Chapter 1 The Development of the EQ-5D-5L and its Value Sets



Nancy Devlin, Simon Pickard, and Jan Busschbach

Abstract This chapter introduces the EQ-5D-5L questionnaire and its development by the EuroQol Group. The availability of the EQ-5D-5L, and the growing evidence of its pivotal role as a measurement system, generated a demand for 'values' to accompany it that would enable the use of EQ-5D-5L data in the estimation of quality-adjusted life-years (QALYs) and other applications where EQ-5D-5L profile data needs to be summarised by a single number. Chapter 1 sets out the main aim of the book: to provide an accessible source of information and guidance to support users of EQ-5D-5L and its value sets. Specifically, the book aims to improve users' understanding of how EQ-5D-5L value sets are generated using the internationally standardised EQ-VT protocol; to raise awareness of the characteristics and properties of value sets; and to inform users' choice of which value set to select for which purpose, and how that choice may affect analysis. The chapter concludes with an overview of the content of the book.

1.1 The EQ-5D as an Instrument for Measuring and Valuing Health

Since the 1990s, the EQ-5D instrument has held a pivotal role in the measurement of self-reported health status and health-related quality of life (HRQoL) (Devlin and Brooks 2017). The availability of a *concise* generic instrument for measuring

S. Pickard

J. Busschbach Section of Medical Psychology, Department of Psychiatry, Erasmus MC, Rotterdam, The Netherlands

N. Devlin (⊠)

Centre for Health Policy, University of Melbourne, Melbourne, VIC, Australia e-mail: nancy.devlin@unimelb.edu.au

Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois at Chicago, Chicago, IL, USA

patients' and population self-reported health¹ meant that it could be included, with minimal responder burden, in clinical trials, observational studies, and population health surveys. More recently, it has become the cornerstone of routine outcomes measurement in health care systems such as the English NHS PROMs programme and Sweden's national quality registers. The ability of the EQ-5D to measure HRQoL in a *generic* manner has the important advantage of yielding data that can readily be compared across disease areas and between patient and population subgroups, and against population norms. This broad comparability of EQ-5D data is particularly crucial in providing evidence that quantifies health benefits in a standardised and transparent manner to inform decisions regarding alternative ways of using health care resources.

The EQ-5D was developed by the EuroQol Group, then a small group of academics which has now grown into an international network of multidisciplinary researchers with more than 100 members worldwide (Devlin and Brooks 2017). The development of the EQ-5D was motivated in part by the specific goal of providing evidence on the outcomes of health care programmes in a manner that would facilitate economic evaluation. One of the considerations underpinning the development of the instrument was that it would be accompanied by the 'values' (sometimes also referred to as 'utilities', 'quality of life weights', the 'EQ-5D Index' or 'EQ Index weights') that would enable the quality adjustment of life years as required for the estimation of quality-adjusted life-years (QALYs) used in cost effectiveness analysis (Drummond et al. 2015). The availability of value sets for this purpose has been a notable part of the success and uptake of EQ-5D instruments.

The value sets that accompany EQ-5D instruments provide a means of summarising, via a single number, how good or bad health status is as described by the EQ-5D. The responses to the EQ-5D instrument – that is, the particular combination of levels which are indicated on each of the five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), by those completing it – can be described as EQ-5D 'profiles' (see Box 1.1). The original version of the EQ-5D, the EQ-5D-3L, has three response levels for each of the five dimensions, describing a total of $3^5 = 243$ possible profiles (Brooks 1996). The focus of this book is on the later five level version, the EQ-5D-5L, development of which is described in more detail in the following section – which describes a total of $5^5 = 3125$ profiles (Herdman et al. 2011).

The value sets for these instruments provide a single value for each of the possible profiles described by them. These values lie on a scale anchored at 1 (full health) and 0 (dead), as is required for the estimation of QALYs. The values are built up from a set of sub-weights which represent the relative importance of each level of problem in each dimension, and indicate how good or bad these are overall, when combined in EQ-5D-5L profiles. The term *value set* refers to a set of values

¹There is ongoing debate about the term health-related quality of life (HRQoL) and whether EQ-5D measures HRQoL or self-perceived health status (Brazier and Karimi 2016). For simplicity, in the remainder of this book, we refer to the EQ-5D-5L value sets as providing values for health states or health as described by the EQ-5D-5L.

Box 1.1: EQ-5D Questionnaires, EQ-5D Profiles and Values EO-5D questionnaires comprise two key parts:

- (i) the EQ-5D descriptive system, as shown below for the EQ-5D-5L. Respondents are asked to indicate the level of problem they experience on each of the five dimensions today. The combination of these ticks describes that person's EQ-5D self-reported health state, referred to as an 'EQ-5D profile'
- (ii) the EQ VAS, a vertical visual analogue scale capturing respondents' overall assessment of their health on a scale from 0 (worst possible health you can imagine) to 100 (best possible health you can imagine) (not shown here).

The EQ-5D-5L questionnaire



© EuroQol Research Foundation. EQ-5DTM is a trademark of the EuroQol Research Foundation. Reproduced by permission of EuroQol Research Foundation. Reproduction of this version is not allowed. For reproduction, use or modification of the EQ-5D (any version), please register your study by using the online EQ registration page: www.euroqol.org.

The EQ VAS is an important part of the questionnaire and provides the patients' overall assessment of their own health on a visual analogue scale. However, many applications of EQ-5D data, including the estimation of QALYs for economic evaluation, focus instead on the use of EQ-5D profile

Box 1.1 (continued)

data. The profile data, and the use of value sets to summarise those data, is the focus of this book, and therefore the EQ VAS is not discussed further. There are in fact many ways of analysing EQ-5D profile data, as detailed in Devlin et al. (2020). One of these ways is by weighting the profile using values sets. This is the most common way of using EQ-5D data in cost effective-ness analysis. This book focuses on the value sets available for the EQ-5D-5L and their use in weighting EQ-5D-5L profile data.

The value sets provide a way of converting the profiles into a single number that reflects how good or bad people think they are. The values are usually obtained using stated preference methods, and yield values that lie on a scale anchored by the value of 1 for full health, and 0 for dead. EQ-5D values cannot be higher than 1, but values <0 are possible, and indicate health states considered on average to be worse than dead (WTD). Value sets are generally intended to represent the average preferences of local/national populations – so EQ-5D value sets differ between countries. See Chap. 4 for a summary of the available value sets for EQ-5D-5L, and Chap. 6 for information about the differences and similarities between them.

for all possible profiles defined by a particular EQ-5D instrument, and is occasionally also referred to by other names, such as an EQ-5D 'tariff' or 'social values.' For the purposes of this book, we will use the terms value and value set.

These values are usually based on the average preferences of the relevant adult general population, obtained using stated preference methods such as the Time Trade-Off (TTO). These stated preference methods aim to elicit values which have the desired properties for estimating QALYs (see Box 5.1 in Chap. 5). Indeed, the availability of EQ-5D values which are suitable for this purpose has led to the EQ-5D being the most widely recommended questionnaire for use in the cost effectiveness evidence submitted to Health Technology Appraisal (HTA) bodies. The EQ-5D is recommended in 85% of HTA guidelines (Kennedy-Martin et al. 2020), including those of the UK's National Institute for Health and Care Excellence (NICE 2013).

The EuroQol Group was and continues to be a pioneer in the development of local/national value sets. The development of EQ-5D-3L value sets was, as an international research effort, unparalleled in the availability of country-specific values (Szende et al. 2007). There are currently EQ-5D-3L value sets available for 35 countries and, for the EQ-5D-5L, 25 countries, with still further value set studies underway or planned. Both the EQ-5D-3L and the EQ-5D-5L, and the value sets which accompany them, continue to be used next to each other in many countries. The value sets facilitated the use of data from EQ-5D instruments in the estimation of QALYs based on local preferences, as well as in other, 'non-economic' applications where EQ-5D profile data are summarised in a way that reflects the relative importance of the different dimensions.

1.2 The Development of the EQ-5D-5L

In 2005 the EuroQol Group initiated efforts to develop an expanded-level version of the EQ-5D-3L. This was motivated by concerns by some stakeholders about limitations of the original instrument, particularly ceiling effects and changes in health that were too small to be detected by the three-level version. Studies which had been undertaken by EuroQol Group members prior to 2005 had shown that various experimental five-level versions of EQ-5D could reduce ceiling effects while at the same increasing reliability and sensitivity (discriminatory power) and maintaining feasibility (Janssen et al. 2008a, b; Pickard et al. 2007a, b).

The development and testing of the EQ-5D-5L is reported in Herdman et al. (2011). A decision was made early in the new instrument's development to retain the same five dimensions as the EQ-5D-3L, but to expand the number of response levels. This could in principle have been achieved simply by adding two 'unlabelled' intermediate levels between the existing three. However, in order to arrive at values for the EQ-5D-5L profiles, each health state to be evaluated by respondents needed to be capable of being described by five sentences. This in turn required a label for each level. An example of an EQ-5D-5L health state, displayed in the manner it might be presented in a stated preference task, is shown in Fig. 1.1. The state described in Fig. 1.1 is the same combination of levels and dimensions as the example in Box 1.1 i.e., it is EQ-5D-5L profile 21325.

Herdman et al. (2011) describe the process by which these labels were established, using both English and Spanish as root languages in order to support further translation and adaptation of the new instrument. Severity labels for 5 levels in each dimension were identified using response scaling. Selecting labels at approximately the 25th, 50th, and 75th centiles produced two alternative 5-level versions. Focus groups were used to investigate the face and content validity of the two versions, including hypothetical health states generated from those versions. This showed evidence in favour of the wording 'slight-moderate-severe' problems, with level one described as 'no problems' in each dimension, and level five being 'unable to' in the EQ-5D functional dimensions (mobility, self-care, usual activities) and 'extreme problems' in the pain/discomfort and anxiety/depression dimensions.

- Slight problems in walking about
- No problems washing or dressing myself
- Moderate problems doing my usual activities
- Slight pain or discomfort
- Extremely anxious or depressed

Fig. 1.1 An example of an EQ-5D-5L health 'state' described by five sentences

The final version of the five-level instrument which emerged from this work is described in the EQ-5D-5L User Guide (EuroQol Group 2019). Beside the increased number of levels of the dimensions, the 5-level version of the EQ-5D has other notable features which represent improvements on the EQ-5D-3L. Most importantly, the wording of the mobility dimension is improved: the most severe level of the mobility dimension of the EQ-5D-3L is 'confined to bed', which means that it cannot capture severe problems with mobility that do not involve being confined to bed. This acts to limit its usefulness both in detecting problems with mobility and in capturing improvements in mobility resulting from treatment (Oppe et al. 2011). In the EQ-5D-5L, the most severe level of mobility has been changed to 'unable to walk about'.

These improvements have yielded a number of advantages for the EQ-5D-5L over the EQ-5D-3L. These are summarised by Devlin et al. (2018) and include:

- (a) A reduction in the ceiling effect: Using the EQ-5D-5L, compared to the EQ-5D-3L, fewer respondents report no problems on any dimension (e.g., see Feng et al. 2015 and Craig et al. 2014).
- (b) Reduced clustering on just a few states: The lack of granularity in the EQ-5D-3L descriptive system imposes constraints on the self-report of health. Observations tend to cluster on a few health states (Devlin et al. 2020). The EQ-5D-5L consistently produces considerably more unique health states than the EQ-5D-3L, as shown by Buchholz et al. (2018).
- (c) Improved ability to discriminate between patient groups/subgroups: The EQ-5D-5L has better discriminative ability, as demonstrated by improved ability to detect differences between subgroups defined by severity at a given sample size (Janssen et al. 2018). EQ-5D-5L users thus benefit from lower sample size requirements within samples of patients (Pickard et al. 2007b). The EQ-5D-5L has improved ability to measure health accurately at the top of the scale and therefore captures finer differences between mild states of ill health and full health at the top of the scale, whereas the EQ-5D-3L has much larger steps between levels 2 and 1.
- (d) Improvements in the EQ-5D-5L with respect to problems with mobility: As noted above, changing the EQ-5D-3L level 3 descriptor 'confined to bed' constitutes an important improvement in the EQ-5D-5L. Level 3 problems on mobility are rarely observed in EQ-5D-3L data. For example, among patients about to receive hip replacement surgery in the English National Health Service, *none* reported a level 3 problem (Devlin et al. 2010). In effect, in most settings, the EQ-5D-3L only has two levels on mobility: no and some problems. Consequently, the EQ-5D-3L will underestimate benefits of treatments that improve severe problems with mobility (Oppe et al. 2011).

Overall, the evidence suggests that the EQ-5D-5L retains the principal benefits of EQ-5D-3L—its brevity and validity in a wide range of conditions—and produces a more accurate measurement of patient health than the EQ-5D-3L (Devlin et al. 2018). These advantages have been recognised by users and use of the EQ-5D-5L has rapidly increased. There are now more than 130 language versions of the EQ-5D-5L available.

1.3 The need for EQ-5D-5L Values

The availability of the EQ-5D-5L, and the supporting evidence of its improved measurement system, generated a demand for values to accompany it, to allow use of its data in the estimation of QALYs and any other applications where EQ-5D-5L profile data need to be summarised by a preference-weighted single number.

In anticipation of the need to provide EQ-5D-5L value sets, the EuroQol Group initiated an ambitious programme of methodological research, running in parallel with the development of the EQ-5D-5L instrument, and aimed at producing an internationally standardised state-of-the-art valuation protocol. This was timely, as most of the EQ-5D-3L value sets were based on the so called 'MVH-protocol' developed in the early 1990s (Dolan 1997). There was a lack of consistency in the design and implementation of that protocol between value sets studies. Furthermore, limitations of the MVH protocol had been recognised, suggesting improved methods were required for valuation of the EQ-5D-5L.

The aim was therefore not just to improve on the instrument, but to also ensure that the valuation of EQ-5D-5L profiles would be based on the best possible stated preference methods – and to provide a well-described, standard valuation study protocol which could be fielded in a consistent way across different countries. This would ensure that the value sets generated for the new instrument would, as far as possible, be comparable across countries. That is, that differences between the EQ-5D-5L value sets which are observed would reflect the local variations in preferences and opinions which they are intended to capture, rather than being confounded by differences in methods.

As it was anticipated that value sets would take several years to be developed and disseminated, an interim solution was to map EQ-5D-5L data to the EQ-5D-3L instrument by linking descriptive systems, and to use the value sets that already existed for the EQ-5D-3L (van Hout et al. 2011) (further explanation of mapping is provided in Chap. 5). While this provided a practical stop-gap means of summarising EQ-5D-5L data, these mapped values were recognised to be suitable only as a temporary solution as these indirect methods introduce additional error variance, and would still rely upon old and non-standardised value sets. Further, one might question whether values sets for the EQ-5D-3L, developed in the 1990s, would be an adequate representation of the average preferences of today's societies. There are numerous reasons to consider the need to update value sets, including changes in the underlying preferences of populations, improvements in the methods available to value health; changes in the distribution of population demographics; and concerns about potential bias in previous studies (Pickard 2015) – these issues are discussed further in Chap. 7.

In order to arrive at an improved and standardised valuation protocol, the EuroQol Group therefore commissioned a substantial programme of research to develop and test methods suitable for creating new value sets for the EQ-5D-5L that was initiated while the descriptive system was under development. The program of research – which is detailed in Chap. 2 – was started with the intention of providing investigators

around the world with the tools to conduct a valuation study that would follow a standardised protocol and produce high quality data based on validated methods that supported comparisons between countries. These efforts culminated in an international protocol for conducting EQ-5D-5L valuation studies, which has been used to produce the 25 country-specific value sets which are summarised in Chap. 4 of this book (Fig. 1.2).

This endeavour is unique in scale and ambition in the field of HRQoL valuation and represents a significant body of work with direct relevance to decision makers and impact on health care policy internationally.

1.4 The Aims of this Book

The book draws together and summarises, for the first time, the body of evidence on EQ-5D-5L value sets that has been produced internationally from the EuroQol Group's programme of research and protocol development.

The primary aim of the book is to provide an accessible source of information and guidance to support users of EQ-5D-5L and its value sets. Specifically, we aim to improve users' understanding of how value sets are generated; raise awareness of the characteristics and properties of value sets; and inform users' choice of which value set to select for particular application, and how that choice may affect their analysis and conclusions. Moreover, the book will also be useful to health economics and outcomes researchers specialising in HRQoL who want to obtain information on the research practises and protocols developed by the EuroQol Group to support EQ-5D-5L valuation.

We begin in Chap. 2 by detailing the process of developing the research protocol underpinning EQ-5D-5L valuation studies. This included a methodological programme of work and international pilot testing; development of a protocol; the first wave of studies and the conclusions drawn from those early studies; modification and strengthening of the protocol and quality assurance processes; and use of the revised protocol in subsequent waves of value set studies. The chapter indicates the considerable learning and progress that was made through this journey of designing and refining the protocol.

Chapter 3 sets out the various aspects of the study design and the basis on which methodological choices were made with respect to the stated preference methods to use; the sub-set of states to value using these methods; and minimum sample size needed.

Chapter 4 provides a reference source and 'thumbnail overview' of the characteristics of the value set in each of 25 countries. In each case, we provide a summary of the value set itself and its characteristics, a worked example of the calculation of values from it; information on the sample from which values were obtained; the methods used in analysing the data and modelling the value set; and the uptake by local HTA bodies and other health care decision makers.





Chapter 5 provides guidance to those who have collected EQ-5D-5L data and want to know how to choose between the value sets reported in Chap. 4. This includes consideration of the purpose of using value sets, how to proceed when there is no value set for a specific country of interest or where there is more than one value set for a given country; and when it is appropriate to use mapping to obtain EQ-5D-5L values.

Chapter 6 draws together the value sets summarised in Chap. 4, and compares and contrasts their characteristics, reporting original comparative analysis undertaken specifically for this book. To what extent are there similarities between EQ-5D-5L value sets across countries – and are there important differences between them? Our intention in Chap. 6 is to encourage users to be aware of the specific properties of the value sets they select to use.

We conclude in Chap. 7 by reflecting on the value sets produced to date and considering a number of questions about future directions for this body of work. For example, what is the 'shelf-life' of a value set – and what factors should prompt an update, in order to ensure that value sets represent an adequate representation of the average preferences of a society? What methodological questions remain – and how are improvements or variations in methods reconciled with the need for consistency in the evidence presented to HTA bodies and other users?

The book includes a glossary of terms for those unfamiliar with the EQ-5D and the valuation of the EQ-5D-5L.

The stated vision of the EuroQol Group is the aim "to improve decisions about health and health care throughout the world by developing, promoting and supporting the use of instruments with the widest possible applicability for the measurement and valuation of health" (EuroQol Group 2021). We hope this book contributes to that aim, and that it supports your use of EQ-5D-5L to provide evidence for better health care decision making.

References

- Brazier J, Karimi M (2016) Health, health-related quality of life, and quality of life: what is the difference? Pharmacoeconomics 34(7):645–649
- Brooks R, on behalf of the EuroQol Group (1996) EuroQol: the current state of play. Health Policy 37(1):53–72
- Buchholz I, Janssen B, Kohlman T, Feng Y-S (2018) A systematic review on studies comparing the measurement properties of the three-level and the five-level version of the EQ-5D. Pharmacoeconomics 36(6):645–661
- Craig BM, Pickard AS, Lubetkin EI (2014) Health problems are more common, but less severe when measured using newer EQ-5D versions. J Clin Epidemiol 67(1):93–99
- Devlin N, Brooks R (2017) EQ-5D and the EuroQol group: past, present and future. Appl Health Econ Health Policy 15(2):127–137
- Devlin N, Parkin D, Browne J (2010) Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. Health Econ 19(8):886–905
- Devlin N, Brazier J, Pickard AS, Stolk E (2018) 3L, 5L, what the L? A nice conundrum. Pharmacoeconomics 36(6):637–640

- Devlin N, Parkin D, Janssen B (2020) Methods for analysing and reporting EQ-5D data. Springer, Cham
- Dolan P (1997) Modelling valuations of EuroQol health states. Med Care 35(11):1095-1108
- Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G (2015) Methods for the economic evaluation of health care programmes, 4th edn. Oxford University Press, Oxford
- EuroQol Research Foundation (2019) EQ-5D-5L User Guide. Version 3.0. EuroQol Research Foundation: Rotterdam. https://euroqol.org/publications/user-guides. Accessed 25 June 2021
- EuroQol Research Foundation (2021). https://euroqol.org/euroqol/vision-and-mission/. Accessed 23 July 2021
- Feng Y, Devlin N, Herdman M (2015) Assessing the health of the general population in England: how do the three- and five-level versions of EQ-5D compare? Health Qual Life Outcomes 13:171. https://doi.org/10.1186/s12955-015-0356-8
- Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, Bonsel G, Badia X (2011) Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 20(10):1727–1736
- Janssen M, Birnie E, Bonsel G (2008a) Quantification of the level descriptors for the standard EQ-5D three level system and a five level version according to 2 methods. Quality Life Res 17(3):463–473
- Janssen M, Birnie E, Haagsma J, Bonsel G (2008b) Comparing the standard EQ-5D three level system with a five level version. Value Health 11(2):275–284
- Janssen MF, Bonsel G, Luo N (2018) Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. Pharmacoeconomics 36(6):675–697
- Kennedy-Martin M, Slaap B, Herdman M, van Reenen M, Kennedy-Martin T, Greiner W, Busschbach J, Boye KS (2020) Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. Eur J Health Econ 21(8):1245–1257
- National Institute for Health and Care Excellence (2013) Guide to the methods of technology appraisal 2013. https://www.nice.org.uk/process/pmg9/chapter/foreword. Accessed 17 July 2021
- Oppe M, Devlin N, Black N (2011) Comparison of the underlying constructs of the EQ-5D and Oxford Hip Score: implications for mapping. Value Health 14(6):884–891
- Pickard AS (2015) Is it time to update societal value sets for preference-based measures of health? Pharmacoeconomics 33(3):191–192
- Pickard AS, Kohlmann T, Janssen M, Bonsel G, Rosenbloom S, Cella D (2007a) Evaluating equivalency between response systems: Application of the Rasch model to a 3-level and 5-level EQ-5D. Med Care 45(9):812–819
- Pickard AS, de Leon M, Kohlmann T, Cella D, Rosenbloom S (2007b) Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. Med Care 45(3):259–263
- Szende A, Oppe M, Devlin N (2007) EQ-5D value sets: inventory, comparative review and user guide. Springer, Dordrecht
- van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, Lloyd A, Scalone L, Kind P, Pickard AS (2011) Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health 15(5):708–715

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.

