Volume 12 • Number 8 • 2009
VALUE IN HEALTH

# International Comparisons in Valuing EQ-5D Health States: A Review and Analysis

Richard Norman, MSc,[1] Paula Cronin, MPH,[1] Rosalie Viney, PhD,[1] Madeleine King, PhD,[2] Deborah Street, PhD,[1,3,4] Julie Ratcliffe, PhD[5]

[1]Centre for Health Economics Research and Evaluation (CHERE), Faculty of Business, University of Technology, Sydney, NSW, Australia; [2]Psycho-oncology Co-operative Research Group (PoCoG), University of Sydney, Sydney, NSW, Australia; [3]Centre for the Study of Choice (CenSoc), Faculty of Business, University of Technology, Sydney, NSW, Australia; [4]Department of Mathematical Sciences, Faculty of Science, University of Technology, Sydney, NSW, Australia; [5]Flinders Clinical Effectiveness, Flinders University, Adelaide, Australia

## ABSTRACT

**Objective:** To identify the key methodological issues in the construction of population-level EuroQol 5-dimensions (EQ-5D)/time trade-off (TTO) preference elicitation studies.

**Method:** This study involved three components. The first was to identify existing population-level EQ-5D TTO studies. The second was to illustrate and discuss the key areas of divergence between studies, including the international comparison of tariffs. The third was to portray the relative merits of each of the approaches and to compare the results of studies across countries.

**Results:** While most articles report use of the protocol developed in the original UK study, we identified three key areas of divergence in the construction and analysis of surveys. These are the number of health states valued to determine the algorithm for estimating all health states, the approach to valuing states worse than immediate death, and the choice of algorithm. The evidence on international comparisons suggests differences between countries although it is difficult to disentangle differences in cultural attitudes with random error and differences as a result of methodological divergence.

**Conclusions:** Differences in methods may obscure true differences in values between countries. Nevertheless, population-specific valuation sets for countries engaging in economic evaluation would better reflect cultural differences and are therefore more likely to accurately represent societal attitudes.

*Keywords:* cost-utility analysis, EQ-5D, health economics methods, health-related quality of life.

## Introduction

Cost-utility analysis (CUA), where outcomes are measured in terms of quality-adjusted life-years (QALYs), is the main approach used to measure and value the impacts of treatments. The US Panel on Cost-Effectiveness in Health and Medicine recommends the use of QALYs [1]; the UK National Institute of Health and Clinical Excellence has most commonly used CUA [2,3] and has recently recommended that it should be the preferred outcome measure; and CUA is increasingly used in Australia in the evaluation of pharmaceuticals and medical services. In the recently released PBAC guidelines, a preference is expressed for the use of CUA [4].

While CUA is simple in concept, it presents challenges in practice. QALYs are designed to allow comparisons across interventions with disparate outcomes across different health-care conditions and population groups. Eliciting valuations for all health states that may be relevant to a disease or intervention is time consuming and costly, and comparison of valuations across interventions and diseases requires comparability of methods. Multiattribute utility instruments (MAUIs), which comprise a generic descriptive quality of life instrument and a scoring algorithm that covers all health states described by the instrument (e.g., the EQ-5D, the Short Form-6 dimensions (SF-6D), Health Utilities Index, and Assessment of Quality of Life), have facilitated comparability [5,6]. The scoring algorithm for these instruments is usually generated from a stated preference experiment, typically time trade-off (TTO), standard gamble conducted in a population sample The key advantage of the MAU approach is that it provides community-based valuation of health states for patients who are experiencing the state.

The role of MAUIs in economic evaluation is increasing. For example, the National Institute of Clinical Excellence has recommended the use of the EQ-5D, and the Pharmaceutical Benefits Advisory Committee in Australia has stated a preference for utility weights generated from the use of a MAUI in a clinical trial setting (without specifying a preference for a particular MAUI). Nevertheless, recent reviews have noted that there are significant differences in the performance of different MAUIs [7], which can be attributed to differences in the dimensions covered by the instruments, differences in preference elicitation techniques, and differences in the methods used to derive the scoring algorithm. These differences can have significant impact on valuations of health states and the resulting cost-effectiveness of interventions [8]. There has been relatively little critical appraisal of the methods of development of MAUIs scoring algorithms. In this article, we examine these issues by considering the EQ-5D [9]. We chose the EQ-5D because it is widely used, and there have been a number of different studies undertaken to develop country-specific scoring algorithms. Because the focus of this review is on one MAUI, we do not consider the psychometric aspects of the instrument but, rather, focus on the methods for development of the scoring algorithm. Many of the issues we raise are relevant to other MAUIs.

## Overview of the EQ-5D

The EQ-5D is a tool developed by the EuroQol group (Rotterdam, The Netherlands) (www.euroqol.org) and has five dimensions

*Address correspondence to:* Richard Norman, Centre for Health Economics Research and Evaluation (CHERE), Faculty of Business, University of Technology, Sydney, PO BOX 123, Broadway, Sydney, NSW 2007, Australia. E-mail: richard.norman@chere.uts.edu.au

10.1111/j.1524-4733.2009.00581.x

  1098-3015/09/1194 1194–1200

**Table 1** The EQ-5D

| Dimension | Description |
|---|---|
| Mobility | |
| 1 | I have no problem in walking about. |
| 2 | I have some problems in walking about. |
| 3 | I am confined to bed. |
| Self-care | |
| 1 | I have no problems with self-care. |
| 2 | I have some problems washing and dressing myself. |
| 3 | I am unable to wash and dress myself. |
| Usual Activities | |
| 1 | I have no problems with performing my usual activities. |
| 2 | I have some problems with performing my usual activities. |
| 3 | I am unable to perform my usual activities. |
| Pain/Discomfort | |
| 1 | I have no pain or discomfort. |
| 2 | I have moderate pain or discomfort. |
| 3 | I have extreme pain or discomfort. |
| Anxiety/Depression | |
| 1 | I am not anxious or depressed. |
| 2 | I am moderately anxious or depressed. |
| 3 | I am extremely anxious or depressed. |

intended to represent the major areas in which health changes can manifest. These areas are mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension contains three levels, classified as "No Problems," "Some Problems," and "Extreme Problems." Details are shown in Table 1. Thus, there are $3^5 = 243$ potential states in the descriptive system. The TTO approach is used to value a selection of these states and then to impute values for the remainder using simple regression. The use of TTO for valuing EQ-5D states is well described in other works [10,11]. For states considered to be preferable to immediate death, a respondent is faced with a choice between 10 years of a particular chronic health state defined in EQ-5D space with a period of $x$ years in full health. The aim of the TTO is to identify a value of $x$ for which the individual is indifferent to the choice. The value for the better-than-death health state is defined as $x/10$.

Regarding Table 1, it should also be noted that we will treat health states with the same levels as identical throughout this article (e.g., health state 12321 is the same irrespective of language). As of March 2009, the EQ-5D has been translated into 100 different languages (with a further 24 awaiting ratification). The comparability of versions is a reasonable assumption because all translations are reviewed by EuroQol Group members and ratified by the EuroQol Group Executive Committee. The Euroqol Web site states that translation consists of two forward translations of the EQ-5D English source version, two back translations, lay assessment, and the production of a full report describing each stage of the process (http://www.euroqol.org).

Our analysis of this EQ-5D/TTO approach involves two strands: first, we look at how to elicit societal valuations for EQ-5D states under the York Research Group on the Measurement and Valuation of Health TTO protocol [9]. We begin by identifying some key themes and issues that run across the population valuation studies. Then, we look at international comparisons and discuss whether it is necessary to provide nationality-specific tariffs for the EQ-5D valuation system.

## Methods

The initial target of this study was to identify all large general population valuations studies employing the EQ-5D as the tool for describing health. EMBASE and MEDLINE were searched for such articles. To be considered for inclusion, the analysis had to present primary research in English and be published since 1995. Because it was expected that a proportion of good quality reports may be unavailable in peer-reviewed publications, the reference lists of articles identified in the main search were used to identify further studies. Because all of these identified nonpeer-reviewed publications were available on the EuroQol Web site (http://www.euroqol.org), the list of EuroQol Plenary Meeting Proceedings was scanned for further studies relevant to this work. To be included, a study had to attempt to value all 243 states described by the EQ-5D. Beyond this constraint, we chose to be conservative in our approach to exclusion because we were seeking to identify divergence in approach.

For each included study, details most relevant to the analysis of the methods used were identified. Key areas for discussion were selected. These areas were the precise formulation of the algorithm, the number of states directly valued in the survey to generate weights, the method to value states worse than death, the influence of time preferences of results, and international comparisons in predicted values across EQ-5D space.

The algorithms were compared by expanding the approach used by Busschbach et al. [12], who compare the directly valued states in the UK, Germany, and Spain. For this, Busschbach et al. used the UK results as the benchmark. The predicted preference scores for the states under the UK algorithm were then ranked in descending order. The preference scores under each of the other algorithms are generated by using the same ordering as in the UK study. We extended this approach by including all identified algorithms. Thus, we can identify any tendency for countries to trade off quantity of life for quality of life, and identify whether countries differ in their relative valuations of the five dimensions.

## Results

10 articles [11,13–21] that met the inclusion criteria were identified, of which eight were published in peer-reviewed journals. These are listed in Table 2. It should be noted that there are, at present, no such results for Canada or Australia, two countries strongly supportive of the use of CUA in health-care decision-making. Two studies utilized the visual analog scale (VAS) as the primary method of valuation [14,15]. Although this technique is widely used in preference elicitation more generally, the age of the two VAS studies in this area suggests that it has been superseded by the TTO although work by Parkin and Devlin suggests that the VAS remains a valuable tool [22].

Three significant methodological differences emerged regarding the survey structure and the development of the algorithm. The first regarded the number of states that need to be directly valued to estimate valuations for the complete EQ-5D space. The second is the approach to valuing states considered to be worse than death. The third is the choice of the algorithm to model those states not directly valued. There were a number of additional issues that might also be considered such as the validity of the TTO method and the assumption of constant proportional trade-off that it is founded on. Nevertheless, it was felt that this had been adequately covered elsewhere [23,24,25].

### *The Number of Directly Valued States*

Given that the EQ-5D has 243 individual possible states, it is unsurprising that no study has attempted to ask respondents to directly value each of these states. Therefore, the pertinent question becomes how best to form a representative fraction of the entire space that allows a good estimation of the remainder of the EQ-5D states in whichever way that is defined. Two approaches have been adopted to form this representative fraction. The