

EUR Research Information Portal

The impact of demographic change on value set validity and obsolescence

Published in:

Quality of Life Research

Publication status and date:

Published: 13/09/2024

DOI (link to publisher):

[10.1007/s11136-024-03770-5](https://doi.org/10.1007/s11136-024-03770-5)

Document Version

Publisher's PDF, also known as Version of record

Document License/Available under:

CC BY

Citation for the published version (APA):

Jonker, M. F. (2024). The impact of demographic change on value set validity and obsolescence. *Quality of Life Research*, 33(11), 3155-3160. <https://doi.org/10.1007/s11136-024-03770-5>

[Link to publication on the EUR Research Information Portal](#)

Terms and Conditions of Use

Except as permitted by the applicable copyright law, you may not reproduce or make this material available to any third party without the prior written permission from the copyright holder(s). Copyright law allows the following uses of this material without prior permission:

- you may download, save and print a copy of this material for your personal use only;
- you may share the EUR portal link to this material.

In case the material is published with an open access license (e.g. a Creative Commons (CC) license), other uses may be allowed. Please check the terms and conditions of the specific license.

Take-down policy

If you believe that this material infringes your copyright and/or any other intellectual property rights, you may request its removal by contacting us at the following email address: openaccess.library@eur.nl. Please provide us with all the relevant information, including the reasons why you believe any of your rights have been infringed. In case of a legitimate complaint, we will make the material inaccessible and/or remove it from the website.



The impact of demographic change on value set validity and obsolescence

Marcel F. Jonker^{1,2}

Accepted: 20 August 2024
© The Author(s) 2024

Abstract

Purpose To investigate the contribution of demographic trends in countries' age and gender composition to value set validity and obsolescence.

Methods Time-trade off (TTO) valuation data from 3 EQ-5D-3L value sets of 20 years or older from the United Kingdom, Japan, and the United States were re-analyzed using Bayesian heteroskedastic Tobit models with sex and age group-specific scale parameters. Original value sets were obtained by weighting the original preference structures with the countries' original demographic composition at the time of the data collection. Updated value sets were created using the original preference structure weighted using the countries' most recent demographic composition. The differences between the original and updated value sets were monitored and compared based on 95% credible intervals.

Results The gender and age composition of the investigated countries changed in all 3 countries over time. The modelled health state preferences also depended on the respondents' gender and age. However, the overall impact of this demographic change on the investigated value sets was negligible in all 3 countries and this finding was robust to accounting for the impact of ethnicity trends in the United States.

Conclusion Value sets may become redundant and obsolete for various reasons, but demographic change was not identified as a contributing factor.

Keywords Value set validity · Value set redundancy · Value set obsolescence · Demographic trends · EQ-5D

Introduction

The use of preference-based value sets in health technology assessment (HTA) and other applications is widespread, and the validity of reported results depends on the value sets reflecting contemporary societal preferences [1]. On the one hand, decision-makers often prioritize consistency in value sets, at times prescribing specific sets for analyses to maintain consistency in between-study comparisons [1, 2]. Consequently, they may hesitate to abandon older value sets. On the other hand, several factors can render value sets increasingly outdated and eventually obsolete over time. These

factors include the development of improved methods for eliciting and modelling preferences, changes in populations' preferences, and shifts in the demographic composition of populations over time.

To date, there are no universally accepted expiration dates or guidelines for reassessing the validity of value sets. In fact, the concept of value set invalidity and subsequent obsolescence has received, with some notable exceptions, very little attention in the literature [3, 4]. At the same time, value sets are becoming older and are seldom updated. Hence there is a growing need to investigate the determinants of value set obsolescence. This paper focuses on a specific aspect of value set obsolescence, which is the impact of demographic shifts in the demographic composition of populations. More specifically, it aims to determine whether changes in the age/sex composition of populations, all other things being equal (*ceteris paribus*), can implicitly generate an expiration date for preference-based value sets.

✉ Marcel F. Jonker
marcel@mfjonker.com

¹ Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

² Erasmus Choice Modelling Centre, Erasmus University Rotterdam, Rotterdam, The Netherlands

Methods

To evaluate the impact of demographic change on preference-based values sets, 3 of the oldest EQ-5D value sets were selected for re-analysis: those of the United Kingdom (UK) [5, 6], Japan (JPN) [7], and the United States (US) [8]. Permission to re-analyze the data was obtained from the corresponding authors and the time trade-off (TTO) valuation data were subsequently obtained from public data repositories and from the corresponding author of the Japanese value set. The obtained data comprised, in addition to the TTO-derived health-state values, also the participating respondents' sex and age, and in the case of the USA, the respondents' race/ethnicity, which was used in a sensitivity analysis that is reported in the online supplemental materials. The respondents' age was recoded into 4 groups, i.e. 18–34, 35–54, 55–74, and 75+.

To allow for an evaluation of the impact of changes in the age and sex composition of the populations of these 3 countries, demographic data were obtained from the World Bank [9]. Both the demographic composition in the year the TTO data were collected and the most recent demographic composition from the year 2022 was obtained. This allowed the original value sets to be obtained by weighting the estimated preference structures with the countries' original demographic composition at the time of the data collection, and updated value sets by weighting the original preference structure with the countries' most recent demographic composition.

This was performed by re-analyzing the TTO valuation using state-of-the-art methods [10]. In the Japanese dataset, each respondent completed the same set of 17 TTO tasks. In the UK and US datasets, each respondent completed 12 TTO tasks that were included in the analysis, which were sampled from an overall design that comprised 45 EQ-5D-3 L health states. First, respondents with fewer than 10 observations were excluded from the analyses. Second, respondents with a positively sloped relationship between their TTO valuations and the misery index of the health states were excluded from the analyses. Third, the data of the remaining respondents was analyzed using Bayesian heteroskedastic Tobit models with censoring at -1 using a model specification that included demographic-specific scale parameters. More specifically, the observed TTO values for respondent i in task t were censored as:

$$TTO_{it} = \begin{cases} TTO_{it}^* & \text{if } TTO_{it}^* > -1 \\ -1 & \text{if } TTO_{it}^* \leq -1 \end{cases} \quad (1)$$

with the latent TTO_{it}^* values assumed to be normally distributed:

$$TTO_{it}^* \sim Normal(\mu_{it}, \sigma_{it}). \quad (2)$$

Here the mean (μ_{it}) and standard deviation (σ_{it}) reflect the average health state value and variation among respondents in their valuation of the health state presented in task t of respondent i , respectively. Similar to Pickard et al. [11], the standard deviation (σ_{it}) was modelled as a 4th-order polynomial of the health state values:

$$\sigma_{it} = \exp(\gamma_0 + \gamma_1 \mu_{it} + \gamma_2 \mu_{it}^2 + \gamma_3 \mu_{it}^3 + \gamma_4 \mu_{it}^4) \quad (3)$$

, which ensured that the variances of the predicted values could flexibly depend on the health state severity.

The mean (μ_{it}) was specified as follows:

$$\begin{aligned} \mu_{it} = & \beta_0 + (\beta_1 MO2_{it} + \beta_2 MO3_{it}) * \varphi_{MO_{sa}} \\ & + (\beta_3 SC2_{it} + \beta_4 SC3_{it}) * \varphi_{SC_{sa}} \\ & + (\beta_5 UA2_{it} + \beta_6 UA3_{it}) * \varphi_{UA_{sa}} \\ & + \beta_7 PD2_{it} + \beta_8 PD3_{it}) * \varphi_{PD_{sa}} \\ & + (\beta_9 AD2_{it} + \beta_{10} AD3_{it}) * \varphi_{AD_{sa}}. \end{aligned} \quad (4)$$

Here β_0 denotes the intercept, β_{1-10} denote the slope coefficients that capture the EQ-5D decrements for levels 2 and 3, and the φ parameters are multiplicative scale parameters that shrink or amplify respondents' slope coefficients based on their sex s and age group a ($a \in \{1-4\}$), which amounts to 8 scale parameters (φ) per EQ-5D dimension. For statistical identification, the scale parameters were constrained to be positive and subject to a mean-of-one constraint.

Bayesian Markov Chain Monte Carlo (MCMC) methods were used to fit the models, which involves the selection of prior distributions for the unknown parameters and updating these via the likelihood of the observed data. Uninformative normal priors with a mean of zero and standard deviation of 10 were assigned to the β and γ parameters and mean-to-one constrained log-normal parameters were assigned to the dimension-specific φ parameters. Similar to Jonker et al. [12], the standard deviation of the log-normal priors was set to 0.4 and the mean of the log-normal priors was defined as $\mu = -\sigma^2/2 = -0.08$ to ensure that the prior distribution had an expectation of 1. The models were implemented in the BUGS language and fitted using OpenBUGS [13]. A custom-implemented Metropolis-within-Gibbs algorithm with antithetic sampling was used to update the β and γ parameters and a custom-implemented slice sampling algorithm was used to update the dimension-specific φ parameters subject to the mean-of-one constraints [12]. All reported estimates were based on 10,000 burn-in iterations to let three chains converge and a total of 30,000

MCMC iterations to reliably approximate the posterior distributions. Note that the model codes are included in the online supplemental and that convergence was evaluated based on a visual inspection of the MCMC chains and the convergence diagnostics as implemented in the OpenBUGS package.

During the Bayesian MCMC estimation, the original value sets were monitored by multiplying the β parameters with the corresponding φ scale parameters and then taking a weighted average using weights determined as the proportion of the population in the respective sex and age categories of the scale parameters at the time of the TTO data collection. Similarly, updated value sets were monitored by taking a weighted average of the original preference structure with weights determined by the countries' most recent demographic composition. Finally, also the difference between the original and updated value sets was monitored during the model estimations. This allowed for a direct assessment of the impact of the change in countries' age and gender compositions on the value sets.

Results

The UK valuation dataset contained 3,395 respondents, 82 respondents who completed fewer than 10 TTO observations and 43 respondents with a positively sloped relationship between their TTO valuations and the misery indices of the health states were excluded, resulting in 3,270 respondents. The Japanese valuation dataset contained 543 respondents, 3 respondents who completed fewer than 10 TTO observations and 4 respondents with positive slopes were excluded, resulting in 536 respondents. And the United States valuation dataset contained 4,048 respondents, all of whom completed 10 or more TTO tasks and 133 respondents with positive slopes were excluded, resulting in 3,915 respondents.

Table 1 presents the demographic composition of the samples and compares them to the nationally representative benchmarks. None of the samples were accurately nationally representative at the time of the original data collection. This underscores the need to obtain regression-weighted value sets, both using the original and updated demographic weights. Also, in all three countries the population compositions have shifted and become significantly older in the last 2 decades; most prominently in Japan but also in the other 2 countries.

The Tobit regression coefficients are presented in the online supplemental. Briefly summarized, in all 3 countries all β parameters and 2 or more γ parameters are statistically different from zero - in the sense that the

Table 1 Survey respondents and national representative benchmarks, by country, sex, age group, and year

Sex	Age	Country															
		United Kingdom				Japan				United States							
		Dataset		National population		Dataset		National population		Dataset		National population					
Male	18-34	(N=3,270)	13.9%	15.6%	13.8%	-1.8%	(N=536)	9.7%	15.2%	15.5%	15.9%	-5.3%	(N=3,915)	16.1%	16.2%	14.9%	-1.3%
Male	35-54	13.9%	17.5%	16.0%	-1.5%	15.5%	17.2%	15.5%	13.3%	14.8%	14.8%	-1.3%	16.5%	19.4%	16.3%	16.3%	-3.1%
Male	55-74	12.5%	11.5%	14.0%	2.5%	15.5%	13.3%	15.5%	2.9%	7.6%	14.8%	1.5%	7.5%	10.1%	13.9%	13.9%	3.8%
Male	75+	3.1%	3.3%	5.1%	1.8%	1.3%	2.9%	1.3%	2.9%	7.6%	7.6%	4.7%	2.0%	2.9%	3.9%	3.9%	1.0%
Female	18-34	18.1%	15.7%	13.3%	-2.4%	12.7%	14.5%	12.7%	14.5%	9.5%	9.5%	-5.0%	22.3%	15.7%	14.5%	14.5%	-1.2%
Female	35-54	16.9%	17.7%	16.4%	-1.3%	25.6%	16.9%	25.6%	16.9%	15.4%	15.4%	-1.5%	22.2%	19.9%	16.2%	16.2%	-3.7%
Female	55-74	16.0%	12.7%	14.8%	2.1%	17.7%	14.7%	17.7%	14.7%	15.5%	15.5%	0.8%	10.0%	11.0%	15.1%	15.1%	4.1%
Female	75+	5.6%	6.0%	6.6%	0.6%	2.1%	5.3%	2.1%	5.3%	11.4%	11.4%	6.1%	3.3%	4.8%	5.2%	5.2%	0.4%
Total	18+	100%	100%	100%	0%	100%	100%	100%	100%	100%	100%	0%	100%	100%	100%	100%	100%

95% credible intervals do not comprise 0. Moreover, in all 3 countries there are sex and age group-specific scale parameters (φ) that are different from 1 – in the sense that the 95% credible intervals do not comprise 1. In the United Kingdom, all five EQ domains have scale parameters that are significantly different from 1. In the United States, all domains except for mobility are significantly different from 1, and in Japan, the pain and discomfort and anxiety and depression domains have scale parameters that are significantly different from 1.

The overall impact of the scale parameters and shifts in demographic compositions (i.e., differential preferences and associated weights) is summarized in Table 2, which provides the original and updated value set decrements as well as their differences. As shown, the value sets themselves are country-specific but the differences between the original and updated value set decrements are close to identical. Most differences are 0.00 or smaller and the maximum difference is 0.01. Most importantly, all differences have 95% credible intervals that comprise 0, implying that the original and updated value sets are not significantly different.

Discussion

There are multiple reasons why value sets can become redundant and obsolete over time. For example, valuation methods have significantly improved in the past 2 decades, HTA bodies may require different value sets or methods to be used, and most importantly, population preferences can and likely have shifted over time. In this paper one aspect of value set obsolescence has been investigated, which is the extent to which demographic changes in the age/sex composition of populations have, keeping all other things constant, implicitly created an expiration date on preference-based value sets. Based on the presented results, we can conclude that demographic change is not an important determinant of value set validity.

Strengths of this study are that the results were obtained using the original TTO valuation data from 3 different countries from 3 different continents, a state-of-the-art Bayesian modelling approach, and a conceptually simple yet effective method to obtain both original and updated value sets. A potential weakness is that only age and gender were considered in the presented analyses; for example, the impact of ethnicity was ignored even though it could particularly

Table 2 EQ-5D health state decrements, original and age/sex corrected, by country

	United Kingdom			Japan			United States		
	1997	2022	Δ	1998	2022	Δ	2002	2022	Δ
MO2	-0.06 (-0.07 — -0.06)	-0.06 (-0.07 — -0.06)	0.00 (0.00 — 0.00)	-0.10 (-0.11 — -0.08)	-0.10 (-0.11 — -0.08)	0.00 (0.00 — 0.00)	-0.04 (-0.04 — -0.03)	-0.04 (-0.04 — -0.03)	0.00 (0.00 — 0.00)
MO3	-0.34 (-0.36 — -0.33)	-0.34 (-0.36 — -0.33)	0.00 (0.00 — 0.00)	-0.43 (-0.46 — -0.4)	-0.43 (-0.46 — -0.4)	0.00 (-0.01 — -0.01)	-0.34 (-0.35 — -0.32)	-0.34 (-0.35 — -0.32)	0.00 (0.00 — 0.00)
SC2	-0.11 (-0.12 — -0.11)	-0.12 (-0.12 — -0.11)	0.00 (0.00 — 0.01)	-0.05 (-0.07 — -0.04)	-0.06 (-0.07 — -0.04)	0.00 (0.00 — 0.00)	-0.07 (-0.08 — -0.07)	-0.07 (-0.08 — -0.07)	0.00 (0.00 — 0.00)
SC3	-0.23 (-0.25 — -0.22)	-0.24 (-0.26 — -0.23)	0.01 (0.01 — 0.01)	-0.09 (-0.11 — -0.07)	-0.09 (-0.12 — -0.07)	0.00 (0.00 — 0.01)	-0.24 (-0.25 — -0.23)	-0.24 (-0.26 — -0.23)	0.00 (0.00 — 0.00)
UA2	-0.08 (-0.09 — -0.07)	-0.08 (-0.09 — -0.07)	0.00 (0.00 — 0.00)	-0.03 (-0.05 — -0.02)	-0.03 (-0.05 — -0.02)	0.00 (0.00 — 0.00)	-0.04 (-0.05 — -0.04)	-0.04 (-0.05 — -0.04)	0.00 (0.00 — 0.00)
UA3	-0.22 (-0.23 — -0.21)	-0.22 (-0.23 — -0.21)	0.00 (0.00 — 0.00)	-0.13 (-0.15 — -0.11)	-0.12 (-0.15 — -0.11)	0.00 (-0.01 — -0.01)	-0.17 (-0.18 — -0.16)	-0.18 (-0.19 — -0.16)	0.01 (0.00 — 0.01)
PD2	-0.09 (-0.09 — -0.08)	-0.09 (-0.09 — -0.08)	0.00 (0.00 — 0.00)	-0.06 (-0.07 — -0.05)	-0.06 (-0.07 — -0.05)	0.00 (0.00 — 0.00)	-0.05 (-0.06 — -0.05)	-0.05 (-0.06 — -0.05)	0.00 (0.00 — 0.00)
PD3	-0.47 (-0.48 — -0.45)	-0.47 (-0.48 — -0.45)	0.00 (0.00 — 0.00)	-0.18 (-0.2 — -0.17)	-0.17 (-0.19 — -0.16)	-0.01 (-0.01 — -0.01)	-0.36 (-0.37 — -0.35)	-0.36 (-0.37 — -0.35)	0.00 (0.00 — 0.00)
AD2	-0.12 (-0.13 — -0.11)	-0.12 (-0.13 — -0.11)	0.00 (0.00 — 0.00)	-0.06 (-0.08 — -0.05)	-0.06 (-0.08 — -0.05)	0.00 (0.00 — 0.00)	-0.08 (-0.09 — -0.07)	-0.08 (-0.09 — -0.07)	0.00 (0.00 — 0.00)
AD3	-0.37 (-0.38 — -0.35)	-0.37 (-0.38 — -0.36)	0.00 (0.00 — 0.00)	-0.12 (-0.14 — -0.11)	-0.12 (-0.14 — -0.11)	0.00 (-0.01 — -0.01)	-0.26 (-0.27 — -0.25)	-0.27 (-0.28 — -0.26)	0.00 (0.00 — 0.01)

for the United States have been an important determinant of value set obsolescence. Fortunately, according to the sensitivity analysis presented in the online supplemental, changes in the ethnic composition of the US population over the last two decades did not affect the validity of the US value set. Another potential weakness is that historic demographic trends may be smaller than future changes in populations' age and sex compositions over time. However, the included countries experienced relatively large demographic shifts in the sex/age composition of their populations [9] and the demographic shifts as experienced over the past 20–25 years are equally large or larger than the predicted shifts in the age/sex distribution in the coming 20–25 years [14]. Accordingly, the preference differentials need to be at least an order of magnitude larger than those presented in this paper before demographic change would be able to significantly contribute to value set obsolescence. In other words, the presented results are robust, although contingent upon the assumption that the original preference structure as observed in the TTO datasets has remained constant over time. Future research is therefore necessary to establish the extent to which health-state preferences have indeed remained constant over time, as well as to determine appropriate criteria and threshold values for value set validity and obsolescence.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11136-024-03770-5>.

Acknowledgements The author would like to thank Claire Gudex, Aki Tsuchiya, and James Shaw for their permission to re-analyze the TTO data and gratefully acknowledges financial support from the EuroQol Research Foundation. The views expressed in this article do not necessarily reflect those of the EuroQol Group.

Funding This work was supported by a grant (EQ Project 1612-RA) from the EuroQol Research Foundation.

Data availability The UK valuation data are available at <https://doi.org/10.5255/UKDA-SN-3444-1>. The US valuation data are available at <https://archive.ahrq.gov/professionals/clinicians-providers/resources/rice/EQ5Dproj.html>. The Japanese valuation data are not publicly available. The demographic data can be obtained from <https://databank.worldbank.org/source/health-nutrition-and-population-statistics>.

Declarations

Ethics statement Permission to re-analyze the legacy valuation data was granted by the corresponding authors of the original publications. The UK and USA data are publicly available, while the Japanese data were obtained in anonymized form from the corresponding author.

Role of the funder/sponsor The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript;

and decision to submit the manuscript for publication.

Conflict of interest Dr. Jonker reported being a member of the Euro-Qol Group. No other disclosures were reported.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Kennedy-Martin, M., Slaap, B., Herdman, M., van Reenen, M., Kennedy-Martin, T., Greiner, W., Busschbach, J., & Boye, K. S. (2020). Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. *The European Journal of Health Economics*, *21*, 1245–1257.
- National Institute for Health and Care Excellence (2022). NICE health technology evaluations: the manual. Process and methods [PMG36].
- Pickard, A. S. (2015). Is it time to update societal value sets for preference-based measures of health? *Pharmacoeconomics*, *33*(3), 191–192.
- Law, E. H., Pickard, A. S., Walton, S. M., Xie, F., Lee, T. A., & Schwartz, A. (2022). Time-Specific Differences in Stated Preferences for Health in the United States. *Medical Care*, *60*(6), 462–469.
- Gudex, C., Dolan, P., Williams, A. H., & Kind, P. (1993). Health State Valuations from the British General Public [data collection]. UK Data Service. SN: 3444. <https://doi.org/10.5255/UKDA-SN-3444-1>
- Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical care*. 1997 Nov 1:1095–108.
- Tsuchiya, A., Ikeda, S., Ikegami, N., Nishimura, S., Sakai, I., Fukuda, T., Hamashima, C., Hisashige, A., & Tamura, M. (2002). Estimating an EQ-5D population value set: The case of Japan. *Health Economics*, *11*(4), 341–353.
- Shaw, J. W., Johnson, J. A., & Coons, S. J. (2005 Mar). US valuation of the EQ-5D health states: Development and testing of the D1 valuation model. *Medical care*, *1*, 203–220.
- World Bank Group Health nutrition and population statistics. <https://databank.worldbank.org/source/health-nutrition-and-population-statistics> Accessed 10/19/2023.
- Rowen, D., Mukuria, C., & McDool, E. (2022). A systematic review of the methodologies and modelling approaches used to generate international EQ-5D-5L value sets. *Pharmacoeconomics*, *40*(9), 863–882.
- Pickard, A. S., Law, E. H., Jiang, R., Pullenayegum, E., Shaw, J. W., Xie, F., Oppe, M., Boye, K. S., Chapman, R. H., Gong, C. L., & Balch, A. (2019). United States valuation of EQ-5D-5L health states using an international protocol. *Value in Health*, *22*(8), 931–941.

12. Jonker, M. F., Donkers, B., de Bekker-Grob, E., & Stolk, E. A. (2019). Attribute level overlap (and color coding) can reduce task complexity, improve choice consistency, and decrease the drop-out rate in discrete choice experiments. *Health Economics*, 28(3), 350–363.
13. Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to Bayesian analysis*. CRC press; Oct 2.
14. World Bank Group Population estimates and predictions. <https://databank.worldbank.org/source/population-estimates-and-projections>. Accessed 11/06/2023.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.