VALUE IN HEALTH ▮ (2017) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

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

Comparing 15D Valuation Studies in Norway and Finland—Challenges When Combining Information from Several Valuation Tasks

Yvonne Anne Michel, Dipl-Psych[1,*], *Liv Ariane Augestad, MD, PhD*[1], *Kim Rand, PhD*[1,2]

[1]*Department of Health Management and Health Economics, Medical Faculty, Institute of Health and Society, University of Oslo, Oslo, Norway;* [2]*Health Services Research Centre, Akershus University Hospital, Lørenskog, Norway*

ABSTRACT

Background: The 15D is a generic preference-based health-related quality-of-life instrument developed in Finland. Values for the 15D instrument are estimated by combining responses to three distinct valuation tasks. The impact of how these tasks are combined is relatively unexplored. **Objectives:** To compare 15D valuation studies conducted in Norway and Finland in terms of scores assigned in the valuation tasks and resulting value algorithms, and to discuss the contributions of each task and the algorithm estimation procedure to observed differences. **Methods:** Norwegian and Finnish scores from the three valuation tasks were compared using independent samples *t* tests and Lin concordance correlation coefficients. Covariance between tasks was assessed using Pearson product-moment correlations. Norwegian and Finnish value algorithms were compared using concordance correlation coefficients, total ranges, and ranges for individual dimensions. Observed differences were assessed using minimal important difference. **Results:** Mean scores in the main valuation task were strikingly similar between the two countries, whereas the final value algorithms were less similar. The largest differences between Norway and Finland were observed for depression, vision, and mental function. **Conclusions:** 15D algorithms are a product of combining scores from three valuation tasks by use of methods involving multiplication. This procedure used to combine scores from the three tasks by multiplication serves to amplify variance from each task. From relatively similar responses in Norway and Finland, diverging value algorithms are created. We propose to simplify the 15D algorithm estimation procedure by using only one of the valuation tasks.

Keywords: 15D, health-related quality of life, value algorithm, visual analogue scale.

Introduction

The 15D is a generic preference-based instrument used to measure health-related quality of life for estimating quality-adjusted life-years in health economic analyses [1]. The 15D was developed in Finland in the late 1970s. Values for 15D health states were derived in general population valuation studies using a set of valuation tasks based on the visual analogue scale (VAS) [2–4]. The 15D has been translated into 30 languages, including Norwegian, and more than 400 articles have been published using the instrument [5,6], 140 of which were in the last 5 years. The 15D is featured alongside the EuroQol five-dimensional questionnaire (EQ-5D-5L), the Assessment of Quality of Life (AQoL), the health utilities index (HUI), and the six-dimensional health state short form (SF-6D) in textbook presentations of health-related quality of life and instruments for measuring quality-adjusted life-years [7,8]. Because the 15D descriptive system covers many dimensions of health, it is often used in studies comparing such instruments. As such, it was recently part of a large multi-instrument comparison survey comparing five multi-attribute utility instruments [9].

Preferences for health states are assumed to vary between cultures. To capture differences in health state preferences between countries, country-specific algorithms are recommended by national guidelines [10–12]. Country-specific 15D value algorithms have been developed in Finland and Denmark [4,13]. The relevance and legitimacy of country-specific algorithms depend on their ability to adequately reflect the health state preferences of particular populations. Culture-dependent differences in health state preferences are still openly debated. Earlier research explores how country-specific differences in wealth, income, religion, health expenditure, and cultural factors such as power distance and individualism explain differences in preferences [14]. It is also possible that health state preferences change over time.

Respondents' health state preferences are not the only driver of differences in value algorithms. Valuation studies include

* *Address correspondence to:* Yvonne Anne Michel, Department of Health Management and Health Economics, Institute of Health and Society, University of Oslo, P.O. Box 1089, Blindern, Oslo 0318, Norway.

E-mail: y.a.michel@medisin.uio.no.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2017.09.018

choices about which methods to use and how to use them. Each valuation method raises questions about how to present the task, which visual aids to use, and which health states to value. There are also considerable differences in how case exclusion is handled [15]. The mode of administration can influence the data, and translation procedures could be a potential source of methodological variation [16]. Norman [17] highlights that "[t]he uncertain element in interpreting [different algorithms] is to identify whether the differences in models are a result of genuine differences in national attitudes toward ill health or whether they are the product of different study designs."

To which extent differences in health state preferences are driven by cultural or methodological variation remains unknown. Although there is a growing body of literature describing how methodological choices influence time trade-off–derived values for the EuroQol five-dimensional questionnaire [7,14,16,18], less is known about values derived for other instruments and other valuation methods. 15D algorithm values are derived by combining information from three VAS-based valuation tasks. Little is known about how country-specific 15D algorithms compare or about how the different valuation tasks contribute to the final algorithm values. Before being able to meaningfully interpret differences observed in 15D algorithm values, a better understanding of how the 15D valuation procedure shapes these values is necessary.

The aim of this study was to compare the results from 15D valuation studies conducted in Norway and Finland. Specifically, we compare the scores assigned in the valuation tasks, the value algorithms derived using the original valuation procedure, and discuss the contributions of each task and the algorithm estimation procedure to the observed differences in algorithm values.

Methods

15D Descriptive System

The 15D descriptive system consists of 15 dimensions, covering physical, mental, and social aspects of health [19]. Each dimension has five response options, with the first level corresponding to full functionality and the remaining levels describing declining levels of functionality.

15D Valuation System

The 15D allows the description of $5^{15} \approx 3.1 \times 10^{10}$ health states. 15D values are calculated using a predefined *value algorithm*, which is generated to reflect the preferences of the target population. The generation of a 15D value algorithm consists of two elements: the valuation tasks and the *value algorithm estimation procedure*. Because of the large number of dimensions and levels, the Finnish 15D value algorithm was derived using assumptions from the multi-attribute utility theory [20].

The Norwegian valuation study was based on the three valuation tasks developed by Sintonen [4]: 1) In the top task, respondents are asked to compare the top levels for all 15 dimensions, using a VAS anchored in "most important" (= 100) and "least important" (= 0, top task; see Appendix 1 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2017.09.018); 2) In the bottom task, the bottom levels for all 15 dimensions, regarding the lowest levels of functioning, are rated on a VAS ranging from "best imaginable health state" (= 100) to "worst imaginable health state" (= 0, bottom task); 3) In the within-dimension task, the respondents are asked to place the five levels of one dimension, plus the state of "being dead," on a VAS anchored in "best imaginable health state" and "worst imaginable health state" (see Appendix 2 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2017.09.018). We refer to the average scores derived from these tasks as top task scores, bottom task scores, and within-dimension scores, respectively. For brevity, L1, L2, L3, L4, and L5 are used to refer to levels 1 through 5 in the within-dimension task.

15D Algorithm Estimation Procedure

We use the term *algorithm estimation procedure* to refer to the steps taken to estimate an algorithm on the basis of scores from the valuation tasks averaged across all included respondents. Unless otherwise explicitly stated, all scores mentioned in this article are such averages. Briefly, the algorithm estimation procedure assigns each of the 15 dimensions a slot of the scale between "full health" (1) and "not being alive" (0), which reflects its relative importance. The levels of each dimension are assigned values within the respective slot. The dimensions are additive; summed up, they represent the full range of the 15D value algorithm. The valuation tasks were designed to provide input to the following value function described by Sintonen [4]:

$$V_H = \sum_j I_j(x_j) w_j(x_j),$$

where V_H is the social value of health state H and I_j (x_j) is a set of positive constants for the jth dimension, representing the relative importance of the dimension at its various levels, constrained such that $\sum_j I_j = 1$ for any level. $w_j(x_j)$ is a numerical function of the jth dimension, representing the relative value of various levels of the dimension, such that the top level = 1 and being dead = 0.

The function was inspired by the multi-attribute utility theory [20] and builds on the idea of a two-stage valuation process in which levels within dimensions are valued in one task and the relative importance of the dimensions is determined separately. Nevertheless, the function developed by Sintonen assumes that the importance assigned to dimensions could vary by level.

We applied the algorithm estimation procedure developed by Sintonen [4]. An overview is presented here, and a more in-depth numerical example is given in Appendix 3 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2017.09.018. The first step generates importance weights for each dimension on the basis of the top task. Averages for each dimension are calculated across respondents and are divided by the sum of all 15 such averages. The result is a set of 15 values (one for each dimension) that sum up to 1. The same procedure is used to generate importance weights from the bottom task (see Table 1 in Appendix 3 in Supplemental Materials for Norwegian top and bottom task scores and importance weights). The following steps are taken for each dimension separately: 1) Within-dimension scores are rescaled such that L1 is anchored in 1 and "being dead" is anchored in 0. These values are reserved for later. 2) L1 to L5 are rescaled, now such that L1 equals the top importance weight for the corresponding dimension and L5 equals the bottom importance weight for the corresponding dimension. 3) The results of step 1 are multiplied with the corresponding results from step 2. The resulting 15D value algorithm consists of 60 values, each referring to one of the five response options of the 15 dimensions. Algorithm values in this article are presented to indicate disutility; a positive value indicates a value loss associated with health problems and negative values indicate value gains.

Samples

The Finnish valuation study that was conducted in 1992 sampled 2500 members of the Finnish general population and is described in detail elsewhere [4]. The Finnish data collection differed from the Norwegian data collection in that there was no Web survey,