Chapter 3

Analysis of EQ VAS Data

The aims of this chapter are

- to explain what is measured by the EQ VAS, and how that affects analysis of EQ VAS data; and
- to demonstrate ways in which EQ VAS data can be analysed and reported.

3.1 Interpreting the EQ VAS

It is important when analysing EQ VAS data to understand the nature of this element of the EQ-5D questionnaire and the measurement properties that it has. (For a more detailed discussion, see Feng et al. 2014.) The EQ VAS has a unique design that does not conform to conventional Visual Analogue Scale (VAS) formats, and the widely-observed properties that VAS data have therefore do not automatically apply to EQ VAS data. It has some features that conventionally belong to a ‘rating scale’ but again, because it has an unconventional design, the properties that rating scale data have may also not apply.

A conventional VAS is a straight line of a specified length with verbal descriptors at each end stating the meaning attached to the end points, without any demarcations of the line or numeric labels at any point. The EQ VAS is also a line that has end-point descriptors, but it also demarcates the line in units of ones and tens, and places number labels on the tens markers. This format for the line is closer to a ‘numerical rating scale’, but such scales usually attach a number to every marker, have many fewer markers, and often do not have verbal end-point descriptors.

The versions of the EQ VAS contained in the EQ-5D-3L and EQ-5D-5L are also unconventional in how the scores are recorded by respondents. For the EQ-5D-3L version, the method of drawing a line from a box that states ‘Your own health state today’ to the scale (see Fig. 1.4, Chap. 1) is unique; it is a feature that was included

for reasons related to the historical development of the EQ-5D¹ rather than evidence about the best way for respondents to record EQ VAS scores (Feng et al. 2014). The EQ-5D-5L uses a more conventional means of recording the score on the line (by marking a cross) but asks respondents also to record the score separately as a number. The aim of this is to overcome problems of imprecise marking on the line, but this introduces the possibility that respondents may respond primarily to the direct numeric estimate and are therefore undertaking a different measurement task, ‘magnitude estimation’, data from which may also have different properties to VAS and rating scale data.

The measurement properties that result from this, and therefore the kinds of statistical analysis that are permitted, are therefore not entirely clear. It is reasonable to assume that the resulting scores are at least ordinal. The line’s design strongly suggests that the respondent is invited to supply interval level data, which is also the intention of magnitude estimation. The provision of a true zero and maximum even suggests providing scores that have ratio level properties. However, those who complete the VAS may in practice not respond to the visual stimuli provided in exactly this way. The evidence is mixed, with some studies finding reasonable interval scale properties; however, EQ VAS responses have very often been found to exhibit ‘end aversion’, which suggests that the data cannot be truly interval, though it is possible that a transformation could be estimated to repair this.

Another consideration is that, as with all health-related quality of life (HRQoL) measurement methods, EQ-VAS responses may not be interpersonally comparable. For example, the end-point labels may mean different things to different respondents, and the meaning that they attach to different numbers may also differ (Devlin et al. 2019).

The guidelines for analysis of EQ VAS data below assume that the numerical values given to the EQ VAS behave as if they have at least an interval scale and are interpersonally comparable, such that it is meaningful to calculate descriptive statistics for a sample or population, such as means; to apply hypothesis testing, such as t-tests of differences in means; and to use estimation procedures, such as regression analysis. However, if there is evidence to suggest that the EQ VAS data are ordinal, then non-parametric versions of the descriptive and inferential statistics described below should be used.

It is also the case that EQ VAS data often exhibit digit preference, which is a tendency to choose numbers ending with 0 and to a lesser extent 5, rather than any others. In the context of sample or population data, this phenomenon may be treated as a lack of precision rather than the existence of bias.

Before beginning to analyse EQ VAS data that have been collected via paper questionnaires for the EQ-5D-3L, it is important to check how those data have been coded. Recall, from Chap. 1, that there are particular issues relating to the range of approaches which respondents have been observed to use in completing the EQ VAS

¹Specifically, the EQ VAS was initially included as a warm-up task in studies to obtain VAS valuations for EQ-5D health states, and the format of the EQ VAS reflects the VAS which was used in those valuation tasks.

in the EQ-5D-3L questionnaire. These may not strictly comply with the instructions but nevertheless represent valid responses. The EQ VAS in the EQ-5D-3L will, in future, be made consistent with the EQ VAS in the EQ-5D-5L, so this issue will no longer arise, but does apply to historic data sets.

3.2 Simple Descriptive Statistics and Inference

With respect to summary measures, the distribution of EQ VAS data within a sample can be reported using a full range of descriptive statistics, such as minimums, maximums, means, medians, quartiles, standard deviations, interquartile ranges, skewness and kurtosis. Descriptions of relationships between EQ VAS data and other variables can also be reported, such as correlation coefficients. These may be subject to appropriate hypothesis testing using, for example, a t-test to test for differences between means or for the significance of a correlation coefficient. Similarly, EQ VAS data can be used for estimation, either as a dependent or independent variable.

The following example uses publicly available data from the Patient Reported Outcome Measures (PROMs) programme of the English National Health Service (NHS) (Devlin et al. 2010). It shows data collected from 38,187 patients before and after they had hip replacement surgery in 2010–11. The following Table 3.1 shows a range of descriptive statistics.

Because the raw data are recorded as integers, it is important to ensure that the figures reported do not have spurious accuracy. In this table, the median and the mode (in this case, there is only one) retain their integer format. The median might be presented to one decimal place, reflecting the possible values that it could take, but because of digit preference a value ending in 0.5 will be rarely observed. Other statistics are presented to three significant figures, since it is unlikely that greater

Table 3.1 EQ VAS score for 38,187 patients before and after they had hip replacement surgery in the English NHS, 2010–11

EQ VAS score	Before surgery	After surgery
Mean	65.3	74.4
Standard error	0.116	0.102
Median	70	80
Mode	80	90
Standard deviation	21.7	19.34
Kurtosis	−0.110	1.37
Skewness	−0.672	−1.17
Minimum	0	0
Maximum	100	100
Range	100	100
Observations	34,716	35,762
Missing (percent)	3,471 (9.09%)	2,425 (6.35%)

Fig. 3.1 EQ VAS scores for hip replacement patients before surgery, English NHS 2010–11

precision than this is either necessary or justified. It is also good practice to report the number and percentage of missing values.

It is also informative to report the full distribution of individual EQ VAS data points, especially graphically. A table showing the frequency of observations taking values from the full range of possible scores is possible but may not be very informative about key features of the distribution and will be affected by the issue of digit preference. It is most useful to use a graphical display, particularly spike plots. An example is shown in Fig. 3.1, again using the before-surgery hip replacement data.

This plot not only shows the shape and central tendency of the distribution, but also the extent of digit preference.

It is possible to reduce frequency tables to categories containing ranges, which makes them more easily read. However, end points for ranges should be chosen carefully, as this may affect the visual appearance of the distribution. It is misleading, for example, to define ranges such as 0–4, 5–9, 10–14 etc., as observations such as 9 are more like 10 than 5, for example. It is better to define ranges such that they cover a midpoint, specifically multiples of 5 and 10. However, at the ends of the distribution it may be better to display individual scores for those below the range around 5 (0, 1 and 2) and above the range around 95 (98, 99 and 100) rather than assume they are all representative of 0 and 100. The following Table 3.2 and Fig. 3.2 show this procedure for the before-surgery hip replacement data.

Table 3.3 shows analyses of the similarity and differences between the two observations. In this example, the two EQ VAS scores (before and after surgery) are paired, but it would be possible to undertake similar analyses for unpaired data.

Table 3.2 EQ VAS scores for hip replacement patients before surgery, English NHS 2010–11

Range	Mid-point	Frequency	Range	Mid-point	Frequency
0	0	253	58–62	60	3161
1	1	15	63–67	65	1202
2	2	3	68–72	70	4438
3–7	5	87	73–77	75	2324
8–12	10	335	78–82	80	5060
13–17	15	151	83–87	85	1677
18–22	20	683	88–92	90	3694
23–27	25	444	93–97	95	1262
28–32	30	1458	98	98	113
33–37	35	611	99	99	77
38–42	40	1855	100	100	707
43–47	45	491	Total observed		34,716
48–52	50	3947	Missing		3,471
53–57	55	668	Total sample		38,187

Fig. 3.2 Mid-point EQ VAS scores for hip replacement patients before surgery, English NHS 2010–11

Table 3.3 EQ VAS scores for hip replacement patients before and after surgery, English NHS 2010–11

	EQ VAS score	
	Before surgery	After surgery
Mean	65.4	74.6
Standard Deviation	21.7	19.2
Observations	32 712	
Missing values	5 475 (14.3%)	
Difference in means		
t-statistic	−69.94	
p value (one- and two-tail)	<0.001	
Pearson Correlation	0.33	

3.3 Modelling Determinants of EQ VAS Scores

It will usually be of interest in analysing EQ VAS data to examine the impact of other variables on the EQ VAS scores. In the example above, a before-and-after comparison was made between scores obtained from the same people at two time periods, but similar comparisons could be made for people according to different characteristics such as age, gender, social circumstances, location etc. Obviously, multivariate comparisons could also be made. Multivariate regression techniques have been applied to EQ VAS data and demonstrated good discriminatory properties, for example Parkin et al. (2004).

An analysis that is always available to users of EQ-5D questionnaire data is to model the relationship between the EQ-5D health state profile and the EQ VAS scores. This makes good use of the full questionnaire data by giving additional insights into the nature of the HRQoL of respondents, highlighting the importance of different aspects of their HRQoL, as described by the profile, on their overall HRQoL, as measured by the EQ VAS. Studies using the EQ-5D-3L have demonstrated a good relationship between these. They have consistently found that coefficients on the levels and dimensions of the EQ-5D-3L health state profile are in the correct direction and follow the expected gradient between levels, such that the coefficients on level 3 are greater than those on level 2 (Jelsma and Ferguson 2004; Whynes 2008, 2013; Feng et al. 2014).

Table 3.4 shows an example from the PROMs hip data used earlier.

Amongst possible interpretations of these results, it is notable that Anxiety & Depression has the biggest impact on the EQ VAS scores and Pain & Discomfort the smallest. Although level 2 Mobility has an impact similar to that of level 2 Self-care, level 3 has a much greater impact, perhaps reflecting the extreme nature of the EQ-5D-3L level 3 descriptor for Mobility ('confined to bed'). Similarly, although level 2 Usual Activities has a much lower impact than level 2 Self-care, the level 3 coefficients for these two dimensions are similar.

It is important to note that the coefficients that will be obtained are specific to the characteristics of the population from which the data are collected. For patient

Table 3.4 regression analysis of EQ VAS score against EQ-5D levels for hip replacement patients before surgery, English NHS 2010–11

	Coefficient	Standard Error
Mobility level 2	−5.64	0.472
Mobility level 3	−14.9	1.85
Self-care level 2	−5.20	0.234
Self-care level 3	−8.25	1.12
Usual activities level 2	−2.99	0.462
Usual activities level 3	−7.76	0.537
Pain & Discomfort level 2	−0.131	0.728
Pain & Discomfort level 3	−5.04	0.743
Anxiety & Depression level 2	−8.61	0.233
Anxiety & Depression level 3	−16.3	0.536
Intercept	83.2	0.814

Number of observations = 34 446
 $R^2 = 0.166$, adjusted $R^2 = 0.166$, $F = 686$, $p < 0.001$
 All coefficients significantly different from 0 at the 0.001 level

data, the evidence is that the profile coefficients may differ according to the type of condition that the patient has. Moreover, other variables, including age and sex, may impact not only on the EQ VAS scores directly but also on the profile coefficients, via interaction effects. Such analyses add further to understanding the impact of different patient characteristics and conditions on HRQoL. An additional implication is that simple comparisons between the coefficients obtained from this analysis and those obtained by modelling valuation data should be avoided.

As a direct illustration of this, Fig. 3.3, which has been generated from a vast amount of EQ-5D data held by the EuroQol Group office in Rotterdam, shows that there is a sharply declining EQ VAS by age for those whose self-reported profile contains at least some problems (Oppe 2013). As age increases, the number and severity of problems reported increase and the EQ VAS decreases. But the EQ VAS declines with age even among patients reporting no problems on any EQ-5D dimension.

Fig. 3.3 The relationship between age and EQ VAS for those with no problems in any EQ-5D dimension and those with problems in at least one dimension (The straight lines are based on linear models for all data, no problems on EQ-5D (11111) and problems in at least one dimension (NOT 11111). The dashed straight lines represent separate linear models for men (♂) and women (♀) for all data. The other lines depict observed mean scores for corresponding groups.)

References

- Devlin N, Parkin D, Browne J (2010) Patient Reported Outcomes in the NHS: new methods for analysing and reporting EQ-5D data. *Health Econ* 19(8):886–905
- Devlin N, Lorgelly P, Herdman M (2019) Can we really compare and aggregate PRO data between people and settings? Implications for clinical trials and HTA. OHE Research Paper. www.ohe.org/publications/can-we-really-compare-and-aggregate-pro-data-between-people-and-settings-implications
- Feng Y, Parkin D, Devlin NJ (2014) Assessing the performance of the EQ VAS in the NHS PROMs Programme. *Qual Life Res* 23(3):977–989
- Jelsma J, Ferguson G (2004) The determinants of self-reported health-related quality of life in a culturally and socially diverse South African community. *Bull World Health Organ* 82(3):206–212
- Oppe M (2013) Mathematical approaches in economic evaluations: applying techniques from different disciplines. Erasmus University Rotterdam
- Parkin D, Rice N, Jacoby A, Doughty J (2004) Use of a visual analogue scale in a daily patient diary: modelling cross-sectional time series data on health-related quality of life. *Soc Sci Med* 59(2):351–360
- Whynes DK (2008) Correspondence between EQ-5D health state classifications and EQ VAS scores. *Health Qual Life Outcomes* 6:94
- Whynes DK (2013) Does the correspondence between EQ-5D health state description and VAS score vary by medical condition? *Health Qual Life Outcomes* 11:155

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.