hybrid TTO and DC model and DC conditional logistic rescaled with TTO values were used and compared.

Introduction

The EQ-5D questionnaire was developed by the EuroQol group to estimate the health related quality of life of patients (HRQoL). The EQ-5D has been tested to be a valid and reliable generic instrument, thus it has been used to assess HRQoL in many different populations. However, certain features of the instrument have been challenged, such as lack of descriptive richness or restricted discriminatory power. This has given rise to an updated version of the EQ-5D instrument: EQ-5D-5L. The standard EQ-5D-3L descriptive system is being compared with a new 5L version, as evidence of confidence in the potential of the instrument.

The increase from 3 to 5 levels implies an increase in the number of possible health states described from 243 for the EQ-5D-3L to 3,125 for the EQ-5D-5L. Parallel to the development work on the 5L itself, the Valuation Task Force of the EuroQol Group has been investigating the potential for improving the valuation methodology used to value 3L health states. Over time, these investigations led to the development of a new valuation protocol [1], and a new Digital Valuation Software [2].

The Spanish EuroQol members have take the initiative to be the first that use that protocol and software to perform the first country specific EQ-5D-5L valuation study.

The objective of this paper is to describe the Spanish Valuation process and its preliminary results.

Methods

Design

A two part design was used on this study. Following the EuroQol valuation protocol the interview has 3 main parts. On the first part the participants is requested to value their own health by EQ-5D-5L, and is asked by their sex, age, and whether they are familiarized with severe illness by their self or family or caring others.

The second part of the interview consists of 10 composite Time to Trade-off (TTO) exercises. The new designed composite TTO is a merger between the classical TTO and the Lead Time TTO tested on the pilots studies previously performed. For the states considered better than death the classical TTO version is used. For the states considered worse than death and in order to avoid the Dolan's transformation or other suggested on the literature [3,4] the Lead Time TTO was used starting at the 0 value. The ratio between the Lead time and the life time was 1:1 in order to make comparable the positive and negative values and to have symmetry around the 0 value.

The last part of the interview consists of 7 Discrete Choice (DC) experiments, where the participants were requested to chose the preferred state of the two presented A and B. Each state is defined by its EQ-5D-5L profile showed by the corresponding labels of each EQ-5D domain.

Sampling and Data collection

The sample size calculations was performed prior on the protocol to be 1,000 of participants on any country which decide to run a valuation study. Those calculations were made in basis to obtain a 0.01 of standard deviation on the possible coefficients of a linear model for the TTO responses. Details of the calculations can be found at Oppe et al [Reference].

Once the total sample size if fix, a two-step stratification procedure were design to obtain a representative sample of the Spanish population. The first stratification step contemplates the Spanish regions. The 50 Spanish regions were ordered by its population size and the first 26 regions were selected due to they represents more than 75% of the total Spanish population size. Since not all Autonomous Community were selected and some of them were over-included, the regions with lower population size of the over-represented Autonomous Community were replaced for the regions on the Autonomous Community with higher Population size. The table 1 show the details of the sample size with all Autonomous Community included.

Table 1.- Sample size by region	
Region	Sample size
ALICANTE	51
ASTURIAS	28
BADAJOZ	20
BALEARES	28
BARCELONA	142
BIZKAIA	30
CADIZ	33
CANTABRIA	20
CORUÑA	22
GIRONA	20
GRANADA	25
GUIPUZKOA	20
LA RIOJA	20
LAS PALMAS	29
MADRID	166
MALAGA	43
MURCIA	33
NAVARRA	20
PONTEVEDRA	25
SEVILLA	51
TARRAGONA	21
TENERIFE	19
TOLEDO	20
VALENCIA	68
VALLADOLID	20
ZARAGOZA	26
Total	1000

The second step of the stratification was performed in basis of the Spanish population structure by sex and age.

The recollecting data process was made by a firm 21rd March-13th July, 2012. 32 interviewers using the offline application were previously recruited and formed on a briefing session to perform the interviews to the sample, following the EoruQol protocol [1].

During the interviews, the phone numbers of the participants were saved on the data base. This numbers were used to perform a quality control of the sample. The 15% of the sample were randomly selected and an independent firm of the firm for the recollecting data process, were contracted to make calls to that phone numbers to ensure the correct recollecting data process. Just some minimal questions were made: date of interview, age, sex, and how was it.

Data analysis

The analysis consists in two main parts. The first part of the analysis is a descriptive analysis of the whole sample.

The second part of the analysis has the aim to find a model to estimate the EQ-5D-5L utilities by profile. Two main approaches have been investigated.

The first one consists on the called the logistic DC conditional model anchor in TTO worth state. In this case the TTO and DC data are modeled separated.

One the one hand, TTO data is analyzed by mean of linear regression with robust estimation.

$$(1) \quad Y_i = \sum_{j=1}^{n} x_{ij} \cdot \beta_j + \varepsilon_i$$

Many specifications for the model were tested in order to choose a model with absolute consistency.

On the second hand, DC data utilities are not directly observable. They have to be calculated from the choice exercise -choosing state A or state B-. We assume that the participants choose the scenario that gives them a higher utility, so this can be modeled as follows:

$$(2) \quad P(Y_i = A) = P(U_{iA} > U_{iB}) = P(V_{iA} + \varepsilon_{iA} > V_{iB} + \varepsilon_{iB}) = P(V_{iA} - V_{iB} > \varepsilon_{iB} - \varepsilon_{iA})$$

Where Y_i is a variable representing the choice of the participant i between A or B. U_{iA} represents the utility of participant i for the state A. Assuming a utility decomposition in two part, one is explainable V_{iA} plus an error ε_i. If errors are assumed to be random and following type 1 extreme value distribution, then this leads to the following conditional logistic model[4,10,11]:

$$(3) \quad P(Y_i = A) = \frac{e^{\mu \cdot V_{iA}}}{e^{\mu \cdot V_{iA}} + e^{\mu \cdot V_{iB}}}$$

Assuming that the component V of the utility can be explained with an additive model:

$$(5) \quad V_{i\cdot} = \sum_{j=1}^{J} X_{i\cdot j} \cdot \beta_j$$

Where X_{iAj} are 20 dummies {0, 1}, per participant i, representing the severity level j for each domain of EQ-5D-5L for the state A. β_j represent the coefficient for each independent variable j.

Substituting the expression for $V_{i\cdot}$:

$$(6) \quad P(Y_i = A) = \frac{e^{\mu \cdot (\sum_{j=1}^{J} X_{iAj} \cdot \beta_j)}}{e^{\mu \cdot (\sum_{j=1}^{J} X_{iAj} \cdot \beta_j)} + e^{\mu \cdot (\sum_{j=1}^{J} X_{iBj} \cdot \beta_j)}}$$

This makes it possible to estimate the coefficients of the model and thus to extrapolate utilities that have not been observed within the population, using the linear part of the $DC_{Conditional}$ model.

The utilities obtained from the linear part of the model above are in an arbitrary scale. In order to rescale the utilities from $DC_{Conditional}$ model into the TTO scale, TTO extreme negative value

(55555) was used as the anchoring point for $DC_{Conditional}$ 55555 state. Finally, the model was restricted to 1 representing full health.

The second main modeling exercise was a hybrid model of both commented above. In this case TTO and DC data are combined as one data file to obtain a common vector of coefficients for both types of data. The model tested was proposed by Oppe and Van Hout [5]. As they describe on that paper the hybrid model could be estimated by mean of maximum likelihood estimation on the product of likelihood function of both approaches. See Oppe and Van Hout [5] for details.

The analysis was performed on STATA 11. For linear regression of TTO data, regress (OLS), xtreg (Random effects), rreg (robust model) commands were tested with different specifications. For Conditional logistic model of DC data clogit command was used with different specifications as well. Finally for the hybrid model of TTO and DC data a specific program was developed using ml (maximum likelihood) command with d0 estimator and many specification for the structure of the model.

Results

Sample characteristic

The mean sample age is 43.9±17.3 years old and 46.8% (461) are males. The 14.5% (143) of the participants have experimented severe illness. The 63.1% (623) have experimented severe illness by some member of his/her family and the 34.6% (341) have experimented severe illness caring for others. The table 2 shows the distribution of self-reported health by domains from respondents, in percentage.

Table 2.- Distribution of participants by his/her own health.

Levels	MOBILITY	SELF CARE	USUAL ACTIVITIES	PAIN/ DISCOMFORT	ANXIETY/ DEPRESSION
NO PROBLEMS	88,7% (875)	95,6% (943)	91,3% (900)	68,9% (679)	79,5% (784)
SLIGHT PROBLEMS	7,1% (70)	3,1% (31)	6% (59)	22% (217)	15% (148)
MODERATE PROBLEMS	3,3% (33)	1% (10)	2,2% (22)	7,4% (73)	3,9% (38)
SEVERE PROBLEMS	0,7% (7)	0,1% (1)	0,4% (4)	1,6% (16)	1% (10)
UNABLE/EXTREME	0,1% (1)	0,1% (1)	0,1% (1)	0,1% (1)	0,6% (6)

Table 3 shows the number of responses per each TTO state. The profile 55555 has 986 observations as it was force to be included for all blocks of states. The profiles 21111, 12111, 11211, 11121 and 11112 were force to have double number of observations than the rest of profiles, trying to reduce the variability for the upper health states. Those profiles have around 200 observations and the rest of states have around 100 observations.

Table 3.- Number of observations per profile included on the TTO design

Profile	N	Profile	N	Profile	N
11112	207	21345	96	32314	104
11121	203	21444	98	32443	96
11122	88	22434	92	33253	104
11211	188	23152	96	34155	96
11212	96	23242	104	34232	88
11221	104	23514	96	34244	96
11235	104	24342	104	34515	104
11414	98	24443	103	35143	98
11421	100	24445	88	35245	104
11425	92	24553	92	35311	88
12111	192	25122	100	35332	92
12112	96	25222	98	51152	92
12121	96	25331	98	51451	104
12244	100	42115	92	52215	96
12334	104	42321	88	52335	88
12344	105	43315	103	52431	103
12513	105	43514	96	52455	100
12514	104	43542	96	53221	105
12543	96	43555	88	53243	98
13122	92	44125	105	53244	98
13224	88	44345	105	53412	104
13313	100	44553	96	54153	103
14113	103	45133	96	54231	104
14554	105	45144	104	54342	105
15151	103	45233	100	55225	104
21111	196	45413	92	55233	100
21112	105	31514	98	55424	96
21315	103	31524	103	55555	986
21334	104	31525	100		

The figure 1 shows the frequencies of responses by profile included in the TTO design. Attending to the health states with lower health problems the empirical distributions show that

many people is not willing to trade-off and just some of them are willing to trade-off. However, many participants consider the health with extreme problems as the death, as the 0 spikes are showed.

The table 4 presents the estimates for the coefficients of some of the models tested. Just raw data was analyzed and no observations were eliminated for the preliminary analysis.

Hybrid model and Conditional logistic model have inconsistencies. In one hand, the hybrid model has inconsistencies between the levels 2 and 3 of self-care domain, and between the levels 4 and 5 of usual-activities domain. In the other hand the conditional logistic model has a little inconsistency between level 2 and 3 of usual-activities domain.

The TTO model is perfectly consistency. The index value for the profile 55555 using TTO model is -0.13 and when the Hybrid model is used this value is - 0.19.

Every other model tested had inconsistencies between levels of the domains self-care or usual-activities, mostly between level 2 and 3 of usual-activities domain.

Table 4.- Parameter estimates for the models based on data derived by discrete choice experiment (DCE) and time trade-off (TTO)							
	Hybrid		TTO		Conditional Logit		Re-scaled Conditional Logit (-12.03-1)/(-0.13-1)
	Number of obs = 23,664 Wald chi2(22) = 20,386 Log likelihood = -6,817.68 Prob > chi2 = 0.0000		Number of obs = 9860 F(21, 9838) = 243.85 Prob > F = 0.0000		Number of obs = 13,804 LR chi2(21) = 4,560.52 Prob > chi2 = 0.0000 Log likelihood = -6827.36 Pseudo R2 = 0.2504		
	B	St. Error	B	St. Error	B	St. Error	B
mo2	-0,048	0,016	-0,007	0,016	-0,874	0,089	-0,076
mo3	-0,086	0,017	-0,022	0,015	-1,078	0,094	-0,093
mo4	-0,239	0,019	-0,169	0,016	-2,296	0,093	-0,199
mo5	-0,243	0,017	-0,225	0,015	-3,131	0,106	-0,271
sc2	-0,082	0,015	-0,024	0,015	-0,297	0,090	-0,026
sc3	-0,045	0,019	-0,025	0,016	-0,578	0,098	-0,050
sc4	-0,181	0,018	-0,146	0,016	-1,420	0,096	-0,123
sc5	-0,189	0,017	-0,180	0,016	-1,660	0,092	-0,144
ua2	-0,082	0,016	-0,016	0,016	-0,361	0,091	-0,031
ua3	-0,110	0,017	-0,057	0,015	-0,331	0,093	-0,029
ua4	-0,191	0,018	-0,133	0,015	-1,096	0,093	-0,095
ua5	-0,147	0,017	-0,142	0,014	-1,335	0,094	-0,116
pd2	-0,099	0,014	-0,039	0,015	-0,677	0,096	-0,059
pd3	-0,104	0,019	-0,047	0,016	-0,945	0,093	-0,082
pd4	-0,186	0,016	-0,181	0,014	-2,361	0,096	-0,204
pd5	-0,312	0,017	-0,269	0,015	-3,655	0,103	-0,316
ad2	-0,115	0,015	-0,041	0,015	-0,631	0,099	-0,055
ad3	-0,145	0,018	-0,089	0,016	-1,250	0,094	-0,108
ad4	-0,248	0,017	-0,174	0,015	-2,524	0,103	-0,219
ad5	-0,294	0,017	-0,238	0,014	-3,255	0,105	-0,282
N2	-	-	-0,010	0,018	-0,145	0,094	-0,013
no1	-	-	-0,075	0,017	-	-	

Discussion

This paper presents the preliminary results for the first EQ-5D-5L valuation study around the world. The Valuation study was conducted over a representative sample of the Spanish population.

As it is mentioned above these results are preliminary and only model over raw data are presented. No transformation and eliminated observations were applied to the data due the limited time to realize the analysis.

Looking the data descriptive analysis the global health of the participants is good and just 14% of them have experimented severe illness. However, 63% of the sample have experimented severe illness by a family member and 34% by caring others.

The distribution of participant's moves and time required to realize the TTO exercises is a right asymmetric distribution, showing that many people do the exercises on a mean and median of six moves, but some people have doubts about some health states and for the time occurs something similar.

The distribution of time associated with DC experiment could not be analyzed due to the Spanish design included an extra task and time was computed overall.

The estimated coefficients for the models presented on this discussion paper looks like just more robust estimation eliminating some extreme values will conduct to a consistency model for each hybrid and DC conditional logistic model as it occurs with the TTO model when a robust estimation is used. However, comparing the TTO lowest index value for the profile 55555 with the previously published by Badia et at [6] for the 3L version of EQ-5D it looks like the composite TTO version used in this study produce higher results than the classical used by Badia et al (-0.13 and -0.654).

The open questions are:

1) Should any levels at any domain be collapse to reach the consistency?
2) Should get estimates from the hybrid model or DC re-scaled model? Or get estimates just from TTO data which is a totally consistent model with robust estimations?
3) Is the composite TTO version comparable with the classical TTO version?
4) More?

References

1. Paul Krabbe, Nancy Devlin, Mark Oppe, Grahame Wilkinson, Paul Kind, Ben van Hout. Protocol to value the EQ-5D-5L
2. http://www.valuationstudy.org/
3. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997 Nov;35(11):1095-108.
4. Craig BM, Oppe M. From a different angle: a novel approach to health valuation. Soc Sci Med. 2010 Jan;70(2):169-74. Epub 2009 Oct 31.
5. Mark Oppe, Ben van Hout. The optimal hybrid: Experimental design and modeling of a combination of TTO and DCE.
6. Badia X, Roset M, Herdman M, Kind P. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. Med Decis Making. 2001 Jan-Feb;21(1):7-16.