# Chapter 4
# Analysis of EQ-5D Values

The aims of this chapter are

- to introduce the properties and use of value sets;
- to highlight points to consider when choosing which value set to use;
- to provide guidance on statistical analysis of EQ-5D values, including descriptive statistics and inference; variance and heteroskedasticity; clustering; and regression methods; and
- to highlight the importance of conducting sensitivity analysis.

## 4.1 Value Sets and Their Properties

Despite the potential richness of the EQ-5D descriptive system demonstrated in the previous chapters, the instrument was originally developed as a measure of health status that could serve as the basis for summarising and comparing health outcomes (Williams 2005). In particular, it was designed to be a brief generic measure that would lend itself for the purpose of assigning a single summary value to each possible health profile (hereafter: 'EQ-5D values' or 'values'). The values, presented in country-specific value sets, are a major feature of the EQ-5D instrument, facilitating the calculation of quality-adjusted life years (QALYs) that are used to inform economic evaluations of health care interventions or policies on health.[1]

It is important to note that these values are a special case, based on people's strength of preference for different health profiles, of an index that generates a single summary number. Such indices in general (other than *values*), for example based on clinically-defined need, may be used for other purposes and, as we discuss below, it

---

[1]Note that it is beyond the scope of this book to offer guidance on how to conduct cost-effectiveness or cost-utility analysis. There are other resources providing detailed guidance on this, such as Drummond et al. (2015) and Sanders et al. (2016).

should not be assumed that using value sets is appropriate for any of those purposes.[2] Other possible uses of indices more broadly defined are to summarise EQ-5D profiles for statistical analysis, describing the health of a population, comparing population health (between regions, populations, or over time), describing severity of illness, and assessing population or patient priorities for treatment (Devlin and Parkin 2006). A relatively newly developed use is in routine outcomes measurement, to assess the performance of healthcare, for use as a hospital performance indicator (used to help patients choose which hospital to be referred to), and for use in measuring the productivity and performance of a healthcare system (Appleby et al. 2015).

A value set consists of weights that can convert each EQ-5D health profile into a value on a scale anchored at 1 (meaning full health) and 0 (meaning a state as bad as being dead). The scale allows negative values to be assigned to health states that are considered worse than dead. The values can be calculated by applying a formula that attaches a weight to each level in each dimension. In some cases, the formula allows for the possibility that combinations of problems might also affect preferences, via interaction effects.

The EQ-5D-3L describes 243 unique health profiles ($3^5$), whereas the EQ-5D-5L describes 3,125 possible unique health profiles ($5^5$). Most of the EQ-5D value sets have been obtained using stated preference data elicited from representative samples of the general public, thereby ensuring that they represent the societal perspective. The normative argument for using so-called "social" value sets is that for resource allocation purposes in publicly-or collectively-funded health care, the valuation of health states should reflect the preferences of the relevant general public (Weinstein et al. 1996; Sanders et al. 2016), since it is the general public who are ultimately funding health care and are the users of the health care system (Dolan 1997).

Value sets are commonly produced by valuing a selection of EQ-5D states and, by using econometric techniques, to extrapolate over the full set of health states. For the EQ-5D-3L, a subset of health states to be used in a valuation study was decided by the Group in 1990 (Rabin et al. 2007) along with a preferred method for obtaining the values, using a visual analogue scale (VAS) approach. For various reasons however, subsequent valuation studies have not always adhered to the standard approach, since these studies were often the result of locally led research initiatives. Apart from the choice of the health state design (i.e. deciding on the subset of states to be valued), studies differed in other ways, such as the valuation method and the (interviewer) protocol used, the number of respondents included, exclusion criteria for valuation responses, and modelling choices in arriving at a final value set. When the EQ-5D-5L was introduced, the EuroQol Group decided to return to having a more standardized approach by developing the EuroQol Valuation Technology platform (EQ-VT) (Oppe et al. 2014). Apart from standardization in terms of health state design, valuation methodology, and a computer-assisted personal interview mode of administration, a strict protocol of interviewer training and quality assurance during the entirety of the data collection process was developed and implemented (Ramos-Goñi et al. 2017a).

---

[2]However, we will only be discussing value sets in this chapter, not any other possible indices.

Values derived for EQ-5D have been based on various stated preference valuation techniques, such as the standard gamble (SG), time trade-off (TTO), VAS, person trade-off or rank-based techniques such as paired comparison, best-worse scaling and discrete choice methods. Since the first publication in 1997 (Dolan 1997), EQ-5D-3L value sets have been derived and published for many countries (www.euroqol.org). EQ-5D-5L valuation studies have been conducted from 2012 onwards, and the published value sets are listed on the EuroQol website at www.euroqol.org. EQ-5D-3L value sets were mainly based on TTO and VAS valuation methodology, although other techniques have also been used (Craig et al. 2009; Bansback et al. 2012). For the valuation of EQ-5D-5L, the EuroQol Group decided to explore the use of rank-based valuation methods to gain additional information (Devlin and Krabbe 2013). The current EQ-VT protocol for the valuation of EQ-5D-5L health states uses composite TTO and discrete choice valuation methodology (Oppe et al. 2014). There has been much discussion about the theoretical and empirical properties of the different valuation methods. In the health economics literature choice-based methods such as SG and TTO are often argued to have a more solid basis in economic theory than a rating approach such as VAS (Brazier et al. 1999; Drummond et al. 2015)— although for an alternative view see Parkin and Devlin (2006)–whereas discrete choice methodology is rooted in mathematical psychology and was further developed into random utility theory (McFadden 1974).

The EQ-5D-5L descriptive system was published before valuation studies were carried out and the subsequent publication of value sets derived from them. As an interim measure, the EuroQol Group coordinated a study that administered both the 3-level and 5-level versions of the EQ-5D to develop a mapping[3] function between the EQ-5D-3L value sets and the EQ-5D-5L descriptive system, resulting in (interim) value sets for the EQ-5D-5L (van Hout et al. 2012). 3,691 respondents completed both the 3L and 5L across 6 countries: Denmark, England, Italy, the Netherlands, Poland and Scotland. Different subgroups were targeted, and in most countries, a screening protocol was implemented to ensure that a broad spectrum of levels of health would be captured across the dimensions of EQ-5D for both the 5L and 3L descriptive systems.

Table 4.1a, b show two existing value sets with examples how to calculate the values for a certain health profile.

Finally, an important consideration is that attaching values to descriptive data introduces an exogenous source of variance, which can bias statistical inference (Parkin et al. 2010; Wilke et al. 2010). This is a special problem for applications where people's preferences are not directly relevant and is a key reason why it should not be assumed that values provide a suitable index for non-economics applications. Conclusions about whether there are statistically significant differences in, for example, the health of 2 regions, or health over time, or between 2 arms of a clinical trial, may be influenced by which value set is used. Furthermore, note that there is no such thing as a neutral value set or index; any weighting of EQ-5D profile data will influence the results, including the equally weighted Level Sum Score

---

[3]For further information on mapping, see Sect. 5.2.

**Table 4.1   a** An example of
applying the EQ-5D-3L value
set for the United Kingdom
(UK) to calculate EQ-5D
values. **b** An example of
applying the English
EQ-5D-5L value set to
calculate EQ-5D values

|  | Central estimate | Value for health profile 21232 |
|---|---|---|
| Constant | 1.000 | 1.000 |
| At least one 2 or 3 | 0.081 | 0.081 |
| At least one 3 (N3) | 0.269 | 0.269 |
| *Mobility* | | |
| Some problems | 0.069 | 0.069 |
| Confined to bed | 0.314 | |
| *Self-care* | | |
| Some problems | 0.104 | |
| Unable to | 0.214 | |
| *Usual activities* | | |
| Some problems | 0.036 | 0.036 |
| Unable to | 0.094 | |
| Pain/discomfort | | |
| Moderate | 0.123 | |
| Extreme | 0.386 | 0.386 |
| *Anxiety/depression* | | |
| Moderate | 0.071 | 0.071 |
| Extreme | 0.236 | |
| The value for health state 21232 | $1-(0.081 + 0.269 + 0.069 + 0.036 + 0.386 + 0.071) = 0.088$ | |
|  | Central estimate | Value for health profile 23245 |
| Constant | 1.000 | 1.000 |
| *Mobility* | | |
| Slight problems | 0.058 | 0.058 |
| Moderate problems | 0.076 | |
| Severe problems | 0.207 | |
| Unable to | 0.274 | |
| *Self-care* | | |
| Slight problems | 0.050 | |
| Moderate problems | 0.080 | 0.080 |
| Severe problems | 0.164 | |
| Unable to | 0.203 | |
| *Usual activities* | | |
| Slight problems | 0.050 | 0.050 |
| Moderate problems | 0.063 | |
| Severe problems | 0.162 | |

(continued)

**Table 4.1** (continued)

| | Central estimate | Value for health profile 23245 |
|---|---|---|
| Unable to | 0.184 | |
| *Pain/discomfort* | | |
| Slight | 0.063 | |
| Moderate | 0.084 | |
| Severe | 0.276 | 0.276 |
| Extreme | 0.335 | |
| *Anxiety/depression* | | |
| Slight | 0.078 | |
| Moderate | 0.104 | |
| Severe | 0.285 | |
| Extreme | 0.289 | 0.289 |
| The value for health state 23245 | $1-(0.058 + 0.080 + 0.050 + 0.276 + 0.289) = 0.247$ | |

index described in Chap. 2. This is not specific to the EQ-5D, applying equally to the scoring systems of other health measures, both generic and condition specific, including measures that simply sum ranked responses. For economic evaluation, the issue is rather different, because the exogenous influence of people's preferences is a desired feature when taking a societal perspective.

## 4.2 Positive and Normative Considerations in Choice of Value Set

In recent decades, a large number of EQ-5D value sets have been published, using a multitude of approaches and valuation techniques, with applications in various fields. Users of EQ-5D often question what the appropriate value set is for their particular use. The aim of this section is to provide advice on this question, largely following the earlier "Guidance to users of EQ-5D value sets" which was published as Chapter 4 of the EuroQol Group Monographs Volume 2: EQ-5D value sets: Inventory, comparative review and user guide (Devlin and Parkin 2007).

An obvious advantage of using a summary value to represent a health profile is that it simplifies statistical analysis. But since all value sets embody preferences about the relative importance of each level of each dimension, it is not possible to offer generalised guidance about which value set to use if the objective is to summarise profiles for descriptive or inferential statistical analysis. If there is not a clear purpose for using a summary value (especially based on social values), but rather an aim to provide information, it may be better if no value is used, but to report the descriptive

information as described in previous chapters. This also applies to describing the health of a population or patient group, or for comparing population health.

One of the most common uses of EQ-5D values remains in economic evaluation, with applications in cost-per-QALY/cost-effectiveness analysis (CEA) or cost-utility analysis (CUA). In CUA, the value set will be used to calculate QALYs, and the weights in the value set should represent "values", meaning that the health profiles described by the instrument should be weighed by the *value* of the health profile. To arrive at QALYs, the values should be anchored at 0 (corresponding to being dead or as bad as being dead), and 1, representing full health. A further requirement, although not essential for all cost-effectiveness analyses, is that the value set should be based on the societal perspective.

Often, economic evaluation is performed to provide evidence for a formal decision-making process. National health technology assessment bodies across the world routinely use economic evaluations to make decisions and recommendations about health care services. At the time of writing the EQ-5D is the preferred (or one of the preferred) health outcome measures recommended by pharmaceutical reimbursement authorities in at least 29 countries, including countries in Europe, North America, South America, Asia and Australia (Kennedy-Martin et al. 2020). When a value set needs to be selected to perform such an evaluation, the first consideration is pragmatic: does the relevant decision maker specify any requirements or preferences regarding which value set should be used? If recommendations will be made to more than one country on the basis of the evaluation's results, for example when performed alongside a multi-country clinical trial, the value set relevant to each separate country should be applied to the effectiveness data and reported to the decision makers in each separate country.

In the absence of specific requirements or guidelines from decision makers, analysts are left to make their own choices, for which broadly there are three main considerations to take into account: relevance to the decision-making context; empirical characteristics of the valuation study and modelling techniques; and the theoretical properties of the valuation methods.

Relevance to the decision-making context entails whether the values reflect the geographical and economic context in which resource allocation decisions are made, and whose values are considered to be relevant in the decision-making process. As mentioned in Sect. 4.1, there is a strong normative argument to opt for social valuations in economic evaluations informing decisions about collectively-funded health care. An alternative would be to use patients' values, because the preferences of patients who are actually experiencing the health states would be more well-informed than values generated from the general public being asked to imagine health states that may be hypothetical to them. Differences between patients' values and social values are widely observed (Zethraeus and Johannesson 1999; de Wit et al. 2000; Brazier et al. 2005; Ogorevc et al. 2019). Since the value set arguably should reflect the preferences of the potential recipients of healthcare, local (i.e. country-specific) value sets should be used when available. For a country for which no value set has been published and no local guidelines are available, practical aspects might be taken into consideration, such as considering a value set of a country that is most similar

in terms of e.g. demographics, geography, language, infrastructure, or health care system. Finally, the time period in which the valuation study has been performed is relevant. The UK EQ-5D-3L value set is still being used extensively at the time of writing, but the data collection for the valuation study dates back to 1993, while the UK has gone through many demographic and economic changes since then which might impact on preferences.

Empirical characteristics should be considered when choosing a value set. It is recommended that users study those characteristics before choosing a value set, looking at e.g. the response rate of the valuation study, whether the sample was representative of the general public, which valuation method was used, whether the health state design was appropriate, which mode of administration was used, the 'quality' of the data (were there many missing values, inconsistencies, low values for very mild health states or vice versa), were the econometric modelling techniques sound and appropriate, was the choice of the final model appropriate? These questions largely apply to EQ-5D-3L, since with the introduction of the EQ-VT platform for the valuation of EQ-5D-5L, many potential issues have been resolved by a high level of standardization and rigorous interviewer training and quality assurance.

The theoretical properties of the underlying valuation methods have been a controversial issue for decades. As mentioned in Sect. 4.1, so-called 'choice-based' methods such as SG and TTO have been preferred over a rating approach such as VAS. For the EQ-5D-3L, mainly VAS and TTO value sets are available. TTO based value sets have generally been preferred for purposes of economic evaluation, although it has been suggested that VAS value sets may be used for non-economics studies (Kind 2003). The EQ-VT protocol for EQ-5D-5L valuation studies uses composite TTO and discrete choice valuation techniques, offering the possibility to model a composite TTO based value set, or a value set based on a hybrid model combining composite TTO and discrete choice data (Feng et al. 2018; Ramos-Goñi et al. 2017b).

Based on the criteria discussed above, there may not be a single 'best' value set for any given application. Therefore, it is recommended to perform sensitivity analysis using other suitable value sets, to assess the impact of the choice of value set on results and conclusions. As mentioned above, many countries do not have a value set of their own and therefore have to use 'foreign' values; Parkin et al. (2010) showed that in a simulated economic evaluation experiment, whether or not an intervention is seen as effective in such a country might depend on which other country's value set it chooses. This stresses the relevance of which value set one chooses, and the importance of performing sensitivity analysis (see Sect. 4.8).

The value sets that are used in economic evaluation have a clear theoretical rationale that is the foundation for the values, the way that they are derived, and their meaning. As mentioned above, this rationale might not be relevant for other uses. The values used in economic evaluation are explicitly regarded as 'utilities', with a very specific definition attached to them. There is a clear meaning for the values 1 and 0 and for negative values. As mentioned above, a recognized stated preference technique such as TTO is often recommended to derive the values. Finally, there is a justification for the use of the general population as a source of EQ-5D values. The

values should be used in other applications only if the same theoretical rationale also
applies.

Figure 4.1 provides an overview of the considerations that should determine your
choice between the EQ-5D value sets. Choosing a value set is not simple, since many
factors are involved, such as the specific nature of the research application, the sort
of decisions it informs, and the context in which the evidence from your research
will be used. In longitudinal studies, the same value set should be applied throughout
the study. When the research aim is to make comparisons across respondents from
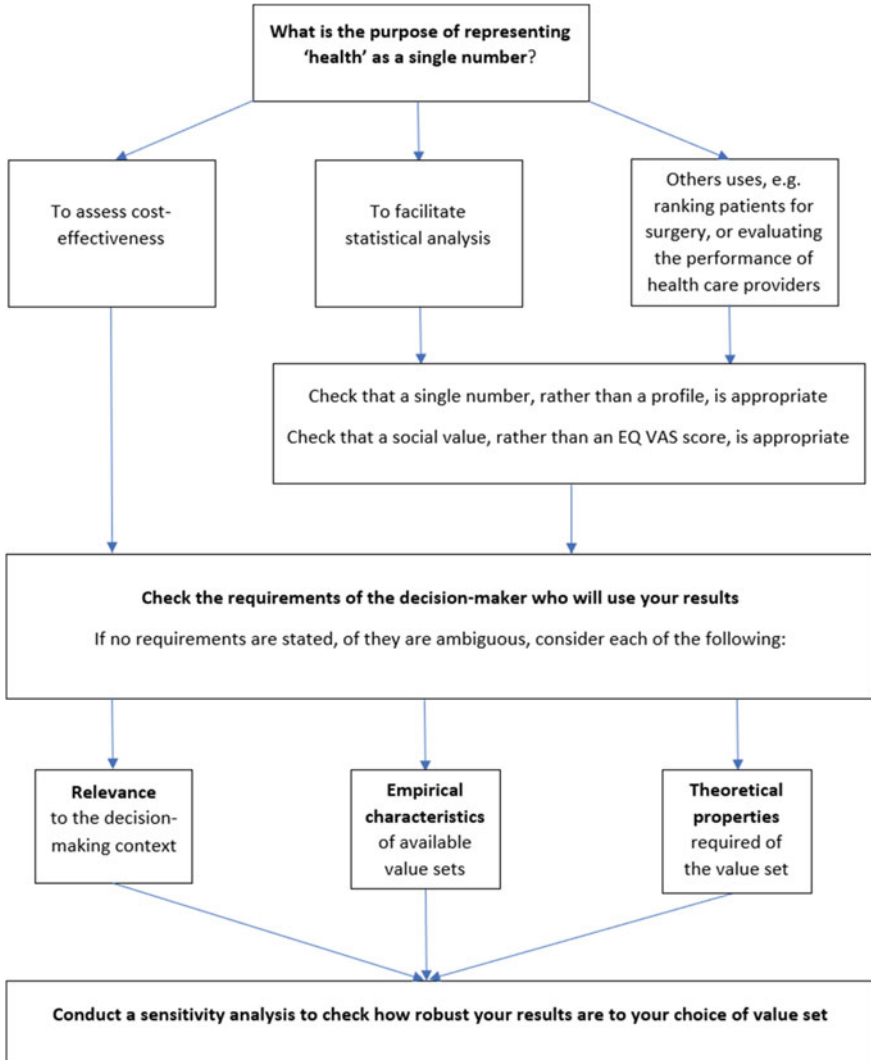


**Fig. 4.1** Guidance on which EQ-5D value set to use

different countries in a multinational cross-sectional study (rather than comparing value set characteristics) it will also be helpful to use a common value set if one is available, otherwise differences in country preferences would be added to the differences between respondents' health status. An example is the European VAS value set (Greiner et al. 2003).

## 4.3   Simple Descriptive Statistics and Inference

EQ-5D values can be presented in much the same way as EQ VAS data. Since the valuation methods underlying the values are meant to provide a scale with cardinal properties, for exploratory data analysis you can present a measure of central tendency (e.g. a mean or median), a spread (i.e. a measure of dispersion such as the standard deviation) and a shape (e.g. skewness, mode, or kurtosis). If the data is skewed, as is often the case with EQ-5D value data for general populations or mildly diseased patients, the median value could be used as measure of central tendency. As measure of dispersion one can also add minimums, maximums, and the inter quartile range (IQR) which is the difference between the 75th and 25th percentiles. If you are interested in the precision of the mean, you can use the standard error of the mean and a 95% confidence interval. Similar to EQ VAS data, a t-test can be used for comparing differences between means of different populations (or the same population over time). When you want to compare more than 2 groups, an Analysis of variance (ANOVA) can be used. The following tables and figures contain 2 examples of how to present EQ-5D value results. Table 4.2 and Fig. 4.2 present the results from a study

**Table 4.2** EQ-5D values before and after treatment

| EQ-5D value | Before treatment | After treatment |
|---|---|---|
| Mean | 0.567 | 0.727 |
| Standard error | 0.017 | 0.015 |
| Median | 0.60 | 0.810 |
| Standard deviation | 0.273 | 0.244 |
| 25th | 0.331 | 0.606 |
| 75th | 0.796 | 0.892 |
| Kurtosis | 2.96 | 5.18 |
| Skewness | −0.734 | −1.59 |
| Minimum | −0.429 | −0.349 |
| Maximum | 1 | 1 |
| Range | 1.429 | 1.349 |
| Observations | 251 | 249 |
| Missing values (percent) | 4 (1.6%) | 3 (1.6%) |

**Fig. 4.2** EQ-5D values
before and after treatment:
mean values and 95%
confidence intervals
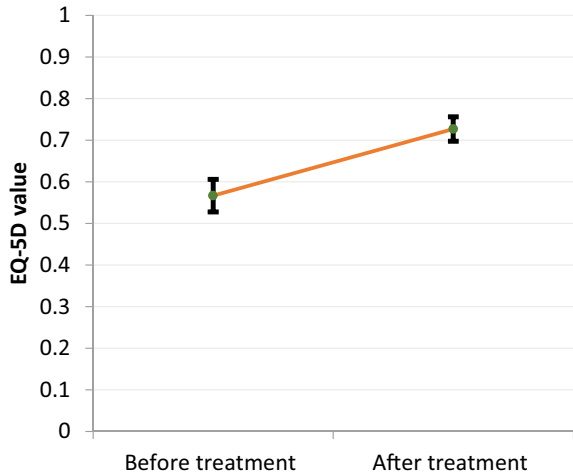


where the effect of a treatment on health status is investigated (the tables and figures
are based on hypothetical data and for illustration purposes only).

Table 4.3 and Fig. 4.3 show results for a patient population and 3 sub-groups.

Below there are two more examples on how to report descriptive statistics for
EQ-5D values. Table 4.4 shows a comprehensive overview of EQ-5D-3L population
norm values for the United States (US), stratified by age and sex and also including
total values. The precision of the estimate of the mean is indicated by the standard
error. The median (50th percentile) is included, being relevant in general population
samples that tend to be skewed towards full health, and a measure of dispersion is
represented by the interquartile range (75th percentile–25th percentile).

An illustrative way to present longitudinal values from different populations is
shown in Fig. 4.4 by a scatter plot for the experimental and comparator arms in an
intervention design. One can simply track the value means for both patient groups
over time, indicating the new treatment causes a more severe drop in health initially
but also displays a quicker recovery and finally leads to a higher level of health than
the comparator treatment.

**Table 4.3** EQ-5D values for the total patient population and the 3 subgroups

| EQ-5D value | All patients | Subgroup 1 | Subgroup 2 | Subgroup 3 |
|---|---|---|---|---|
| Mean | 0.660 | 0.450 | 0.550 | 0.900 |
| Standard error | 0.010 | 0.013 | 0.015 | 0.010 |
| Median | 0.550 | 0.400 | 0.550 | 0.950 |
| 25th | 0.500 | 0.300 | 0.500 | 0.800 |
| 75th | 0.700 | 0.500 | 0.600 | 1.000 |
| N | 300 | 100 | 75 | 125 |

**Fig. 4.3** Mean EQ-5D values and 95% confidence intervals for the total patient population and 3 subgroups
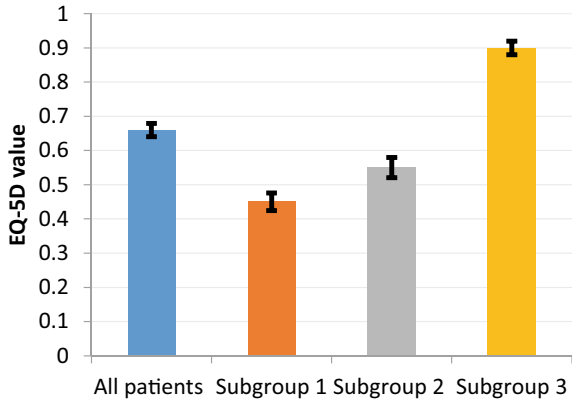


**Table 4.4** General population EQ-5D-3L norm values for a representative sample of the US (Szende et al. 2014, Springer open access)

| EQ-5D value | | Age | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (TTO value set) | | 18–24 | 25–34 | 35–44 | 45–54 | 55–64 | 65–74 | 75+ | Total |
| Total | Mean | 0.925 | 0.912 | 0.888 | 0.855 | 0.827 | 0.813 | 0.754 | 0.866 |
| | Standard error | 0.002 | 0.002 | 0.002 | 0.002 | 0.003 | 0.003 | 0.004 | 0.001 |
| | 25th Percentile | 0.83 | 0.83 | 0.83 | 0.80 | 0.78 | 0.78 | 0.71 | 0.80 |
| | 50th Percentile (median) | 1.00 | 1.00 | 1.00 | 0.83 | 0.83 | 0.83 | 0.80 | 0.84 |
| | 75th Percentile | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.83 | 1.00 |
| Males | Mean | 0.935 | 0.921 | 0.900 | 0.864 | 0.842 | 0.825 | 0.773 | 0.880 |
| | Standard error | 0.003 | 0.003 | 0.003 | 0.003 | 0.004 | 0.005 | 0.007 | 0.001 |
| | 25th Percentile | 0.84 | 0.83 | 0.83 | 0.81 | 0.80 | 0.78 | 0.71 | 0.82 |
| | 50th Percentile (median) | 1.00 | 1.00 | 1.00 | 0.84 | 0.83 | 0.83 | 0.81 | 1.00 |
| | 75th Percentile | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.84 | 1.00 |
| Females | Mean | 0.914 | 0.904 | 0.877 | 0.846 | 0.812 | 0.803 | 0.741 | 0.854 |
| | Standard error | 0.003 | 0.003 | 0.003 | 0.003 | 0.004 | 0.005 | 0.005 | 0.001 |
| | 25th Percentile | 0.83 | 0.83 | 0.81 | 0.80 | 0.78 | 0.77 | 0.71 | 0.80 |
| | 50th Percentile (median) | 1.00 | 1.00 | 0.84 | 0.83 | 0.83 | 0.82 | 0.78 | 0.84 |
| | 75th Percentile | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.86 | 0.83 | 1.00 |

It is important to note that EQ-5D values are often not symmetrically distributed, and tend to be divided into multiple groups (clusters), which might mean that standard statistics such as means and standard deviations are harder to interpret. This will be discussed in more detail in Sects. 4.4 and 4.6.
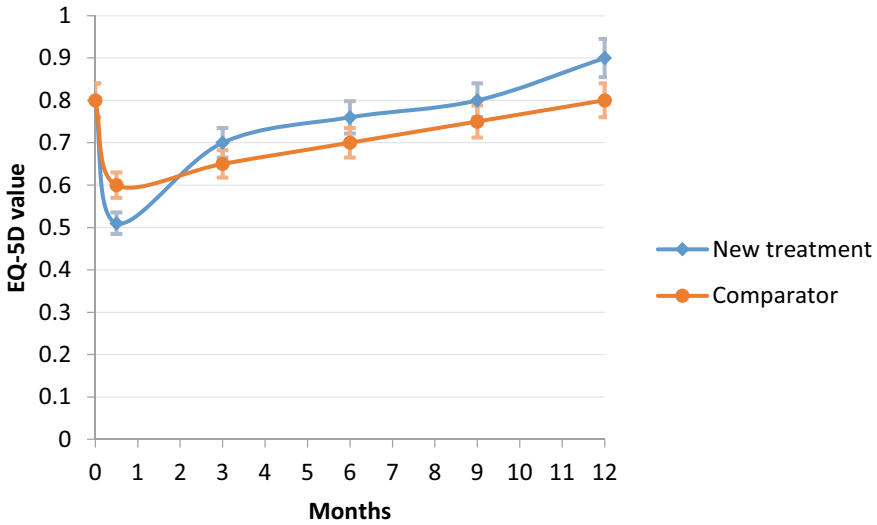
**Fig. 4.4** Example of presentation of longitudinal EQ-5D values (hypothetical data with smoothed lines and confidence intervals)

EQ-5D values are often used to calculate QALYs, for use in CUA. Although QALYs are commonly used in an evaluative context, for example when comparing two or more health programmes. An example is shown here to calculate QALYs for descriptive purposes, e.g. for a single individual. In the standard QALY model, values are simply multiplied by the time period for the corresponding health state, and when different health states occur over time, these are added, as shown in Fig. 4.5, where two health states in suboptimal health occur with values ('utilities') of 0.4 and 0.8 after which health gradually improves after the initial event.



$$\text{QALYs} = (6\times1 + (2\times0.4 + 2\times0.4\times0.5) + (4\times0.8 + 4\times0.2\times0.5))/12 = \mathbf{0.90}$$
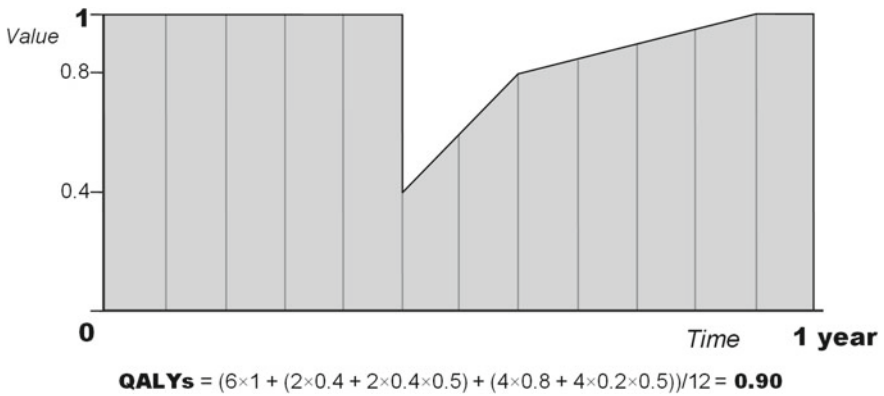
**Fig. 4.5** QALY calculation of an event-like condition with a recovery period

## 4.4   Examining the Distribution of the EQ-5D Values

Examining EQ-5D value distributions can be done in a graphical as well as in a numerical manner. First, we present graphical ways of exploring distributions. Distributions of EQ-5D values often show gaps and spikes or clusters of observations in certain parts of the scale. At the upper part of the scale there is often a gap which can be quite substantial, especially in EQ-5D-3L value distributions. This gap is caused by the ceiling often present in EQ-5D data and the intercept in the value function. In general population samples, but also in mildly or moderately diseased samples, often a relatively large proportion of respondents score no problems on all five dimensions: the ceiling. A large ceiling will result in a skewed distribution. For many value sets, there is a relatively large constant (or intercept) in the value set, leading to a gap between full health and the second-best health state. In distributional terms this may result in at least two clusters in the distribution. Apart from this "upper gap", more gaps may appear in EQ-5D value distributions. Parkin et al. (2016) demonstrated that two or three clusters often occur in value distributions for EQ-5D-3L. The left panel in Fig. 4.6 shows an example with 3 clusters caused by the ceiling and the intercept (the upper gap) and a low and high cluster which are caused by differences between levels 2 and 3 value decrements being greater than those between levels 1 and 2, and also because of the so-called N3 term[4] used in the many EQ-5D-3L value sets, as shown by Parkin et al. The right panel in Fig. 4.6 shows a distribution of EQ-5D-5L in the same patient group, resulting in a much smoother distribution. Note that these data were derived from a single patient sample: these respondents scored both the EQ-5D-3L and EQ-5D-5L descriptive systems, and subsequently the corresponding value sets (UK for EQ-5D-3L and English for EQ-5D-5L) were applied to the health profile data.

There are several differences between value sets across countries, but overall it was shown that EQ-5D-5L distributions resulted in smoother and more natural
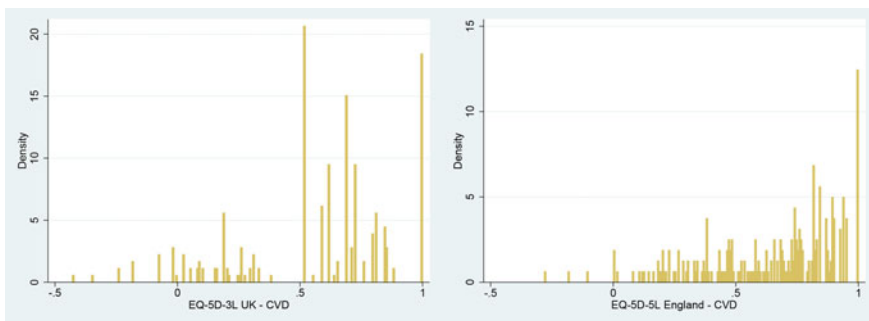


**Fig. 4.6**   Distribution of EQ-5D-3L and EQ-5D-5L values in a sample of cardiovascular disease (CVD) patients (N = 251)

---

[4]The N3 term results in an additional decrement of the value when at least one level 3 is present in the health profile.
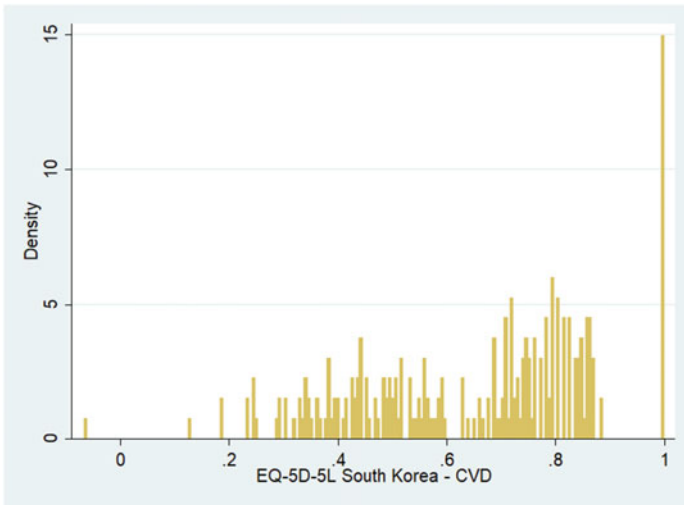
**Fig. 4.7**  Distribution of EQ-5D-5L values in a sample of cardiovascular disease patients (N = 251)

looking distributions than EQ-5D-3L (Janssen et al. 2018). Interestingly, an exception occurred for an EQ-5D-5L value set including a model term similar to N3. Here again three clusters appear in the distribution, as depicted in Fig. 4.7.

Sometimes histograms of distributions are not easy to assess, especially with large datasets in heterogeneous populations, e.g. showing a large spread of observations and perhaps spikes or clusters across the value scale. It becomes even more difficult when you want to compare two distributions in a single figure. In these cases, it might help to use a smoothing function such as the kernel density estimation. Figure 4.8 shows an example of an EQ-5D-3L and EQ-5D-5L kernel density plot in a large heterogeneous dataset. Note that also here the EQ-5D-5L distribution resulted in a much smoother plot when compared to the EQ-5D-3L distribution plot which is much more irregular.

When depicting a single distribution one can also combine a histogram with a smoothing function, such as shown in Fig. 4.9.

A final comment in regard to graphical presentation by histograms is that the choice of number of interval ranges ("bins", each bar represents 1 interval range) might influence the density in areas with a high concentration of observations, e.g. the ceiling (proportion of 11111) will result in a larger spike when more bins are opted for. Figure 4.10 shows an example for an EQ-5D-5L value distribution in a pooled dataset of 9 condition groups with 35 bins in the left panel versus 100 bins in the right panel.

Many EQ-5D-3L value set distributions will result in a distribution with clusters and gaps. These patterns in the distribution are considered to be undesirable as they can diminish the sensitivity and accuracy of the instrument (Janssen et al. 2018). Moreover, they can lead to estimation problems if distributions result in a
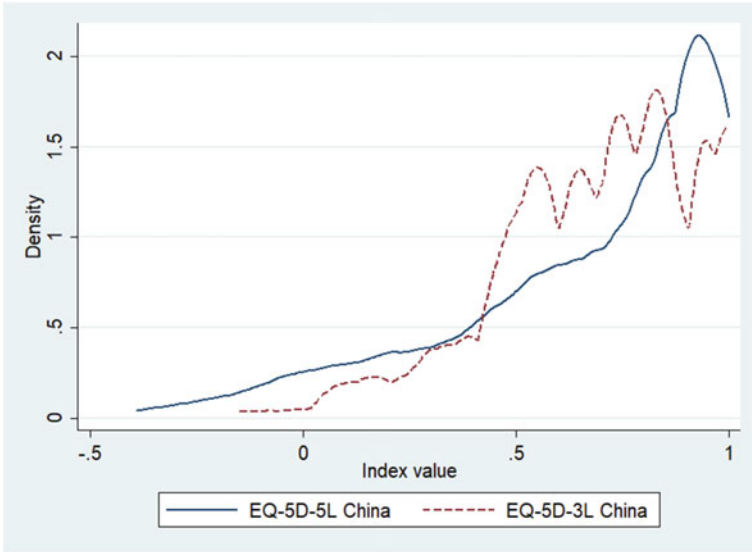
**Fig. 4.8** Distribution of EQ-5D-3L and EQ-5D-5L values in a pooled dataset of 9 condition groups (N = 3,790)
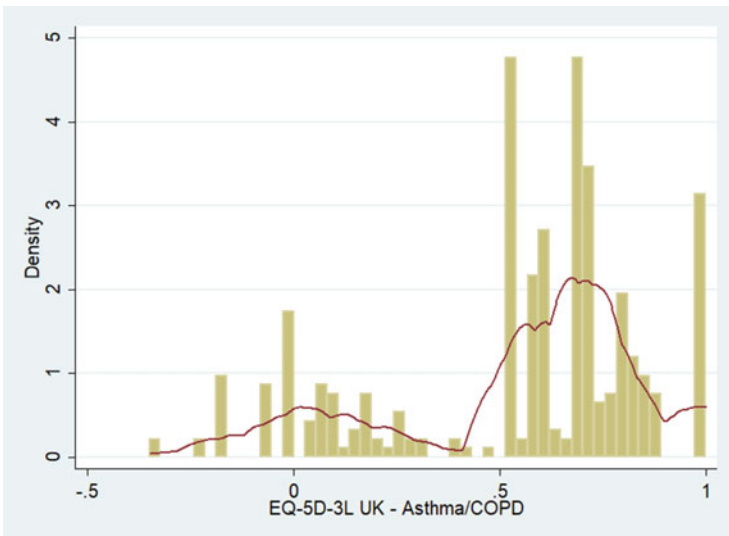


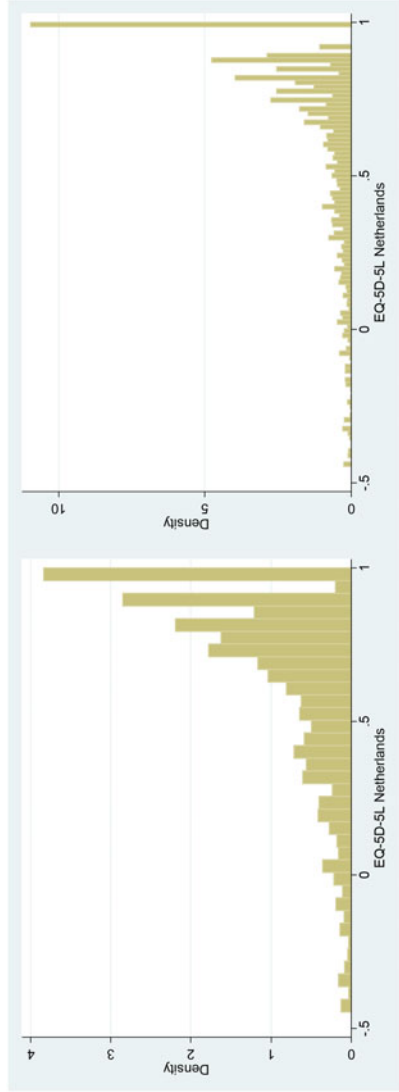**Fig. 4.9** Distribution of EQ-5D-3L values in a sample of Asthma/COPD patients (N = 342)

**Fig. 4.10** Distribution of EQ-5D-5L values applying 35 versus 100 bins

violation of homoscedasticity[5] when the values are used as dependent variable in regression analysis. With the introduction of the EQ-5D-5L the clusters and gaps largely disappeared, although to a lesser extent they still might occur. An extreme example is shown in Fig. 4.7 where clusters were caused by the large intercept and the interaction term in the value function. For other EQ-5D-5L country-specific value sets clusters and gaps hardly occur. Overall the interim ('mapped') EQ-5D-5L value distributions tend to be more similar in shape to the EQ-5D-5L value set distributions, although the range is identical to the EQ-5D-3L distributions (Feng et al. 2019; Mulhern et al. 2018). In Sects. 4.6 and 4.7 guidance is provided on how to deal with a clustered data distribution.

A final remark can be made in regard to the terms bimodal and even trimodal that are often used to describe distributions with 2 or 3 clusters respectively. Parkin et al. (2016) point out that in regard to EQ-5D-3L data these terms are misleading, since the modes of the groups are not their most interesting feature. The groups do not always have a single local mode, and in practice these modes are never actually identified, reported, or analysed.

There are several numerical ways of assessing EQ-5D value distributions. A simple way is to report the proportion of the ceiling and the floor. More comprehensive methods are evenness measures such as the Shannon indices, or the Health State Density Index as described in Sect. 2.8. Note that the total number of unique *values* might be (almost) equal to the number of unique possible *health profiles*. In these cases, the resulting indices will be equal to or close to the indices applied to the profile data.

## 4.5   Variance and Heteroskedasticity

As described above, EQ-5D value data is often defined by some specific characteristics. By nature, the data are censored due to the upper bound at 1 (full health) and the lower bound for the most severe health profile (33333 in 3L and 55555 in 5L). Because 11111 describes full or "normal" health as indicated by having no problems across the five dimensions, there often is a ceiling present and the data distribution might be skewed. A consequence of these factors is that variances might vary across the value space, leading to heteroskedasticity. Heteroskedasticity refers to the situation where the variance of a variable is unequal across the range of values of a second variable that describes or predicts it. Figure 4.11 shows an example of observed values paired with self-reported EQ VAS ratings. Typically EQ-5D variances will be unequal across the scale, which is at least partly due to the censored nature of the value scale (e.g. the figure clearly shows reduced variance in the upper right corner of the Fig. 4.11).

A graphical way of depicting heteroskedasticity (or homoscedasticity) is by using a residual-versus-predictor plot, which is a scatter plot of residuals against the
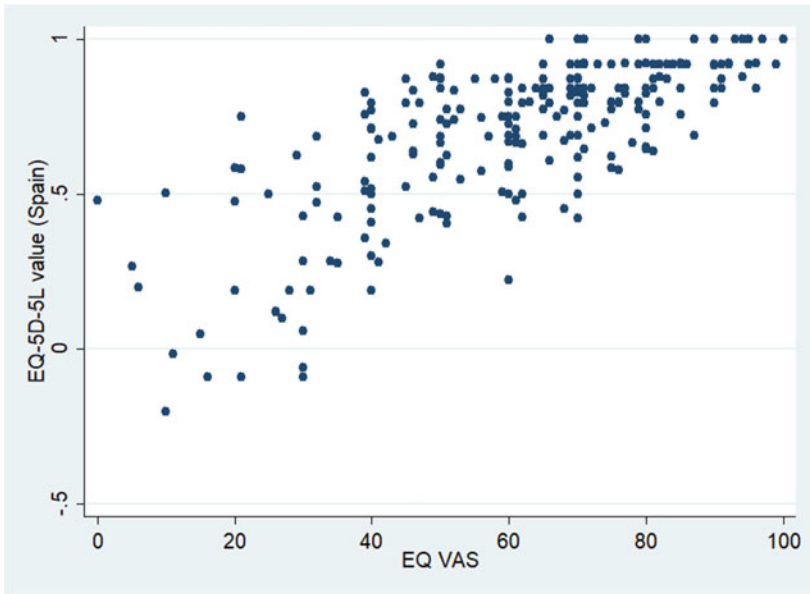
---

[5]See also Sect. 4.6.

**Fig. 4.11** EQ-5D-5L values (Spanish value set) plotted against EQ VAS for a sample of personality disorder patients (n = 384)

predicted values. One can easily detect if there are any patterns visible in the scatter plot. If there are no visible patterns or the plot shows roughly a rectangular shape, or both, the data are likely to be homoscedastic. Note that a pattern could be present in a residual-versus-predictor plot but the data could still be homoscedastic, in which case the data is likely to be biased. Figure 4.12 shows an example of a residual-versus-predictor plot for the same data used in Fig. 4.11. Clearly residuals are distributed unequally across the value scale which means that heteroskedasticity is present.

There have been many reported cases of heteroscedasticity in EQ-5D data. Section 4.7 provides further information on how to deal with heteroscedasticity in EQ-5D data.

## 4.6  Exploring Clusters in EQ-5D Value Distributions

As described in Sect. 4.4, EQ-5D value distributions often show clusters of observations. Sometimes these can be clearly detected graphically as is the case for many EQ-5D-3L distributions. In other cases, one can use statistical methods to test for the presence of clusters. A distribution with multiple clusters might imply that there are actually multiple patient populations that should be analysed separately. The mean value might actually refer to a point on the value scale were there are hardly any observations, so perhaps a better way to inform about these data would be to report
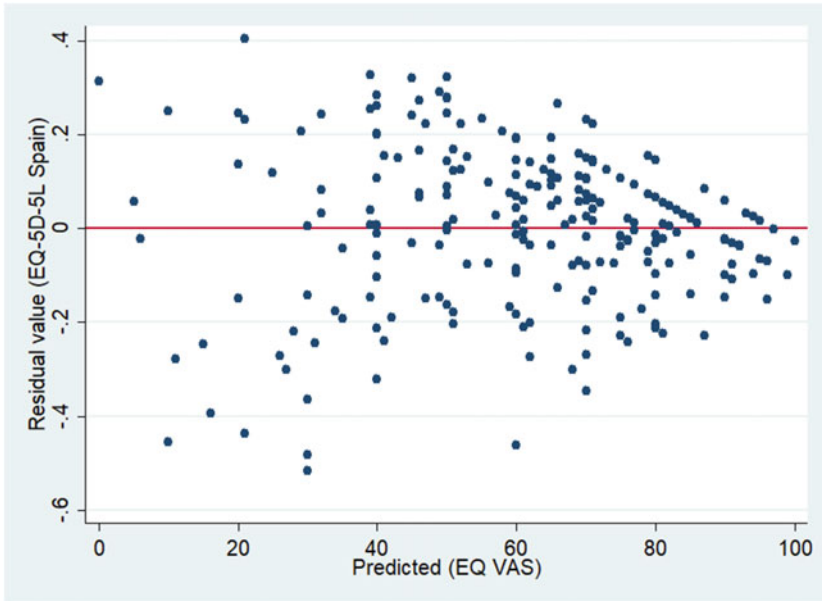
**Fig. 4.12** Residual-versus-predictor plot of EQ-5D-5L residual values (Spanish value set) plotted against EQ VAS (ordinary least squares regression) in a sample of personality disorder patients (n = 384)

simple descriptive statistics such as the mean, median, mode, range and standard deviations for the clusters *separately.*

Parkin et al. (2016) and Feng et al. (2019) used statistical techniques to detect clusters, first by applying k-means clustering to demonstrate the presence of clusters in EQ-5D-3L and EQ-5D-5L distributions. The k-means cluster algorithm searches for the optimal partition in *k* clusters. There are many stopping rules available for determining the optimal number of clusters. Feng et al. identified the Calinski-Harabasz pseudo-F index as the most suitable stopping rule for EQ-5D value data. Before applying the k-means procedure, the number of clusters must be decided upon. Subsequently the stopping rule may be applied to determine the optimal number of clusters. For more detail see Feng et al. (2019). Table 4.5 shows an example of applying this method on the EQ-5D-5L value set for England in a large pooled dataset across 2 patient groups. There are different clusters apparent, with different mean values and different dispersion and shape statistics. Note that different clusters are found for the different patient groups, and also the optimal number of clusters varies over the different patient groups.

Although this approach can be used as a useful exploratory tool, it does involve arbitrary judgments. Therefore, a careful examination of the data and resulting cluster statistics is advised before making conclusions in regard to what the optimal clusters are, if any. Testing for clusters and identifying clusters can be useful before using the data for different applications, such as health technology assessment and health

**Table 4.5** Identifying clusters in EQ-5D-5L data (English value set) in 2 patient groups (Feng et al. 2019)

| Cluster | Musculoskeletal Patients | | | Specialist nursing patients | | | | |
| | K = 2 | | | K = 4 | | | | |
| | 1 | 2 | Total (unclustered) | 1 | 2 | 3 | 4 | Total (unclustered) |
|---|---|---|---|---|---|---|---|---|
| N | 5051 | 14551 | 19602 | 253 | 571 | 916 | 1182 | 2922 |
| Min | −0.285 | 0.555 | −0.285 | −0.285 | 0.246 | 0.535 | 0.770 | 0.285 |
| Max | 0.553 | 0.950 | 0.950 | 0.243 | 0.531 | 0.766 | 0.950 | 0.950 |
| Mean | 0.329 | 0.781 | 0.664 | 0.092 | 0.396 | 0.669 | 0.867 | 0.645 |
| Median | 0.374 | 0.795 | 0.732 | 0.119 | 0.399 | 0.674 | 0.864 | 0.715 |
| SD | 0.178 | 0.099 | 0.233 | 0.123 | 0.082 | 0.063 | 0.055 | 0.252 |
| Skewness | −0.925 | −0.297 | −1.235 | −0.920 | −0.100 | −0.353 | −0.082 | −0.982 |
| Kurtosis | 3.255 | 2.234 | 4.096 | 2.941 | 1.726 | 2.079 | 1.699 | 3.324 |
| Range | 0.838 | 0.395 | 1.235 | 0.528 | 0.285 | 0.231 | 0.180 | 1.235 |

care management processes. The statistical techniques one intends to use should take account of clustering, in order to ensure that inferences drawn from the results are not biased.

An exploratory potential use of cluster analysis is to provide a means of identifying distinct pre-and post-treatment patient groups, and to use that information to predict which patients might benefit the most from the treatment and for which the treatment is less successful.

## 4.7   Regression Analysis

Regression analysis is a commonly used statistical technique for analysing EQ-5D values, quantifying the influence on values of their underlying determinants, such as clinical and socioeconomic characteristics. Applying multivariate regression enables multivariate comparisons, similar to the analysis of EQ VAS scores, as described in Sect. 3.3. The main uses are within economic evaluation, where the interest is in the values generated by different health care interventions, and in mapping studies, where the interest is in the values attached to different health states.

In Table 4.6 an example is shown of applying regression techniques for economic modelling for a treatment for relapsed or refractory multiple myeloma (NICE 2017). EQ-5D-3L data (UK value set) resulting from a randomized controlled trial were modelled by regression analysis for use in CUA. A repeated measurement mixed model was used to predict EQ-5D-3L values based on three types of response, whether a patient was ≤3 months prior to death, hospitalizations, (treatment related) adverse events, and new primary malignancies. The occurrence of adverse events and

**Table 4.6** Utility coefficients for parameters obtained using the EQ-5D-3L (UK value set)[a]

| Parameter | Coefficient | Standard error | 95% Confidence Limits | 95% Confidence limits | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|
| Intercept | −1.245 | 0.038 | −1.319 | −1.170 | −32.950 | <0.001 |
| PD | 0.182 | 0.054 | 0.077 | 0.287 | 3.400 | 0.001 |
| PR | 0.122 | 0.056 | 0.012 | 0.232 | 2.180 | 0.029 |
| SD | 0.187 | 0.061 | 0.068 | 0.305 | 3.080 | 0.002 |
| Hospitalisation | 0.219 | 0.203 | −0.178 | 0.617 | 1.080 | 0.279 |
| Grade 3 or 4 TRAE | 0.055 | 0.036 | −0.016 | 0.127 | 1.52 | 0.13 |
| New Primary Malignancy | 0.713 | 0.052 | 0.611 | 0.815 | 13.70 | <0.0001 |
| EOL 0-3 months pre-death | 0.378 | 0.081 | 0.219 | 0.537 | 4.65 | <0.0001 |

Key: EOL, end of life; PD, progressed disease; PR, partial response; SD, stable disease; TRAE, treatment related adverse events

[a]EQ-5D-3L data were transposed into a utility decrement using "decrement = 1–utility". The decrements were used as dependent variables in the regression model with response status, hospitalisation, adverse events, new primary malignancy, whether a patient is within 3 months prior to death, treatment allocation and time as independent variables, with interactions between time and response status

hospitalisation were included as covariates. The model used a log link and a Gamma distribution. The results from this regression showed that new primary malignancies and whether a patient is ≤3 months prior to death had the largest effects on utility. Variables associated with response status also had a significant impact. The coefficients associated with adverse events and hospitalisations were not significant. The utility coefficients can be used for the calculation of QALYs for inclusion in a CUA model.

As we have seen in Sect. 4.4, EQ-5D data is characterized by its censored nature with bounds at full health and the worst health state. Moreover, for many country-specific EQ-5D value sets, there is a gap between full health and the second best health state. For EQ-5D-3L, there are often clusters present, which only occurs for certain country-specific value sets for EQ-5D-5L data. Given this specific nature of EQ-5D values, many different regression techniques have been applied.

Ordinary Least Squares (OLS) regression is the most commonly used regression technique. As always, it is necessary to test for violations of its underlying assumptions, although it is robust to small violations, especially in large samples. These include the assumption that the residuals are normally distributed[6] and homoscedastic, violations of which affect statistical testing of regression coefficients,

---

[6]Note that the data itself do not need to be normally distributed due to the Central Limit Theorem. The distribution of the means of non-normal distributions will still be normal as long as the samples are large enough, large being roughly above 30 (Norman and Streiner 2000, p. 28).

though not the estimates themselves. However, EQ-5D values data may be subject to clustering, which violates the assumption that all of the observations in the data are independent, and censoring, which could affect the consistency of the OLS estimator, generating estimates that may be biased.

Various statistical tests are available to verify which regression techniques are most suitable for a given dataset. Normality of residuals can be assessed by several formal tests, including skewness and kurtosis estimates, the Shapiro-Wilk test, or the Jarque-Bera test. When using EQ-5D data in regression analysis, it is recommended to test for heteroskedasticity, for which many formal tests are available, such as the Breusch-Pagan test or the White test. When comparing two or more groups one also has to take the possibility of unequal variances into account. Again, it is recommended to test for unequal variances, e.g. by using the F-test of equality of variances. Note that the assumption of homoscedasticity is related to the residuals and not the dependent and independent variables included in a regression itself. Graphical and numerical approaches as described in Sects. 4.4 and 4.6 can be applied to test for the presence of clusters. Based on these results, one can determine which regression technique is most suited for the analysis of interest.

Many regression modelling techniques are available to deal with the typical nature of EQ-5D value data, such as Tobit, censored least absolute deviation (CLAD) or other median models, two-part models, latent-class models, and limited dependent variable mixture models[7] (Austin 2002; Fu and Kattan 2006; Huang et al. 2008; Pullenayegum et al. 2010; Hernández Alava et al. 2012). These different models aim to deal with various characteristics of the data. Tobit and CLAD models can take account of the censored nature of EQ-5D data. Two-part models specifically take the ceiling effect and the upper gap into account. Pullenayegum et al. (2010) suggested that Tobit and CLAD models might lead to biased results and propose OLS coupled with robust standard errors or the nonparametric bootstrap as a simpler and more valid approach which corrects for heteroskedasticity. Hernández Alava et al. (2012) demonstrated that an adjusted limited dependent variable approach combined with a mixture model can also account for the typical nature of EQ-5D-3L data (censored, large upper gap, and clustering). Figure 4.13 shows how the various models relate to different distributions, and we can indeed see that the adjusted limited dependent variable mixture model might be a good fit for various EQ-5D-3L distributions. For EQ-5D-5L, less complex models might suffice.

The mixture model approach applied by Hernández Alava et al. can also be used to identify latent classes, which bears a resemblance to identifying clusters as described in Sect. 4.6. A latent class model might be applied in regression to account for the different classes or clusters.

We end this section by providing an example. An innovative technique to develop a "catalogue" of EQ-5D-3L values by applying regression techniques to a large representative population survey database collected in the US was introduced by Sullivan and Ghushchyan (2006). CLAD regression was used to estimate the marginal

---

[7]Note that Hernández Alava et al. (2012) use a wider term (limited dependent variable) for EQ-5D data being censored at 1.
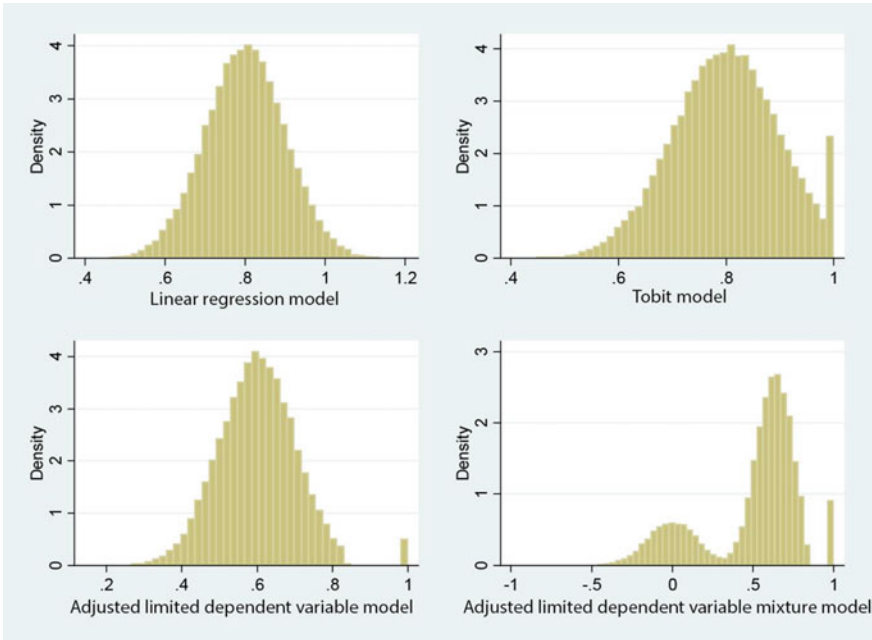
**Fig. 4.13**   Illustrative histograms of possible model distributions (Hernández Alava, copyright Value in Health)

disutility of conditions classified by International Classification of Diseases codes (Ninth Revision), controlling for age, comorbidity, gender, race, ethnicity, income, and education. The resulting list of EQ-5D-3L values could serve as an "off-the-shelf" catalogue that might be used by analysts to estimate QALYs in CUA.

## 4.8   Uncertainty and Sensitivity Analysis

As mentioned in Sect. 4.2, since there may not be a single 'best' value set for any given application, it is recommended to perform sensitivity analysis using other suitable value sets in order to assess the impact of the choice of value set on results and conclusions. Parkin et al. (2010) showed that the choice of value set might determine whether an intervention is seen as effective or not. Since many countries do not have a value set of their own, the choice of value set as well as performing sensitivity analysis, is very important. The analyst conducting CUA should treat the values in an economic evaluation as uncertain parameters which, just like other non-stochastic uncertain variables such as the discount rate, should be subject to sensitivity analysis, in order to improve confidence in the obtained results.

It must be noted that the magnitude of differences between value sets, and their implications for estimates of QALYs, is not always obvious. As one value set might contain values that are systematically higher (or lower) than another for the health states relevant to a given therapy, these differences may even out in economic evaluation, which focuses on the incremental change in health resulting from that therapy.

Due to the preference structure of a certain country-specific value set, an intervention might be considered effective in one country and not effective in another country, based on identical EQ-5D data resulting from clinical trials. In applications where the societal perspective is not relevant, but the values are used as a convenient way of summarizing the EQ-5D descriptive data, one has to be even more careful, since the influence on any given country-specific value set might bias the results. Sensitivity analysis will also give the researcher a sense of how stable the results are and whether robust conclusions might be drawn.

# References

Appleby J, Devlin N, Parkin D (2015) Using Patient Reported Outcomes to improve health care. Wiley Blackwell

Austin PC (2002) A comparison of methods for analyzing health-related quality of life measures. Value Health 5:329–37

Bansback N, Brazier J, Tsuchiya A, Anis A (2012) Using a discrete choice experiment to estimate health state utility values. J Health Econ 31(1):306–318

Brazier J, Deverill M, Green C (1999) A review of the use of health status measures in economic evaluation. J Health Serv Res Policy 4(3):174–184

Brazier J, Akehurst, R, Brennan, A Dolan P, Claxton K, McCabe C, Sculpher M, Tsuchyia A (2005) Should patients have a greater role in valuing health states? Appl Health Econ Health Policy 4(4):201–208

Craig BM, Busschbach JJ, Salomon JA (2009) Keep it simple: ranking health states yields values similar to cardinal measurement approaches. J Clin Epidemiol 62(3):296–305

Devlin N, Parkin D (2007) Guidance to users of EQ-5D value sets. In: Szende A, Oppe M, Devlin N (eds) EQ-5D value sets: inventory, comparative review and user guide. Springer, Dordrecht, The Netherlands

de Wit GA, Busschbach JJ, de Charro FTH (2000) Sensitivity and perspective in the valuation of the health status: whose values count? Health Econ 9:109–126 www.euroqol.org, Accessed June 2nd 2019

Devlin NJ, Krabbe PF (2013) The development of new research methods for the valuation of EQ-5D-5L. Eur J Health Econ 14(Suppl 1):S1–S3

Dolan P (1997) Modeling valuations for EuroQol health states. Med Care 35(11):1095–1108

Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW (2015) Methods for the economic evaluation of health care programmes, 4th edn. Oxford University Press, Oxford

Feng Y, Devlin N, Bateman A, Zamora B, Parkin D. (2019) Distribution of the EQ-5D-5L Profiles and Values in Three Patient Groups. Value Health 22(3):355–361

Feng Y, Devlin NJ, Shah KK, Mulhern B, van Hout B (2018) New methods for modelling EQ-5D-5L value sets: an application to English data. Health Econ 27(1):23–38

Fu AZ, Kattan MW (2006) Racial and ethnic differences in preference-based health status measure. Curr Med Res Opin 22(12):2439–2448

Greiner W, Weijnen T, Nieuwenhuizen M, Oppe S, Badia X, Busschbach J, Buxton M, Dolan P, Kind P, Krabbe P, Ohinmaa A, Parkin D, Roset M, Sintonen H, Tsuchiya A, de Charro F (2003) A single European currency for EQ-5D health states. results from a six-country study. Eur J Health Econ 4(3):222–31

Hernández Alava M, Wailoo AJ, Ara R (2012) Tails from the peak district: adjusted limited dependent variable mixture models of EQ-5D questionnaire health state utility values. Value Health 15(3):550–561

Huang IC, Frangakis C, Atkinson MJ et al (2008) Addressing ceiling effects in health status measures: a comparison of techniques applied to measures for people with HIV disease. Health Serv Res 43(1 Pt 1):327–339

Janssen MF, Bonsel GJ, 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. Pharmacoecononmics 36(6):675–697

Kennedy-Martin M, Slaap B, Herdman M, van Reenen M et al (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 doi: 10.1007/s10198-020-01195-8 Online ahead of print

Kind P (2003) Guidelines for value sets in economic and non-economic studies using EQ-5D. In: Brooks R, Rabin R, de Charro F (eds) The measurement and valuation of health status using EQ-5D: a European perspective. Kluwer Academic Publisher, Dordrecht, pp 29–41

McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed) Frontiers in Econometrics. Academic Press, New York

Mulhern B, Feng Y, Shah K, Janssen MF, Herdman M, van Hout B, Devlin N (2018) Comparing the UK EQ-5D-3L and English EQ-5D-5L value sets. Pharmacoecononmics 6(6):699–713

NICE 2017. https://www.nice.org.uk/guidance/ta505/documents/committee-papers

Norman GR, Streiner DL (2000) Biostatistics: the bare essentials, 2nd edn. BC Decker Incorporated, Hamilton

Ogorevc M, Murovec N, Fernandez NB, Rupel VP (2019) Questioning the differences between general public vs. patient based preferences towards EQ-5D-5L defined hypothetical health states. Health Policy 123(2):166–172

Oppe M, Devlin NJ, van Hout B, Krabbe PF, de Charro F (2014) A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. Value Health 17(4):445–453

Parkin D, Devlin N (2006) Is there a case for using visual analogue scale valuations in cost-utility analysis? Health Econ 15(7):653–664

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

Parkin D, Devlin N, Feng Y (2016) What determines the shape of an EQ-5D index distribution? Med Decision Making 36(8):941–951

Pullenayegum EM, Tarride J, Xie F, Goeree R, Gerstein HC, O'Reilly D (2010) Analysis of health utility data when some subjects attain the upper bound of 1: are Tobit and CLAD models appropriate? Value Health 13:487–494

Rabin R, de Charro F, Szende A (2007) Introduction. In: Szende A, Oppe M, Devlin N (eds) EQ-5D value sets: inventory, comparative review and user guide. Springer, Dordrecht, The Netherlands

Ramos-Goñi JM, Oppe M, Slaap B, Busschbach JJ, Stolk E (2017a) Quality control process for EQ-5D-5L valuation studies. Value Health 20(3):466–473

Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabasés JM, Serrano-Aguilar P, Rivero-Arias O (2017b) Valuation and modeling of EQ-5D-5L health states using a hybrid approach. Med Care 55(7):e51–e58

Sanders GD, Neumann PJ, Basu A et al (2016) Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. JAMA 316(10):1093–1103

Sullivan PW, Ghushchyan V (2006) Preference-based EQ-5D index scores for chronic conditions in the United States. Med Decis Making 26(4):410–420

Szende A, Janssen MF, Cabases J (2014) Self-reported population health: an international perspective based on EQ-5D. Springer, Dordrecht

van Hout B, Janssen MF, Feng Y et al (2012) Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health 15:708–715

Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB (1996) Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. JAMA 276(15):1253–1258

Wilke CT, Pickard AS, Walton SM, Moock J, Kohlmann T, Lee TA (2010) Statistical implications of utility weighted and equally weighted HRQL measures: an empirical study. Health Econ 19(1):101–110

Williams A (2005) The EuroQol instrument. In: Kind P, Brooks R, Rabin R (eds) EQ-5D concepts and methods: a developmental history. Springer, Dordrecht

Zethraeus N, Johannesson M (1999) A comparison of patient and social tariff values derived from the time trade-off method. Health Econ 8:541–545