

Chapter 1

An Introduction to EQ-5D Instruments and Their Applications



The aims of this chapter are

- to introduce the EQ-5D ‘family’ of questionnaires: what they are for, how they are used and what they measure;
- to explain the nature of the data that the EQ-5D questionnaires generate and how that affects the way that EQ-5D data should be analysed;
- to examine how the purposes for which EQ-5D data are collected affect the ways that they should be analysed and reported; and
- to describe good practice in data handling and preparing for statistical analysis of EQ-5D data.

Our focus, throughout this book, is on the analysis of EQ-5D data. The book is designed to meet the needs of those who have, or are planning to collect, EQ-5D data. Our hope is that this book will encourage all analysts, both those new to the EQ-5D and those experienced in using EQ-5D questionnaires, to make full use of the data provided by respondents, and to maximise the insights possible from those data.

It is also important to say what this book does not address. We do not provide guidance on methods of Patient Reported Outcome (PRO) data collection or PRO study design. For such guidance, you may wish to consult resources such as the SPIRIT-PRO¹ guidelines on inclusion of PROs in clinical trials (Calvert et al. 2018), the United States Food and Drug Administration (FDA) guidance to industry on the use of PRO measures in evidence to support labelling claims (FDA 2009); the European Medicines Agency (EMA) guidance regarding use of health-related quality of life (HRQoL) in labelling studies (EMA 2006); and the various good practice guidelines published by the International Society for Pharmacoeconomics & Outcomes Research (ISPOR), for example on electronic PROs (Zbrozek et al. 2013), and on collection of PROs in paediatric studies (Matza et al. 2013). Also, we do not offer

¹SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials.

guidance on which EQ-5D questionnaire to use in what circumstances—for example, in what populations to use the youth version of the EQ-5D (the EQ-5D-Y); whether to use the three- or five-level version; and how and when to use the paper, telephone, proxy or digital versions. Information on these Issues is provided in the User Guides available online at: www.euroqol.org.

A glossary of the EQ-5D terms used in this and subsequent chapters is in an appendix.

1.1 Measuring Health Using the EQ-5D

The EQ-5D is a concise, generic measure of self-reported health which is accompanied by weights reflecting the relative importance to people of different types of health problems. The concept of health being measured by EQ-5D is variously described as health status or HRQoL,² the latter of which might be defined as:

The value assigned to duration of life as modified by the impairments, functional status, perceptions and social opportunities that are influenced by disease, injury, treatment or policy. (Patrick and Erickson 1993)

The EQ-5D is ‘generic’ because it measures health in a way that can be compared across different sorts of patients, disease areas, and treatments. The researchers who developed it—the EuroQol Group—aimed to develop a questionnaire which was brief, minimised the burden of data collection, and could be used in a wide variety of health care sector applications (Devlin and Brooks 2017). The ‘5D’ in its name refers to its use of 5 dimensions for describing health states: Mobility, Usual Activities, Self-care, Pain & Discomfort and Anxiety & Depression. In the original EQ-5D questionnaire (Fig. 1.1), now known as the EQ-5D-3L, three levels of problems are described in each dimension, representing no, moderate, or extreme problems in the Pain & Discomfort and Anxiety & Depression dimensions and no, some, and inability to in the Mobility, Usual Activities and Self-care dimensions.³ In the more recent EQ-5D-5L (Fig. 1.2), the number of levels has been expanded from three to five and these are explicitly expressed as no, mild, moderate, severe and extreme or unable to (Herdman et al. 2011). A version of the instrument, the EQ-5D-Y (Fig. 1.3), has been developed for young people and children, retaining the same five dimensions (Wille et al. 2010).

In each case, the questionnaires are designed mainly for self-completion, either by people who are receiving treatment (for example patients in a clinical trial) or people in other settings (for example a sample of the general public in a population health survey). (As well as the self-report questionnaire, there are also ‘interview’

²For a discussion of definitional and conceptual issues relating to HRQoL, see Morris et al. (2012), Sect. 11.3.

³For the Mobility dimension the worst level is ‘confined to bed’.

MOBILITY

- I have no problems in walking about
- I have some problems in walking about
- I am confined to bed

SELF-CARE

- I have no problems with self-care
- I have some problems washing or dressing myself
- I am unable to wash or dress myself

USUAL ACTIVITIES (e.g. work, study, housework, family or leisure activities)

- I have no problems with performing my usual activities
- I have some problems with performing my usual activities
- I am unable to perform my usual activities

PAIN / DISCOMFORT

- I have no pain or discomfort
- I have moderate pain or discomfort
- I have extreme pain or discomfort

ANXIETY / DEPRESSION

- I am not anxious or depressed
- I am moderately anxious or depressed
- I am extremely anxious or depressed

Fig. 1.1 EQ-5D-3L descriptive system. *Source* EuroQol Research Foundation. *EQ-5D-3L User Guide, 2018*. Latest version available from: <https://euroqol.org/publications/user-guides>

and ‘proxy’ versions, designed for special cases where people whose EQ-5D data are being collected cannot complete a self-report questionnaire themselves.) For this reason, the EQ-5D belongs to a category of questionnaires often referred to as PROs and sometimes as Patient Reported Outcome Measures (PROMs). PROs aim to measure people’s subjective assessment of their own health in a manner that is systematic, valid and reliable. There is growing recognition that such data from

Under each heading, please tick the ONE box that best describes your health TODAY.

MOBILITY

- I have no problems in walking about
- I have slight problems in walking about
- I have moderate problems in walking about
- I have severe problems in walking about
- I am unable to walk about

SELF-CARE

- I have no problems washing or dressing myself
- I have slight problems washing or dressing myself
- I have moderate problems washing or dressing myself
- I have severe problems washing or dressing myself
- I am unable to wash or dress myself

USUAL ACTIVITIES (e.g. work, study, housework, family or leisure activities)

- I have no problems doing my usual activities
- I have slight problems doing my usual activities
- I have moderate problems doing my usual activities
- I have severe problems doing my usual activities
- I am unable to do my usual activities

PAIN / DISCOMFORT

- I have no pain or discomfort
- I have slight pain or discomfort
- I have moderate pain or discomfort
- I have severe pain or discomfort
- I have extreme pain or discomfort

ANXIETY / DEPRESSION

- I am not anxious or depressed
- I am slightly anxious or depressed
- I am moderately anxious or depressed
- I am severely anxious or depressed
- I am extremely anxious or depressed

Fig. 1.2 EQ-5D-5L descriptive system. Source EuroQol Research Foundation. *EQ-5D-5L User Guide, 2019*. Latest version available from: <https://euroqol.org/publications/user-guides>

Under each heading, please tick the ONE box that best describes your health TODAY

Mobility (*walking about*)

- I have no problems walking about
- I have some problems walking about
- I have a lot of problems walking about

Looking after myself

- I have no problems washing or dressing myself
- I have some problems washing or dressing myself
- I have a lot of problems washing or dressing myself

Doing usual activities (*for example, going to school, hobbies, sports, playing, doing things with family or friends*)

- I have no problems doing my usual activities
- I have some problems doing my usual activities
- I have a lot of problems doing my usual activities

Having pain or discomfort

- I have no pain or discomfort
- I have some pain or discomfort
- I have a lot of pain or discomfort

Feeling worried, sad or unhappy

- I am not worried, sad or unhappy
- I am a bit worried, sad or unhappy
- I am very worried, sad or unhappy

Fig. 1.3 EQ-5D-Y. *Source* EuroQol Research Foundation. *EQ-5D-Y User Guide, 2014*. Latest version available from: <https://euroqol.org/publications/user-guides>

patients provides important information that complements the clinical endpoints traditionally used in medical care, and can pick up problems and issues missed by them (Appleby et al. 2015). For example, Robert Temple from the FDA stated that “The use of Patient Reported Outcome instruments is part of a general movement toward the idea that the patient, properly queried, is the best source of information about how he or she feels” (Bren 2006). The EQ-5D is one of the most widely used PRO measures internationally, and by 2016 the EQ-5D-3L was available in 176 language versions the EQ-5D-5L 123 and the EQ-5D-Y 40 (Devlin and Brooks 2017).

The EQ-5D questionnaire comprises two parts. The first is the EQ-5D descriptive system, as shown in Figs. 1.1, 1.2, and 1.3. Respondents are asked to tick boxes to indicate the level of problem they experience on each of the five dimensions. The combination of these ticks under each dimension describes that person’s EQ-5D self-reported health state, often called an ‘EQ-5D profile’, which is described in more detail below.

The second part of the questionnaire is the EQ VAS, so called because it incorporates a Visual Analogue Scale. This captures the respondent’s overall assessment of their health on a scale from 0 (worst health imaginable) to 100 (best health imaginable). The current versions of the EQ-5D-3L and 5L use the same EQ VAS, shown in Fig. 1.4, but the original version of the 3L had a slightly different format, as does the EQ-5D-Y.

The EQ-5D profile data can also be supplemented by using a ‘scoring’ or ‘weighting’ system to convert profile data to a single number—EQ-5D values. These scoring systems are usually based on preferences—that is, the problems on each dimension are weighted to reflect how good or bad people think they are. So, for example, many studies have shown that problems with pain and discomfort often carry more weight than problems with self-care as reported by the EQ-5D (see Szende et al. 2007), and this is reflected in the way questionnaire respondents’ profile data is summed. These EQ-5D values—which are sometimes referred to in the literature as the EQ-5D Index, or quality of life weights or utilities—are constructed to lie on a scale anchored by the value 1, full health, and 0, dead. EQ-5D values cannot take a value higher than 1, but values less than 0 are possible for health states considered to be worse than dead.

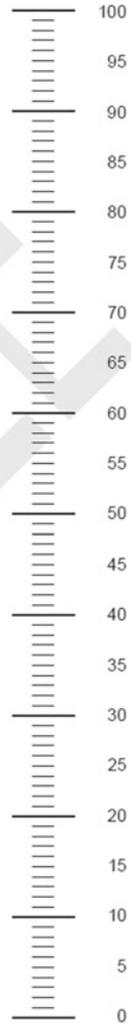
A full set of values for each possible EQ-5D profile is often called a ‘value set’. These values are obtained from stated preference studies, where members of the general public⁴ are asked to imagine living in health states described by the EQ-5D descriptive system, and to engage in a series of tasks designed to gauge how good or bad they consider those health states to be. A variety of methods can be used to elicit these preferences and to model them to create weights for the components of

⁴By convention, and for normative reasons, the general public’s stated preferences are usually argued to be those relevant to constructing these value sets (see, for example, Neumann et al. 2017). Value sets and their use are discussed in more details in Chap. 4.

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY =

The best health
you can imagine



The worst health
you can imagine

Fig. 1.4 EQ VAS (current EQ-5D-5L and EQ-5D-3L version). *Source* EuroQol Research Foundation. *EQ-5D-5L User Guide, 2019*. Latest version available from: <https://euroqol.org/publications/user-guides>

the EQ-5D profiles. The resulting ‘value sets’—the complete lists of values for each of the 243 profiles described by the EQ-5D-3L and EQ-5D-Y, and for the 3125 states described by the EQ-5D-5L—differ depending on what methods were used to elicit and model the preferences. They may also differ by country, reflecting differences in preferences across cultures and regions. Being aware of the properties of these value sets, and the difference they might make to your analysis of EQ-5D profile data, is important, and we discuss this further below and in Chap. 4.

1.2 What does the EQ-5D Measure?

The two parts of the EQ-5D questionnaire, combined with the value sets, means that the instrument generates three distinct types of data: the EQ-5D profile; the EQ VAS; and the EQ-5D values.

Each of these elements measures a somewhat different underlying construct of health. It is important to understand the nature of what is being measured in each case, since this affects hypotheses both about the expected relationship between these elements and between them and other data collected on respondents’ health and other characteristics.

1.2.1 *The EQ-5D Profile*

A respondent’s EQ-5D profile is a summary of the responses that they give to the descriptive system component of the EQ-5D self-report questionnaire. It can be described as five sentences, or summarised as a series of numbers representing the levels of problems in the order that the dimensions appear. Boxes 1.1, 1.2, and 1.3 give a fuller description.

Box 1.1. What are EQ-5D profiles?

A set of responses to the statements given in the descriptive system element of the EQ-5D questionnaire describes a health state or ‘profile’ as a combination of dimensions and levels within dimensions. For example, a completed questionnaire may be like this:

<p>Under each heading, please tick the ONE box that best describes your health TODAY.</p> <p>MOBILITY</p> <p>I have no problems in walking about <input checked="" type="checkbox"/></p> <p>I have some problems in walking about <input type="checkbox"/></p> <p>I am confined to bed <input type="checkbox"/></p> <p>SELF-CARE</p> <p>I have no problems with self-care <input checked="" type="checkbox"/></p> <p>I have some problems washing or dressing myself <input type="checkbox"/></p> <p>I am unable to wash or dress myself <input type="checkbox"/></p> <p>USUAL ACTIVITIES (e.g. work, study, housework, family or leisure activities)</p> <p>I have no problems with performing my usual activities <input type="checkbox"/></p> <p>I have some problems with performing my usual activities <input checked="" type="checkbox"/></p> <p>I am unable to perform my usual activities <input type="checkbox"/></p> <p>PAIN/DISCOMFORT</p> <p>I have no pain or discomfort <input type="checkbox"/></p> <p>I have moderate pain or discomfort <input type="checkbox"/></p> <p>I have extreme pain or discomfort <input checked="" type="checkbox"/></p> <p>ANXIETY/DEPRESSION</p> <p>I am not anxious or depressed <input type="checkbox"/></p> <p>I am moderately anxious or depressed <input checked="" type="checkbox"/></p> <p>I am extremely anxious or depressed <input type="checkbox"/></p>	<p>Levels of perceived problems are coded as follows:</p> <p><input checked="" type="checkbox"/> Level 1 is coded as a '1'</p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Level 2 is coded as a '2'</p> <p><input checked="" type="checkbox"/></p> <p><input type="checkbox"/></p> <p><input type="checkbox"/> Level 3 is coded as a '3'</p> <p><input type="checkbox"/></p> <p><input checked="" type="checkbox"/></p>
--	--

This profile can be described as a series of five sentences. For example, this respondent has:

- No problems in walking about
- No problems with self-care
- Some problems with performing their usual activities
- Extreme pain or discomfort
- Moderate anxiety or depression

In Box 1.2 we describe how these profiles may be more concisely summarised.

Box 1.2. Summarising EQ-5D profiles

A simpler way than using five sentences to summarise a profile is to assign each level a number and describe the profile as a five-number string, representing the level of each dimension in the order in which they appear in the questionnaire. The numbers used are: no problems = 1; some problems = 2; and extreme problems or unable to = 3. So, for example, no problems in any dimension is 11111, some problems in every dimension is 22222, and extreme problems in every dimension is 33333. The profile shown in Box 1.1 is 11232.

EQ-5D-5L profile data can be summarised in the same way. 11111 again means no problem on any of the five dimensions of health and the worst health state is 55555. The profile labels are not directly equivalent between the 3L and the 5L, except for 11111, which means no problems on any dimension. The worst health profiles, 33333 and 55555, describe different underlying health states because the worst level for mobility in the 3L is ‘confined to bed’ whereas in the 5L it is ‘unable to walk about’. Similarly, the ‘middle’ states, 22222 and 33333, mean different things, as 3L level 2 refers to ‘some’ problems, but 5L level 3 refers to ‘moderate’ problems.

The numbers given to levels within dimensions are ordinal—for example, 3 is worse than 2 and 2 is worse than 1. However, the profile labels are categories, not numbers, and do not even have ordinal properties. They do have a limited logical ordering—see Devlin et al. (2010) and Parkin et al. (2010) for further details—and in some cases can be used to compare profiles. For example, profile 11111 is better than profile 11112 (it logically dominates it) and 11112 is better than 11122. But we cannot say anything about *how much* better 11111 is compared to 11112. Moreover, we cannot say whether 11112 is better or worse than a profile such as 11121. That depends on the relative importance attached to some problems with anxiety & depression compared with some problems with pain & discomfort.

Chapter 2 demonstrates how health profiles can be compared to make judgements about whether health has improved, using only the ordinal properties of the levels within profiles. But to compare health profiles such as 11112 and 11121 and to measure the magnitude of the difference between any profiles requires a scoring system that assigns weights to each profile. EQ-5D value sets achieve that, using data from stated preferences studies to convert the profile data into a single, cardinal number. We examine the use of value sets in detail in Chap. 4.

Box 1.3. How many EQ-5D profiles?

For the EQ-5D-3L, there are $3^5 = 243$ possible profiles. There are three groups of profiles that include only two levels (1 and 2, 2 and 3 or 1 and 3), with $2^5 = 32$ profiles (13% of all profiles) in each group. Therefore, for each level there are $3^5 - 2^5 = 211$ profiles that include at least one of that level. So:

- 32 (13%) do not include a level 3 in any dimension
- 32 (13%) include only level 2 and 3
- 211 (87%) include at least one level 1
- 211 (87%) include at least one level 3

The number of unique profiles described by the EQ-5D-5L is $5^5 = 3125$. There are five groups of profiles that include only four levels, with $4^5 = 1024$ profiles (33% of all profiles) in each group. Therefore, for each level there are $5^5 - 4^5 = 2101$ profiles that include at least one of that level, $5^5 - 3^5 = 2882$ that contain at least one of each of two different levels and $5^5 - 2^5 = 3100$ that contain at least one of each of three different levels. So:

- 1024 (33%) do not include a level 1 in any dimension
- 1024 (33%) do not include a level 5 in any dimension
- 2101 (67%) include at least one level 5
- 2882 (92%) include at least one level 4 or a level 5
- 32 (1%) include only levels 1 and 2
- 32 (1%) include only levels 4 and 5
- 3093 (99%) include levels 1, 3 and 5
- 243 (8%) include only levels 1, 3 and 5.

In practice, not all profiles have an equal probability of being observed. For example, data obtained from the general population often contain a large proportion of profile 11111. In patient data sets, observations are often clustered on a sub-set of profiles relevant to those patients' condition; and some profiles are almost never observed because they contain unusual combinations of levels—for example the EQ-5D-3L profile 33133, in which there are extreme problems with everything except usual activities, where there are no problems.

The profile element of the EQ-5D questionnaire can be categorised as an example of a Health Status Measurement questionnaire, broadly defined (Bowling 2001, 2004). As noted earlier, the EQ-5D is often also described in the literature as measuring HRQoL. However, the concept of quality of life, and which aspects of it are seen as health-related, is often not precisely defined. Because the EQ-5D is a generic instrument, the EQ-5D profile will not capture everything that matters to all people with respect to their health status or HRQoL, and does not claim to do so. That means that, for some diseases and patients, there may be aspects of health that are important which the EQ-5D does not fully reflect, and this may be important to consider in your analysis of the data.

1.2.2 EQ VAS

The EQ VAS can be thought of as showing how patients feel about their own health overall. Their overall score will reflect both the relative importance that they place on the different aspects of their health that are included in the EQ-5D descriptive system and other dimensions of health that are not. The EQ VAS therefore provides information that is complementary to the EQ-5D profile. For example, it is often observed that some people who report no problems in any EQ-5D dimension rate their health as less than 100 on the EQ VAS (for example, see Devlin et al. 2004). Chapter 3 discusses other evidence for this, for example that the average EQ VAS scores decline with age even for those whose profile is 11111. Further, although profiles are systematically related to the EQ VAS scores in regression analyses, they only partially explain them (Feng et al. 2014).

1.2.3 EQ-5D Values

As noted above, EQ-5D values data are produced by applying value sets to summarise the EQ-5D profile data. The nature of these value sets, and their characteristics, are influenced by their principal application, which is in the estimation of quality-adjusted life years (QALYs). It is their use in this context that determines the anchors for the scale of 1 for full health and 0 for dead.⁵

It is important to note that using these value sets to generate EQ-5D values data introduces a source of exogenous variance into the analysis of profile data which can bias statistical inference (Parkin et al. 2010). Each value set places a different weight on the various levels and dimensions of the profile data, reflecting underlying differences in preferences, the methods used to elicit them, or both. This means that whether there are statistically significant differences in the EQ-5D values between, for example, two arms of a clinical trial, or between two regions in a national health survey, may depend on which value set is used, and the relative importance it puts on the different types of health problems and improvements in them.

More generally, there is *no* neutral way to summarise the data from the EQ-5D profile into a single number. This is not an issue that is only relevant to the EQ-5D instruments: these same points are relevant to the scoring and weighting systems used in *all* generic or condition specific PROs. Any method of combining responses to multiple questions must entail some weight being placed on each question. Even if preference-based weights were not used, and the dimensions of a PRO were equally weighted, that would imply a strong value judgement about the relative importance of various kinds of health problems that may or may not reflect the views of the people who self-reported their health on that PRO. Analysts should be aware of this, and check for the sensitivity of results to the choice of value set.

⁵The convention of anchoring at dead = 0 is very widely accepted, but could be debated—see Sampson et al. (2019).

1.2.4 Which Aspect of the Information Provided by the EQ-5D Should be the Primary Focus of My Analysis?

When considering which element of the EQ-5D data should be the primary focus of analysis, and what methods of analysis should be used, users should be guided by the *purpose* of collecting EQ-5D data and how the results will be used. Table 1.1 provides an overview of the main contexts in which EQ-5D data are collected, and implications regarding the analysis of the resulting data.

There are advantages in being able to summarise and represent a health profile by a single number like the EQ-5D values—for example, it simplifies statistical analysis. However, as we have already emphasised, there is no neutral set of weights that can be used for that purpose: they all embody judgements about what is meant by importance and the appropriate source of information for judging importance. It is therefore not possible to offer generalised guidance about which set of weights should be used if the sole purpose is to summarise profile data for descriptive or inferential statistical analysis. Users should consider the wider purpose for which the summary will be used. If the purpose is simply to provide descriptive information, then it may be better not to use EQ-5D values, but to focus analysis on the profile data themselves (see Chap. 2). This may also be preferable because the EQ-5D value provides less detailed information than the EQ-5D profile it is summarising. Focussing on the EQ-5D values may obscure the underlying information on the type and severity of problems affecting patients that the profile data provide (for example, see Gutacker et al. 2013).

Further, in some cases where a single number is required to represent health, for example, in the generation of population norms (Kind et al. 1999), it may be more appropriate to focus on the EQ VAS data provided by patients or populations, rather than applying the EQ-5D value sets to their profile data.

Economic Evaluation

Where the economic evaluation of treatment is the main goal of analysis, this has implications for the analysis of EQ-5D data. A key requirement for a health measure to use in cost effectiveness analysis is that it should provide an unambiguous measure of effectiveness. That is, higher EQ values should represent a better state of health and the same differences between EQ values should have the same level of importance. For example, the difference between 0.87 and 0.91 should represent the same degree of change as between 0.22 and 0.26. However, there is arguably a further requirement if the measure of effectiveness is to be based on economics principles, such as those embodied in cost utility analysis—essentially, that the weights need to represent ‘values.’ Just as costs represent the total value of resources used, that is the volume of each type of resource weighted by their individual value, effectiveness in the context of economic evaluation should represent the value of health output, that is the amount of health generated weighted by its value.

Table 1.1 Example of types of studies and some considerations for analysis

Types of studies or health care contexts in which EQ-5D data are collected	What questions are being asked?	What are the implications for data analysis?
Clinical trials	Is this technology effective and cost-effective relative to the comparator in the sample of patients included in this trial?	EQ-5D values are required for estimation of QALY gains. The EQ-5D profile and EQ VAS can provide additional evidence on relative effectiveness. Cluster analysis can be used to identify responder/non-responder groups
Observational studies of patient populations	The focus of these studies varies but could include: how does self-reported health change through time in a given patient group? How do patients' health compare to the general public? What evidence is there of response to treatment?	Descriptive analysis of EQ-5D profile and EQ VAS at each observation and analysis of changes between repeated observations. EQ-5D values will be required if estimation of QALYs is a goal. Cluster analysis can be used to identify responder/non-responder groups
Population health surveys	How does the health of a population compare with that of others? What is the burden of ill health?	Comparisons of EQ-5D profile and EQ VAS between sub-populations. EQ-5D values can provide a means of summarising profile data as a single number (although there are caveats about the use of values in this context, as we note in the following paragraphs)
Routine data collection in the health care system ('PROMs programmes')	How much variation is there between providers in improving patient health? How do patients' health and health improvements compare between different conditions and treatments? How does the cost effectiveness of different procedures compare?	Comparisons of profile and EQ VAS. EQ-5D values can be used as a way of summarising profile data as a single number, although caution is required (see p. 12) EQ-5D values are relevant where QALY estimation is required
Shared decision making between a patient and their doctor	What problems is this patient reporting? How difficult do they find these problems overall? How should this effect choice of treatment?	The individual patient's profile and EQ VAS are the focus. These may be benchmarked against evidence from other patients

There is ongoing debate over the extent to which the commonly-used stated preferences methods used adequately reflect underlying notions of ‘value’, and about the adequacy of QALYs as a measure of societal benefit from treating ill health. However, there appears to be general acceptance (for example, among Health Technology Appraisal bodies, like the National Health Care Institute (Zorginstituut) in The Netherlands, and the United Kingdom’s National Institute for Health and Care Excellence) that value sets available for EQ-5D instruments, based on the preferences of adult members of the general public, are usually appropriate for use in cost effectiveness analysis (NICE 2013; Zorginstituut Nederland 2016; Neumann et al. 2017).

Further detail on EQ-5D values, including which value set to use, and the analysis of EQ-5D values data, is provided in Chap. 4.

1.3 EQ-5D Data Collection and Data Handling

Where EQ-5D data are captured electronically, manual data entry is not required. However, in many cases, EQ-5D questionnaires are still completed in paper format. Where this is the case, data will need to be coded and entered manually. As this process is subject to human error, best practice for EQ-5D questionnaires is the same as any other self-completed paper questionnaire and entails double entry—that is, data being entered twice, and files compared for anomalies, which are then checked against the hardcopy.

Coding and data entry for the descriptive system are relatively straightforward. It is recommended that levels are coded as 1, 2 and 3 (for the EQ-5D-3L) and 1, 2, 3, 4 and 5 (for the EQ-5D-5L) in each dimension, to enable easy generation of the conventional 5-number profile label. Missing data need to be flagged as do any unusual responses, for example if more than one level is ticked on a dimension, although the latter are relatively rare.

EQ VAS data collected electronically are also very straightforward. However, the paper format of the original and current versions of the EQ VAS used in the EQ-5D-3L and EQ-5D-5L (see Figs. 1.4 and 1.5) and the current version of the EQ-5D-Y (see Fig. 1.6) require respondents to draw a line or mark a cross on the VAS to record their response. The resulting data can require a considerable degree of interpretation in coding responses. For example, Feng et al. (2014) noted, from qualitative analysis of a sub-sample of English National Health Service (NHS) PROMs data, a number of common response types with respect to the EQ VAS data (see Table 1.2).

Whereas a type 1 response in Table 1.2 is the only response which strictly complies with the EQ VAS instructions, Feng et al. (2014) argue that types 2 and 3 also provide unambiguous responses that can be captured accurately and reflect the same meaning to the score intended by respondents. Together, types 1–3 covered 88% respondents in the data presented in Table 1.2. Other types, including missing and ambiguous responses (types 5 and 6) require separate codes to flag these issues in analysis. Similar issues may exist with EQ VAS data from the EQ-5D-Y.

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked **100** and the worst state you can imagine is marked **0**.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

**Your own
health state
today**

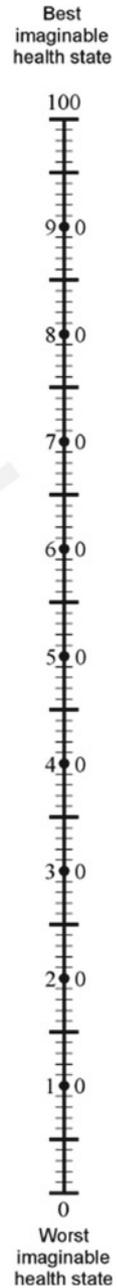


Fig. 1.5 EQ VAS (Original EQ-5D-3L version). *Source* EuroQol Research Foundation. *EQ-5D-3L User Guide*, 2015. Latest version available from: <https://euroqol.org/publications/user-guides>

How good is your health TODAY

The best health you can imagine



- We would like to know how good or bad your health is TODAY.
- This line is numbered from 0 to 100.
- 100 means the best health you can imagine. 0 means the worst health you can imagine.
- Please mark an X on the line that shows how good or bad your health is TODAY.

Fig. 1.6 EQ VAS (EQ-5D-Y version). *Source* EuroQol Research Foundation. *EQ-5D-Y User Guide, 2014*. Latest version available from: <https://euroqol.org/publications/user-guides>

Table 1.2 Types of responses to the original EQ-5D-3L EQ VAS

	EQ VAS response type, from most to least frequent	% responses
1	Drew a line from the box towards the EQ VAS, sometimes touching or crossing it. This is the way that the EuroQol Group intends the EQ VAS to be completed	45
2	Indicated precisely a horizontal level on the VAS, but did not draw a line to it. For example, ticks, crosses, lines, arrows, asterisks on or beside the VAS, or a tightly drawn circle around a specific number or tick mark	32
3	Drew a vertical line extending from 0 up to a point parallel with a point on the VAS	11
4	Missing	8
5	Drew a vertical line parallel to the VAS, but not extending from 0, or circled an area of the VAS. This indicated a range rather than a single point	4
6	Gave an unclear response. For example, multiple markings on the VAS or vertical lines drawn from 100 downwards	1

Source Feng et al. (2014). Response types have been combined across both pre-and post-surgery responses and re-ordered by frequency

The current format of the EQ VAS in the EQ-5D-5L and EQ-5D-3L (see Fig. 1.4) entails respondents both noting a number in the box and marking a cross on the scale. In electronic data capture, the two are identical. In paper completion, there is potential for the two responses to differ, and best practice would suggest capturing both and reporting any such discrepancies.

1.4 Before Starting Your Analysis

1.4.1 Treatment of Missing Data—What to Do, What Not to Do

There are broadly two types of missing EQ-5D data. Data can be missing altogether—for example, where an elective surgery patient in the English NHS fails to complete and return their post-surgery PROMs questionnaire. Or data can be missing in part—for example, where the patient completes an EQ-5D questionnaire, but provides incomplete profile data, or does not complete the EQ VAS.

General guidelines (i.e. relating to PRO data, rather than specifically the EQ-5D) often indicate that a substantial amount of missing data can compromise the validity of analysis—but what constitutes ‘substantial’ is a matter of opinion. For example, based on the German Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen) standard approach, data from at least 70% of patients at both baseline and one follow up visit are needed to consider analysis of that data valid for its purposes. However, ‘percent missing’ is not

defined consistently across the literature and different definitions on how to estimate the amount of missing data may lead to different practices and results (Coens et al. 2020). Further, even where there are high rates of missing data, analysis of available data may still yield insights into the sub-group who did respond, even if results cannot be generalised to non-responders. In short, there are no hard and fast rules. However, it is important for analysts to report missing data, and to be mindful of potential limitations arising from loss of generalisability.

In general, you should provide data descriptions, state the assumptions underlying the handling of the missing EQ-5D data, and conduct sensitivity analyses to the selected assumption. Included in the data description should be the amount of missing data, missing data patterns, and the association between missing data and observed data, for example respondents' age, gender and any previously observed EQ-5D data for that respondent (Faria et al. 2014).

Analytical methods used for missing data in general are applicable to the EQ-5D; users are advised to consult a statistical text for details. Essentially, it is necessary to consider the assumed form that missingness takes for the data—Missing Completely At Random (MCAR), Missing At Random (MAR) or Missing Not At Random (MNAR) (Little and Rubin 1987)—and to select a method for dealing with this appropriate to that form.

If MCAR, where a respondent's missing data are not related to that person's socio-demographic or other characteristics, analysis can assume that the missing data follow the same patterns as the non-missing data.

If MAR, where a respondent's missing data is related to their observed characteristics, but not any unobserved characteristics, analysis can assume that we have a random sample of respondents with those characteristics and make inferences from that sample about the data that are missing. Multiple Imputation (MI) has been increasingly used in recent years for EQ-5D data with MAR (Ratcliffe et al. 2005; Kaambwa et al. 2012; Simons et al. 2015).

If MNAR, where a respondent's data are missing because of their characteristics, we do not have random samples of people with different characteristics and require more complex analytical methods to deal with resulting selection bias. The Heckman selection model has been applied to EQ-5D values data that are assumed to be MNAR (Kaambwa et al. 2012).

Recent guidance suggests that data analysts should evaluate the sensitivity of the analysis to the MAR assumption using methods such as the weighting or pattern mixture approaches (Faria et al. 2014; Simons et al. 2015). In particular the evaluation should examine how the results might change when a MNAR assumption is made to the missing EQ-5D data.

There are two missing data issues specific to EQ-5D data. First, there is the issue of what should be done where the user wishes to analyse profiles and some but not all of the profile items are missing. Bad practice includes substituting for a respondent's missing profile items an average derived from their non-missing items and substituting an average derived from the non-missing items in the sample as a whole. It might be possible to use MI in this context, but there are currently no

examples on which to base guidance. Conservative guidance is therefore to treat as missing any profiles based on missing profile items.

The second is where some or all of the profile items are missing, and the user wishes to analyse EQ-5D values. For this, MI may be an appropriate method if the data are assumed MAR, but an issue is whether this should be applied to profile items, from which an EQ-5D value is calculated, or to EQ-5D values directly (Faria et al. 2014). In practice, the decision depends on the observed missing data pattern and the sample size available for analysis (Simons et al. 2015).

1.4.2 Planning Your Analysis

A systematic review of the use of PROs in oncology conducted by the Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data (SISAQOL) Consortium (Pe et al. 2018) showed a widespread lack of clearly-specified a priori research hypotheses and a link with the design and statistical methods to be employed. New guidelines for protocol development (for example SPIRIT-PRO) and reporting of PROs (for example CONSORT-PRO⁶—see Calvert et al. 2013) also recognise this to be a common issue in PRO studies generally.

Before beginning analysis of EQ-5D data, you should therefore consider what questions you want to answer with your data. What are your hypotheses about, for example, how a treatment arm is expected to behave relative to a reference arm in a clinical trial? What assumptions underpin these hypotheses, for example what is your rationale and what evidence has informed that? This, in turn, should inform the statistical analysis plan (SAP) developed prior to analysis. Note that the content of SAPs will vary depending on the study type and study aims.

1.5 Guide to the Rest of this Book

In the remainder of this book, we explain in detail how each element of the data generated from using EQ-5D instruments—the profile data, EQ VAS and EQ values — can be analysed. We provide both a basic introduction to analysis in each case, assuming no prior knowledge of analysis of EQ-5D data, as well as introducing more advanced topics relating to analysis of EQ-5D data.

⁶CONSORT: Consolidated Standards Of Reporting Trials.

References

- Appleby J, Devlin N, Parkin D (2015) Using patient reported outcomes to improve health care. Wiley. ISBN: 978-1-118-94860-6
- Bowling A (2001) Measuring disease: a review of disease specific quality of life measurement scales, 2nd edn. Open University Press
- Bowling A (2004) Measuring health: a review of quality of life measurement scales. McGraw-Hill
- Bren L (2006) The Importance of Patient-Reported Outcomes ... It's All About the Patients. *FDA Consum* 40(6):26–32
- Calvert M et al (2018) Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. *JAMA* 6; 319(5):483–494
- Calvert M, Blazeby J, Altman DG et al (2013) Reporting of patient reported outcomes in randomised trials. *J Am Med Assoc* 309(8):814–822
- Coens C, Pe M, Dueck AC, Sloan J et al (2020) International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium. *Lancet Oncol* 21(2):e83–e96
- Devlin N, Brooks R (2017) EQ-5D and the EuroQol Group: past, present, future. *Appl Health Econ Health Policy* 15(2):127–137
- Devlin N, Hansen P, Selai C (2004) Understanding health state valuations: a qualitative analysis of respondents' comments. *Qual Life Res* 13(7):1265–1277
- Devlin NJ, 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
- European Medicines Agency (2006) Reflection paper on the regulatory guidance for the use of health related quality of life (HRQL) measures in the evaluation of medicinal products. EMA: Committee for medicinal products in human use. https://www.ema.europa.eu/documents/scientific-guideline/reflection-paper-regulatory-guidance-use-healthrelated-quality-life-hrql-measures-evaluation_en.pdf. Accessed 18 Dec 2018
- Faria R, Gomes M, Epstein D, White IR (2014) A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics* 32(12):1157–1170
- FDA (2009) Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. <https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf>. Accessed 18 Dec 2018
- Feng Y, Parkin D, Devlin N (2014) Assessing the performance of the EQ VAS in the NHS PROMS programme. *Qual Life Res* 23(3):977–989
- Gutacker N, Bojke C, Daidone S, Devlin N, Street A (2013) Hospital variation in patient-reported outcome at the level of EQ-5D dimensions: evidence from England. *Med Decis Making* 33(6):804–818
- Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D,onsel 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
- Kaambwa B, Bryan S, Billingham L (2012) Do the methods used to analyse missing data really matter? An examination of data from an observational study of Intermediate Care patients. *BMC Res Notes* 5:330
- Kind P, Hardman H, Macran S (1999) UK population norms for EQ-5D. Working Papers 172 Centre for Health Economics, University of York
- Little RJA, Rubin DB (1987) Statistical analysis with missing data. Wiley, New York
- Matza LS, Patrick D, Riley AW et al (2013) Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. *Value Health* 16:461–479
- Morris S, Devlin N, Parkin D, Spencer A (2012) Economic analysis in health care, 2nd edn. Wiley
- Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG (2017) Cost effectiveness in health and medicine, 2nd edn. Oxford University Press

- NICE (2013) Guide to the methods of technology appraisal. <https://www.nice.org.uk/process/pmg9/chapter/foreword>. Accessed 18 Dec 2018
- Parkin D, Rice N, Devlin N (2010) Statistical analysis of EQ-5D profiles: does the use of value sets bias inference? *Med Decis Making* 30(5):556–565
- Patrick DL, Erickson P (1993) *Health status and health policy: quality of life in health care evaluation and resource allocation*. Oxford University Press, New York
- Pe M, Dorme L, Coens C, Basch E, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Dirven L, Dueck AC, Devlin N, Flechtner HH, Gotay C, Griebisch I, Groenvold M, King M, Koller M, Malone DC, Martinelli F, Mitchell SA, Musoro JZ, Oliver K, Piau-Louis E, Piccart M, Pimentel FL, Quinten C, Reijneveld JC, Sloan J, Velikova G, Bottomley A (2018) Setting international standards in analyzing patient-reported outcomes and quality of life endpoints data consortium (SISAQOL). *Lancet Oncol* 19(9):e459–e469
- Ratcliffe J, Young T, Longworth L, Buxton M (2005) An assessment of the impact of informative dropout and nonresponse in measuring health-related quality of life using the EuroQol (EQ-5D) descriptive system. *Value Health* 8(1):53–58
- Sampson C, Parkin D, Devlin N (2019) Drop dead: is anchoring at ‘dead’ a theoretical requirement in health state valuation? Paper presented to the 35th scientific plenary meeting of the EuroQol Group, Barcelona, Spain
- Simons CL, Rivero-Arias O, Yu LM, Simon J (2015) Multiple imputation to deal with missing EQ-5D-3L data: should we impute individual domains or the actual index? *Qual Life Res* 24(4):805–815
- Szende A, Oppe M, Devlin N (2007) *EQ-5D value sets: inventory, comparative review and user guide*. Springer
- Wille N, Badia X, Bonsel G et al (2010) Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res* 19(6):875–886
- Zbrozek A, Hebert J, Gogates G et al (2013) Validation of electronic systems to collect patient-reported outcome (PRO) data—recommendations for clinical trial teams: report of the ISPOR ePRO systems validation good research practices task force. *Value Health* 16:480–489
- Zorginstituut Nederland (2016) *Guideline for economic evaluation in health care*. https://tools.ispor.org/PEGuidelines/source/Netherlands_Guideline_for_economic_evaluations_in_health_care.pdf. Accessed 3 July 2019

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.

