TECHNICAL SUPPORT DOCUMENT 23:

A GUIDE TO CALCULATING SEVERITY SHORTFALL FOR NICE EVALUATIONS

REPORT BY THE DECISION SUPPORT UNIT

January 2024 (Updated March 2026)

Allan Wailoo,

Sheffield Centre for Health and Related Research (SCHARR), University of Sheffield, Sheffield, UK.

Decision Support Unit, ScHARR, University of Sheffield, Regent Court, 30 Regent Street Sheffield, S1 4DA Tel (+44) (0)114 222 0734 E-mail dsuadmin@sheffield.ac.uk

ABOUT THE DECISION SUPPORT UNIT

The Decision Support Unit (DSU) is based at the University of Sheffield with members at the universities of York, Bristol, Leicester, Warwick and the London School of Hygiene and Tropical Medicine. The DSU is commissioned by The National Institute for Health and Care Excellence (NICE) to provide a research and training resource to support the Institute's Centre for Health Technology Evaluation Programmes. Please see our website for further information: https://sheffield.ac.uk/nice-dsu

The production of this document was funded by the National Institute for Health and Care Excellence (NICE) through its Decision Support Unit. The views, and any errors or omissions, expressed in this document are of the authors only. NICE may take account of part or all of this document if it considers it appropriate, but it is not bound to do so.

ABOUT THE TECHNICAL SUPPORT DOCUMENT SERIES

NICE describes the methods it follows when carrying out health technology evaluations in its process and methods manual. This provides an overview of the key principles and methods of health technology assessment and appraisal for use in NICE appraisals. The manual does not provide detailed advice on how to implement and apply the methods it describes. The DSU series of Technical Support Documents (TSDs) is intended to complement the manual by providing detailed information on how to implement specific methods.

The TSDs provide a review of the current state of the art in selected topic areas. They make recommendations on the implementation of methods and reporting standards where it is appropriate to do so. They aim to provide assistance to all those involved in submitting or critiquing evidence as part of NICE technology evaluations, whether companies, assessment groups or any other stakeholder type.

We recognise that there are areas of uncertainty, controversy and rapid development. It is our intention that such areas are indicated in the TSDs. All TSDs are extensively peer reviewed prior to publication (the names of peer reviewers appear in the acknowledgements for each document). Nevertheless, the responsibility for each TSD

lies with the authors and we welcome any constructive feedback on the content or suggestions for further guides. The TSDs will be amended and updated whenever appropriate. Where minor updates or corrections are required, the TSD will retain its numbering with a note to indicate the date and content change of the last update. More substantial updates will be contained in new TSDs that entirely replace existing TSDs.

Please be aware that whilst the DSU is funded by NICE, these documents do not constitute formal NICE guidance or policy.

Prof Allan Wailoo, Director of DSU and TSD series editor.

ACKNOWLEDGEMENTS:

For extremely helpful comments on previous drafts, thanks are due to Paul Tappenden, Nick Latimer, Matt Stevenson, Ben Kearns, Carl Prescott and Lorna Dunning. Donna Davis provided expert assistance with formatting and presentation of this document.

This report should be referenced as follows:

Wailoo, A., A guide to calculating burden of illness for NICE Evaluations: Technical Support Document 23. (2024) Updated March 2026. [Available from http://www.nicedsu.org.uk]

EXECUTIVE SUMMARY

NICE has introduced Quality Adjusted Life Year (QALY) weights for health technologies for patient groups where the condition is considered to impose a high degree of severity. Severity is measured in terms of absolute and proportional shortfall (AS and PS). Both measure the difference between the number of QALYs patients would be expected to experience over the remainder of their lives under current care compared to the general population of the same age and sex. AS expresses this in terms of the number of QALYs lost. PS is a ratio of the number of QALYs lost compared to the total number of QALYs expected.

Consistency is an important aspect of NICE decision making. A “reference case” set of data sources and methods should be considered to ensure the assessment of AS and PS across appraisals meet this criterion. However, since part of the calculation of AS and PS relies on estimates submitted to NICE as part of existing cost effectiveness models (the number of QALYs for patients under current NHS care) the issue of consistency within an appraisal must also be considered in any recommendations. This report considers key sources of evidence for severity calculations.

Expectancy of life by age and sex should be estimated from life tables for England published by the ONS. Estimates should be weighted according to the proportions of males and females in the relevant patient population. The relevant population is those patients with the health condition(s), as defined in the appraisal for the patient subgroup of interest, at the same point in the disease management pathway where the technology under appraisal is being considered. At the time of writing, we recommend period life tables for 2017-2019 be used, due to uncertainty about the long-term impact of Covid-19 on the most recently released data for 2020.

NICE currently recommends that the value set for EQ-5D-3L be used in reference case economic analyses. We propose that the most suitable source for EQ-5D-3L health utilities by age and sex is based on data from the Health Survey for England (HSE), 2014, and modelled results reported in Hernández et al (2020). We show that using slightly more up-to-date survey data (HSE 2018) and mapping responses to the

EQ-5D-5L instrument to the EQ-5D-3L value set using NICE recommended methods, does not lead to substantially different estimates of health utility by age. Therefore, where the economic model is based on EQ-5D-5L mapped to EQ-5D-3L, the same 2014 directly observed EQ-5D-3L utility values can be used for the primary analysis.

A new EQ-5D-5L value set for the UK was published in March 2026 by Rowen et al. (2026). NICE has committed to use this value set, subject to the completion of a public consultation (see https://www.nice.org.uk/news/articles/changes-to-nice-s-cost-effectiveness-thresholds-confirmed). It is therefore appropriate to provide updated estimates of health utility by age to reflect this. The 2018 HSE provides the most appropriate source for directly observed EQ-5D-5L for this purpose. These data have been modelled by Hernández et al (2026).

Those aged under 16 years were not included in the HSE. Where estimates for children or adolescents are required, the results of the Hernández et al model should be extrapolated by applying the estimates at age 16 years for all lower ages.

Where other preference-based instruments have been used in the economic model, there are no obviously superior approaches. In some situations, either published mapping models to EQ-5D-3L/5L or the use of different population estimates of health utility by age could be used as additional sensitivity analyses. In many cases, such as where disease specific instruments have been used, these options will either not be available or appropriate. Committee judgement will be required in these non-reference case analyses.

Carer quality of life, if included in the estimated QALYs for current NHS care, should be excluded. This applies both to the calculation of AS/PS and the incremental QALYs that attract any non-unitary weight. The reference case discount rate of 3.5% should be used for severity calculations.

This guidance will not cover all scenarios for the assessment of AS and PS. Committee judgement will inevitably be required. Unforeseen complexities will arise. NICE should monitor the implementation of severity weights, identify those situations where challenges arise and support work to address those challenges including the updating

of this TSD where relevant. A dataset of all accepted AS and PS estimates that are accepted should be maintained and referred to as part of the appraisals process to ensure face validity across the programme. Pre-programmed software should be provided to implement AS and PS estimates easily.

CONTENTS

1. INTRODUCTION 8
2. ABSOLUTE AND PROPORTIONAL SHORTFALL 9
   2.1 Defining terms 9
   2.2 Data requirements 11
3. CALCULATION OF GENERAL POPULATION QALYS 12
   3.1 Expectation of life 12
   3.2 Health utility weights 15
   3.3 Using the EQ-5D-3L 16
   3.4 Using the EQ-5D-5L 16
   3.5 Utility weights for children 17
   3.6 Discounting 17
4. CALCULATION OF COMPARATOR QALYS 17
   4.1 Discounting 17
   4.2 Comparators 18
   4.3 Lifetime horizon 18
   4.4 Carer Quality of Life 19
   4.5 Source of utility values 19
5. WORKED EXAMPLE 22
   5.1 Using EQ-5D-3L 22
   5.2 Using EQ-5D-5L 23
6. SUMMARY 24

REFERENCES 26

APPENDIX 1: 27

TABLES

Table 1: QALY weightings for severity 8
Table 2: AS and PS for a hypothetical example (using EQ-5D-3L) 23
Table 3: AS and PS for a hypothetical example (using EQ-5D-5L) 24

FIGURES

Figure 1: Burden of Illness 10
Figure 2: Period versus cohort expectation of life years by age, for males in England 2020 14
Figure 3: EQ-5D-3L by age for females: HSE 2014 compared to HSE 2018 21
Figure 4: EQ-5D-3L by age for males: HSE 2014 compared to HSE 2018 21

1. INTRODUCTION

NICE published an updated set of processes and methods in January 2022[1] that apply to health technology evaluations ("the manual"). One significant change from previous methods is the incorporation of a decision modifier for severity. The manual specifies that the appraisal committee will consider the severity of the condition and apply a higher weight to health benefits where severity is judged to be higher than those where severity is lower. The manual covers highly specialised technologies (HST), diagnostics, medical technologies, and health technology appraisals. However, the severity modifier applies primarily to health technology appraisals (it is not applied in HST, is "not normally" applicable in diagnostics, and it is "considered deliberatively" in medical technologies, see sections 6.2.19 to 6.2.21).

Severity is to be calculated as "Absolute Shortfall" (AS) and "Proportional Shortfall" (PS). Both express the difference between Quality Adjusted Life Years (QALYs) over remaining life expectancy for the general population compared to those with the health condition, conditional on age and sex. AS expresses this difference in absolute QALYs. PS expresses this as the ratio of general population QALYs / QALYs for those with the condition.

Weights should be applied to incremental QALYs. The weight that is to apply is specified in Table 6.1 of the manual and reproduced below. The definition that provides the highest weight is the one that should be used.

Table 1: QALY weightings for severity

QALY weight	Proportional QALY shortfall	Absolute QALY shortfall
1	Less than 0.85	Less than 12
x1.2	0.85 to 0.95	12 to 18
x1.7	At least 0.95	At least 18

The manual also states that both AS and PS should be calculated using the reference case discount rate. The QALYs for those with the health condition should be based on current established NHS practice.

The introduction of severity weights is a significant change. Whilst other factors may be relevant to committee decision making, severity is the only factor in the updated

manual that has a specified quantitative definition and weight to be applied. This replaces “end-of-life” which, since 2009, allowed additional weight to be assigned to health benefits for technologies that offer life extension above a minimum threshold for patients with very short life expectancy under current NHS care.

The manual does not provide detail on how the severity calculations should be performed. The purpose of this document is to outline the different components of information that are required to calculate AS and PS, provide information on different sources of data, discuss the advantages and disadvantages of these sources, and provide guidance on how to estimate AS and PS.

2. ABSOLUTE AND PROPORTIONAL SHORTFALL

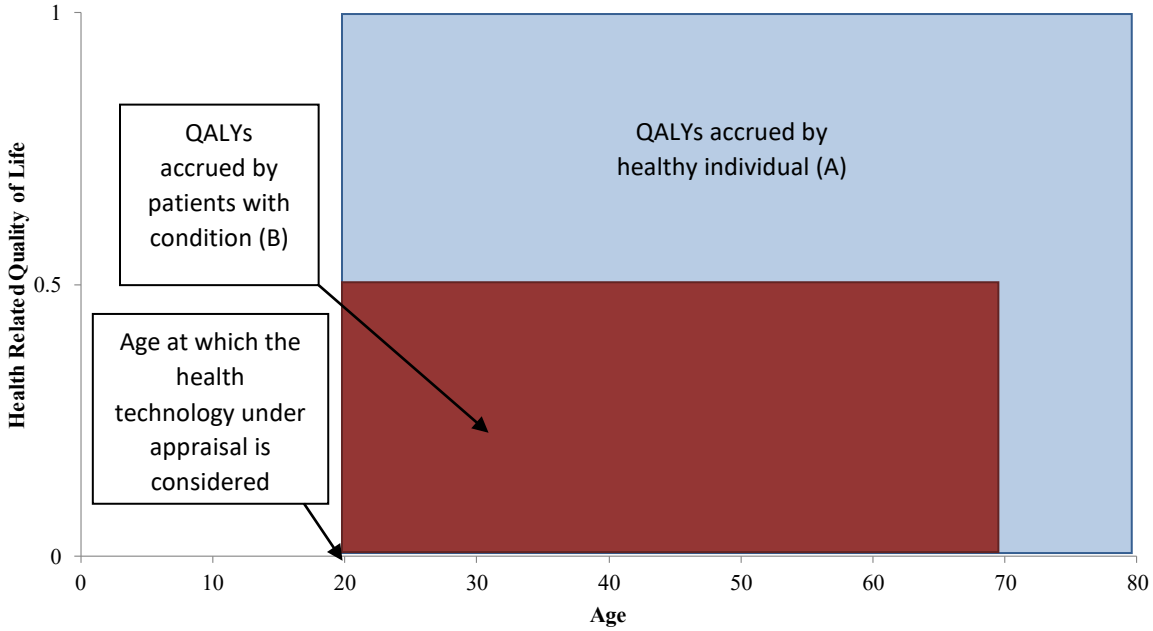
2.1 Defining terms

Severity of disease, as described in the manual and measured by AS and PS, refers to the difference between the average number of QALYs that patients with the condition would be expected to receive over the remainder of their lifetimes if they received current NHS care, compared to the number of QALYs they would receive if they did not have this condition. We assume that expected QALYs in the absence of the condition is equivalent to the QALYs that members of the general population of England would receive over the remainder of their lifetimes.

This is a definition that focusses on health from the point at which the health technology in question is being considered for use. It does not take into account past health.

Figure 1 illustrates the concepts in simplified form.

Figure 1: Burden of Illness



In this situation, the health technology under consideration is for a position in the disease treatment pathway where patients have a mean age of 20 years. For those patients, assume that the mean impact of the condition going forward is to reduce quality of life to 0.5 and to reduce length of life to 70 years. This is represented by the red shaded area (labelled B) and is equivalent to 25 undiscounted QALYs (50 years *x* 0.5). (We use the undiscounted figure here for ease of illustration). Compare this to the expected number of QALYs for a member of the general public at the same age: a life expectancy of 80 years at full health, in this example. This is represented by the red and blue shaded areas combined (labelled A) and is equivalent to 60 undiscounted QALYs (60 years *x* 1). Note that the comparison is between the expected (mean) QALYs with and without the condition (this assumes that the general population estimates can be used to estimate health in the absence of the condition). Use of the mean starting age (and mean proportion of males versus females) rather than some other measure of central tendency is consistent with the other parts of this calculation.

$$\textit{Absolute shortfall (AS)} = Q_h - Q_c$$

$$= 60 - 25$$

$$= 35 \text{ QALYs.}$$

Where, Q = QALYs, h = age/sex equivalent general population, c = the population for whom the new technology is being considered.

AS is expressed in terms of QALYs. It lies within a range that is limited at the bottom by zero (where the relevant condition imposes no reduction in either quality or quantity of life, that is $Q_h = Q_c$). The upper limit, in most cases, is the number of QALYs for the general population, Q_h (where the condition causes immediate death i.e. $Q_c = 0$). In the example above this would be 80 QALYs. It is theoretically possible for AS to be greater than Q_h in those circumstances where Q_c is negative. This would occur if the health condition imposed a large impact on quality of life such that a negative health utility weight was assigned for a sufficient period that this outweighed any period spent in states valued above zero. This is unlikely to occur in practice. The number of QALYs for the general population is greater at lower ages. A higher AS indicates greater severity.

$$\textit{Proportional shortfall (PS)} = (Q_h - Q_c) / Q_h$$

$$= (60 - 25) / 60$$

$$= 0.583$$

PS is a ratio that should normally lie in the interval 0 to 1. PS is equal to zero when the condition in question imposes no reduction in either quality or quantity of life, that is $Q_h = Q_c$. PS is equal to 1 if the condition causes immediate death, that is $Q_c = 0$. Theoretically, PS could be greater than 1 in the situation where the condition in question causes quality of life to reduce below zero and this is not offset by any subsequent period spent in health states considered rated above zero. Higher PS indicates a greater degree of severity.

2.2 Data requirements

Both AS and PS are relatively simple concepts but to estimate these values in practice requires a number of different pieces of information to be brought together.

Future QALYs for the general population (Q_h) require estimates of expectancy of life from the relevant mean age and for each year of that life to be adjusted for quality. The manual states that all estimates should be discounted.

Future QALYs for those with the condition under current NHS treatment (Q_c) requires more complex modelling. The basic requirement is the same as for the general population (expectancy of life from the same age, with each year adjusted for quality) but these estimates are specific to the condition and the effectiveness of current NHS care.

We outline the data requirements for these calculations in turn below.

3. CALCULATION OF GENERAL POPULATION QALYS

3.1 Expectation of life

Life tables for England describe mortality throughout the life cycle and are published by the Office for National Statistics [ONS]. They are published in three-year bands[2]. The most current set of life tables (published Sept 2021) is for 2018 – 2020 based on mid-year population estimates for 2018, 2019 and 2020.

We use the probability of death (q_x) between age *x* and *x+1*, by sex, to calculate the proportion of individuals from a given starting age that will remain alive at each age between 0 and 100 years (the age range covered by the life tables).

Deaths are assumed to occur linearly over each year of age.

“Period” and “cohort” life expectancy (and life tables) are reported. Period life expectancy is based on age-specific death rates for the given period and assumes that mortality rates remain constant into the future. Cohort life expectancy allows for known or projected changes in mortality over the remaining lifetime. The ONS provide details of the difference between these two approaches here[3].

For the purposes of calculating the proportions of individuals alive at each age, stakeholders submitting these analyses to NICE need to estimate the proportion of patients that are male/female at the mean age *x*, who have the relevant condition (for which the new technology is under consideration) at the point in the disease pathway where the technology is being assessed and, where appropriate, for the relevant subgroup. This information would typically be included as standard in most current assessments of cost effectiveness in some form. Often, such estimates are drawn

from pivotal trials of the health technology in question with some recognition that the true NHS population may differ somewhat. For example, trial populations are often younger and fitter than the broader patient population. The analysis will need to use the age and sex distribution used in the economic model to ensure consistency. It is therefore of increased importance that decision models endeavour to reflect English, NHS, real-world populations and there is transparency where this is not achieved.

Cohort life tables use a combination of observed mortality rates for past years and projections for future years. The ONS provide an example to illustrate this from 2019:

"For example, cohort life expectancy at age 65 years in 2016 would be worked out using the mortality rate for age 65 years in 2016, for age 66 years in 2017, for age 67 years in 2018 and so on. This uses observed mortality rates up to 2018 and projected mortality rates from the most recently published projections for 2019 onwards."

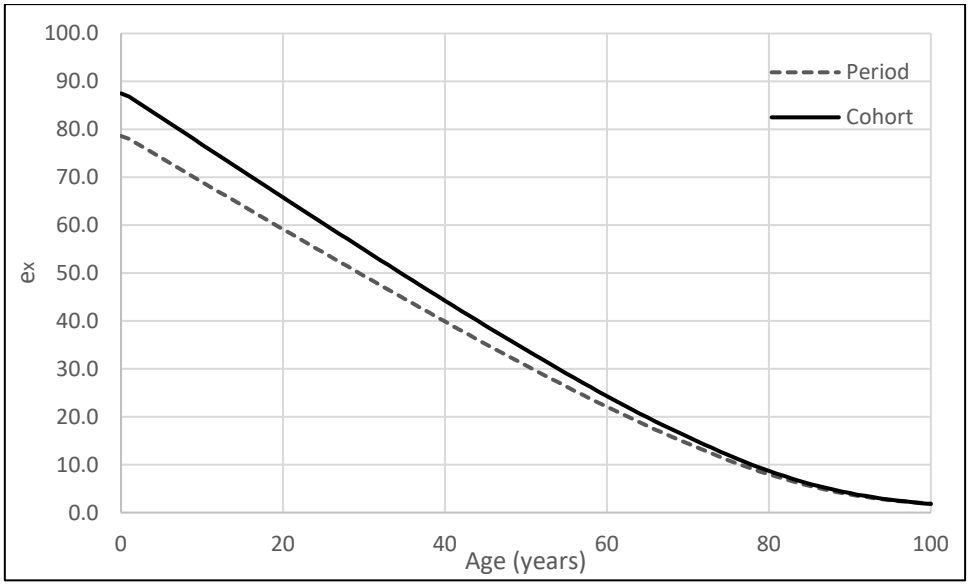
Observed mortality rates can of course, only be used when we are dealing with life expectancy at age x from some point in the past. If, in the above example, we wanted to estimate cohort life expectancy for someone aged 65 years in 2019, that would be based entirely on projections of how mortality rates may change and are therefore uncertain. They entail assumptions about future changes to mortality rates. Expert opinion is engaged to help determine how long historical trends will prevail but unforeseen changes, such as the Covid-19 pandemic, can impact on these projections. ONS publish a range of these projections to indicate sensitivity to the assumptions and the degree of uncertainty.

Both the absolute difference and the difference as a proportion of remaining years of life between period and cohort expectation of life are greatest at younger ages. Figure 2 illustrates this for males using figures for 2020. Note that these are the most recently published life tables for single years from ONS at the time of writing, published on 12th January 2022[4]. There is an absolute difference of approximately 9 years from birth (78.6 years vs 87.5 years for period and cohort expectations of life respectively). This difference is 5.5 years at the age of 30 years and 2.2 years at the age of 60 years. The potential impact of these differences on subsequent estimates of AS and PS will depend on the age of the relevant patient population, the impact of the technology on mortality and the discount rate. Health technologies that have mortality benefits for

conditions that affect children or young adults would find that estimates of severity would be most affected by the choice of period or cohort life expectancy.

Period expectation of life is based on currently observed data and does not entail extrapolation. This approach uses mortality estimates from the current period (a year or set of years) and assumes that those rates will continue to apply throughout the rest of a person's life. These are in more widespread use, including in economic evaluation, and are subject to less subjective extrapolation. For these reasons, and to retain consistency with the calculation of comparator QALYs, period life tables should be used.

Figure 2: Period versus cohort expectation of life years by age, for males in England 2020



Note: e_x is the average period expectation of life at exact age x

Best practice for inputs of this type used for calculations in submissions to NICE would use the most up to date source of evidence available. At the time of writing, published life tables cover up to the year 2020. The latest three-year band is 2018 – 2020. However, the Covid-19 pandemic has impacted mortality figures and it may be more appropriate for earlier periods to be used for NICE assessments at the current time.

Period expectation of life for males was lower in 2020 than any other year from 2010 onwards for all ages between 0 and 98 years. Expectation of life was lower in 2020 compared to 2019 by 1.3 years at age 0 years, by 1.2 years at age 50 years and 0.7 at age 80 years. Similar findings were found for females. Expectation of life for females

was lower in 2020, for all ages up to 100 years, than in each of the preceding years. 2020 expectation of life was lower by 1 year at birth and 0.7 years at age 80.

The impact of Covid-19 on long term mortality rates will remain unknown for some years. For current NICE assessments, it would be prudent to use figures from the three-year band 2017 - 2019 and for the Institute to commit to regular reviews of recommended data sources.

It is also worth noting that standard life tables are based only on age and sex. Other factors, such as socio-economic group and ethnicity, are known to be related to life expectancy. Where a particular condition impacts individuals from groups unequally compared to the overall general population, life expectancy estimates may be misleading. For example, where a condition disproportionately affects individuals from lower socio-economic groups, the proposed approach of using general population life expectancy will bias the estimates of AS and PS upwards: the impact of the condition will appear more severe than is truly the case. This is likely an acceptable limitation but will need to be considered if QALY weights for these factors are explicitly introduced by NICE in the future.

3.2 Health utility weights

Health related quality of life within the general population changes with age. As part of a cost effectiveness analysis, it is often the case that adjustments to health utilities for specific health states are required to take account of the impact of ageing within the population. The current methods guide states: “Adjustment should be based on a recent and robust source of population health-related quality of life” but it does not provide more specific guidance.

The EQ-5D is the preferred measure of health outcome for NICE and NICE accepts data from the descriptive systems of either the 3L or 5L instruments. NICE does not recommend using the EQ-5D-5L value set for England published by Devlin et al. (2018)[5]. Instead, NICE currently specifies that where data from the EQ-5D-5L is available, values that align to the EQ-5D-3L should be estimated using a mapping model. A new value set for EQ-5D-5L was published in March 2026[6] and, at the time of writing this update, NICE has indicated its intention to recommend the use of this value set, subject to the completion of a public consultation[7]. Therefore, this report has

been updated to include recommendations for calculating general population and comparator QALYs using the EQ-5D-5L and its value set.

3.3 Using the EQ-5D-3L

The DSU published a report on sources of evidence for estimating age and sex-based population norms for EQ-5D[8]. Suitable, large-scale surveys of the population that include EQ-5D-3L have become very limited in their availability in recent years. The Health Survey for England (HSE) last included EQ-5D-3L in 2014 and those data (n=7,085 usable responses) were used to estimate a model of EQ-5D-3L as a function of age and sex. Consistent with previous estimates, this showed declining health utility with increasing age. A very large-scale UK study (n=50,000), conducted in early 2020 was also analysed and showed a more complex relationship with age. Data collection coincided with the early stages of the Covid-19 pandemic and the imposition of strict social distancing measures in the UK, including the requirement to stay at home. It is possible that these results are therefore not representative of the health utilities that would be observed currently in the absence of such measures. Further investigation is required into this issue, and, in the meantime, the recommendation is to use the modelled results derived from the 2014 HSE.

There are datasets that include EQ-5D-5L, including the HSE in 2018. Mapping methods could be used to estimate EQ-5D-3L values. However, mapping does not generate data. It is typically used to generate predicted values. The distribution of these predictions is not equivalent to data, and this is a potential source of inconsistency with directly observed EQ-5D-3L values in this setting. In addition, there is uncertainty inherent in the use of statistical methods for modelling EQ-5D-3L by age and sex that differs from that associated with direct observation of data.

3.4 Using the EQ-5D-5L

The most suitable source for directly observed, large, UK, general population samples of EQ-5D-5L is the HSE for years 2017 and 2018. Hernandez et al[9] provide full details of statistical modelling used to estimate EQ-5D-5L conditional on age and sex using the value set for the UK[6]. The final estimation samples consist of 8,094 females and 6,319 males.

3.5 Utility weights for children

The HSE administered the EQ-5D-3L/5L to participants over the age of 16 years and therefore, estimates for ages <16 years by extrapolation should be treated with some degree of caution. Non-linear models were used for the estimation of EQ-5D-3L/5L and these can be used to predict at these out of sample younger ages. Considering 3L, for females, the non-linear relationship with age is more pronounced than for males and, at ages 16 and below, the relationship shows slightly decreasing quality of life with decreasing age, though this is marginal and broadly flat. For males the relationship implies slightly higher quality of life for younger individuals. Based on these results, calculations in the DSU report chose to apply the same EQ-5D-3L at ages less than 16 as were predicted by the model for those aged 16, conditional on sex. This is also the case for the updated EQ-5D-5L based estimates.

3.6 Discounting

Future QALYs should be discounted at the reference case rate of 3.5% in order to reflect their current value.

4. CALCULATION OF COMPARATOR QALYS

It is expected that few additional calculations will be required to estimate expected QALYs for patients with the relevant condition under current NHS care (Q_c), beyond those conducted routinely as part of the cost effectiveness analysis in most cases. In general terms, developers of cost effectiveness analyses should try to ensure that the analyses are consistent with the methods described in section 3 for estimating Q_h. This applies to the source of life expectancy for the general population, health utility data and discount rates.

4.1 Discounting

The reference case discount rate of 3.5% should be applied to future health benefits. This ensures that the assessment of severity across all appraisals is applied in a consistent manner, even for those appraisals where the committee may accept an alternative, non-reference case discount rate for the purposes of the cost-effectiveness analysis. A rate of 1.5% can be used in non-reference case cost effectiveness analyses in some circumstances. However, for the assessment of

severity in a fair manner across appraisals, the calculations of AS and PS only should be made in the same way, using 3.5%, and comparability ensured.

4.2 Comparators

Standard cost effectiveness analysis conducted as part of a NICE evaluation should include established practice in the NHS as the comparator to the technology in question. In some situations, there is uncertainty about current NHS care or a degree of heterogeneity in clinical practice. Where differences of comparator are considered to be appropriate for specific patient subgroups, these should be considered as separate decisions requiring estimates of AS and PS (and cost effectiveness) for each subgroup. This is consistent with the use of subgroups for cost effectiveness analysis.

Where multiple comparators are presented and all represent practices that occur to some degree within the NHS, estimates of AS and PS should be presented for each. It is committee judgement that should determine which is the most appropriate comparator as is currently undertaken when interpreting cost effectiveness results.

Often, standard of care will include a variety of treatment options that are similar in terms of clinical effectiveness. This creates a challenge for cost effectiveness analysis when those components differ in terms of their costs. AS and PS estimates only consider benefits and therefore will not need to be disaggregated in this situation.

4.3 Lifetime horizon

Most but not all decision models used for cost effectiveness analysis will use a lifetime horizon since this is the period over which there can reasonably be expected to be differences in either the costs or QALYs (or both) generated by the new intervention and comparator. The manual specifically refers to the use of a shorter than lifetime horizon where there is no mortality effect between technologies and the differences in costs and outcomes relate to a shorter time period. Where a lifetime horizon has not been used (as for example in TA213, Aripiprazole for schizophrenia in adolescents) there is no estimate of lifetime QALYs for the comparator arm (Qc) and further extrapolation is required.

The most straightforward situation is where the same decision model can be run for a longer period than has been used for the base case cost-effectiveness analysis. Recall that the interest here is solely the estimates of QALYs for the comparator arm. The

new intervention arm, and validity of differences in either costs or QALYs between the two arms, is not the issue of interest for the assessment of AS and PS. Those comparisons may provide some insight into the validity of estimates of Qc but equally may not be relevant. In some situations, a lifetime horizon is not adopted in the economic analysis because of uncertainty about the duration of treatment effects. Such concerns may not impact the validity of using the modelling framework to estimate QALYs for the comparator arm. Where there are valid concerns about the absolute estimates of survival and quality of life from the submitted economic model, these will need to be documented and addressed in sensitivity analysis if feasible.

4.4 Carer Quality of Life

There are a small number of technology appraisals where the impact of a condition on carer health related quality of life is included as part of the economic evaluation. A review of NICE appraisals conducted in 2019 found 12 TAs from a total of 422 did so[10]. Eight of these were in appraisals of treatments for people with multiple sclerosis. The near absence of quantitative assessment of impacts on carer quality of life in historical Technology Appraisals that formed the basis of the calculation both of the cut-offs and weights for severity in current use is an important issue.

Carer impacts should be excluded from the calculation of AS and PS. Severity weights should be applied only to the incremental QALYs for patients. The comparison being made is between the QALYs people would be expected to receive with and without the condition, where “without the condition” is taken from general population estimates. If the estimates of QALYs for patients include both patients and their carers, then there is an inconsistency between the two parts of the calculation.

4.5 Source of utility values

NICE currently recommends the EQ-5D-3L value set be used. At the time of writing this update, NICE has indicated its intention to adopt the EQ-5D-5L value set, pending the completion of a public consultation[7]. The following analyses and recommendations are relevant only to the period where the EQ-5D-3L value set remains the preferred method for estimating QALYs. In those situations where the descriptive system of the EQ-5D-5L has been used, NICE recommends values be estimated using a published mapping model, to link those responses to the EQ-5D-3L value set.

The recommendations above in relation to the estimation of QALYs for the general population favour directly observed EQ-5D-3L for the base case analysis rather than EQ-5D-3L values mapped from responses to the EQ-5D-5L descriptive system. This is because mapping introduces an additional aspect of uncertainty into predicted EQ-5D values and those predictions do not have the same distributional characteristics as observed data, which raises issues about the optimal methods for modelling health utility as a function of age and sex. Where the estimates for QALYs for current NHS treatment(s) are based on mapped EQ-5D-3L values, the primary issue of concern for NICE is that both parts of the calculation (comparator QALYs and general population QALYs) are being conducted in terms of EQ-5D-3L values. Nevertheless, additional confidence may be provided through confirmatory sensitivity analyses which use EQ-5D-5L data mapped to EQ-5D-3L values for the general population estimates of QALYs.

Figure 3 and Figure 4 compare EQ-5D-3L by age, separately for males and females, from two sources: EQ-5D-3L from the HSE 2014 and EQ-5D-3L mapped from EQ-5D-5L using the NICE recommended approach for mapping (Hernandez et al[11]) and data from the HSE 2018. This is the most recent wave of the HSE that contained EQ-5D of any version. It should be noted that the HSE 2014 plot is for modelled results whereas the 2018 data is simply a plot of the mean observed EQ-5D by age. The 2018 data are supplied with age as a categorical variable, and we have taken the mid-point of these categories for reporting purposes. For females there are small differences in EQ-5D-3L between the two data sources. Between the ages of 16 and 67 years, the HSE 2018 estimates are below those of the HSE 2014 by an average of less than 0.05 and from above 67 years the difference is 0.014. A similar picture is seen for males. The average difference is less than 0.04 up to 67 years, and 0.011 up to aged 90 years.

Figure 3: EQ-5D-3L by age for females: HSE 2014 compared to HSE 2018

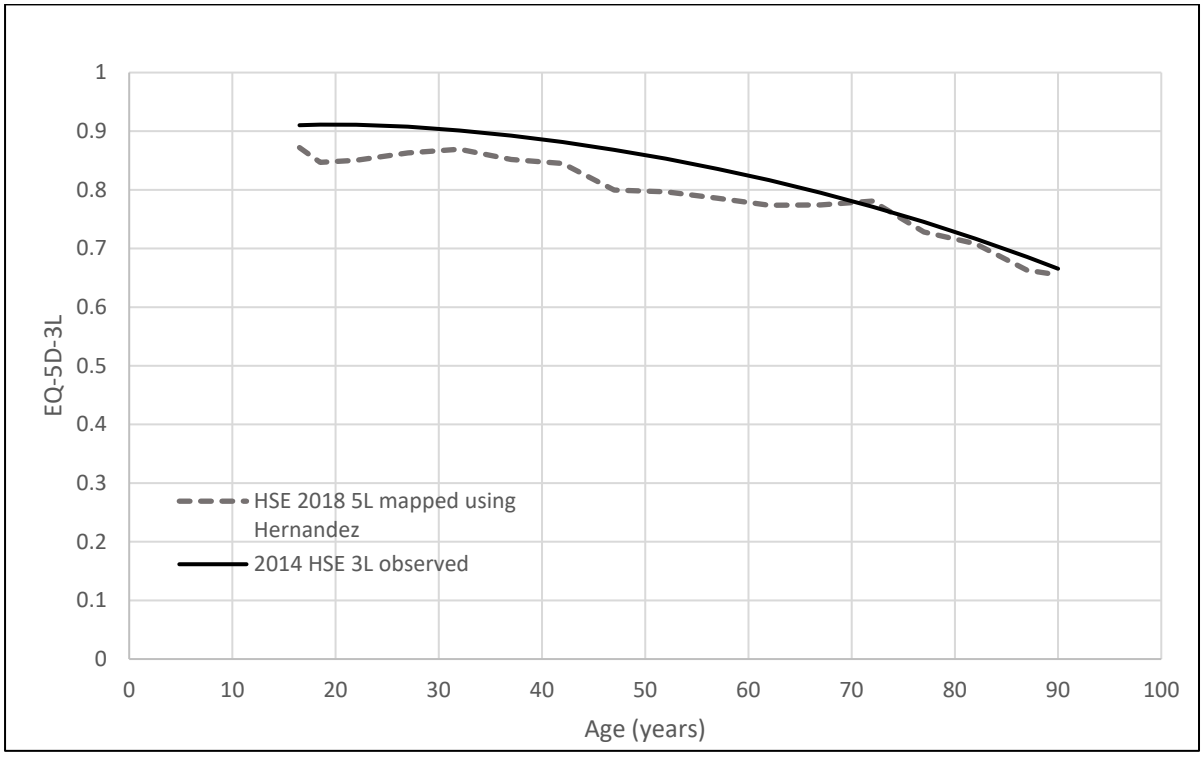
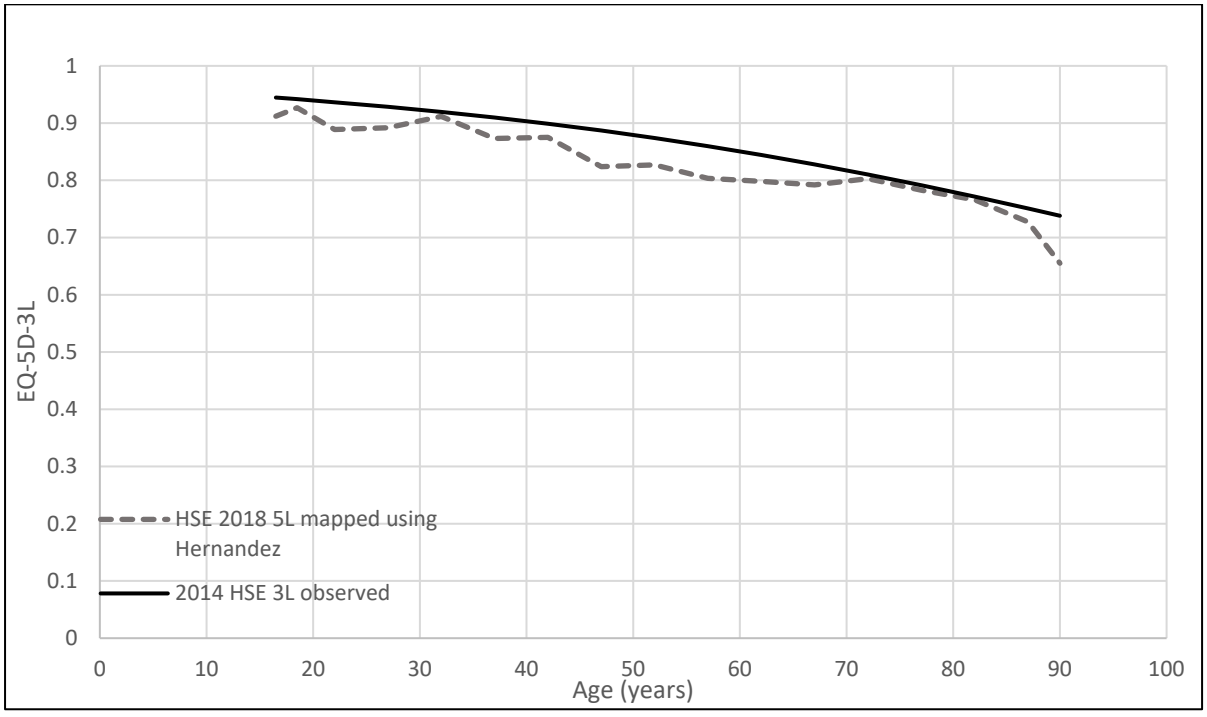


Figure 4: EQ-5D-3L by age for males: HSE 2014 compared to HSE 2018



Where other sources of utility values have been used, the issue of potential comparability between estimates of QALYs for patients treated under current NHS practice (the comparator arm of the economic model) and QALYs for the general population remains important.

Where other generic, preference-based measures of quality of life rather than EQ-5D-3L have been used in the economic model there are several potential options:

i) it may be feasible to estimate health utility by age and sex from appropriate sources. For example, the SF12 is included in “Understanding Society”, a large, UK longitudinal study.
ii) Mapping from one generic instrument to EQ-5D-3L. Published studies have reported mapping from SF-12[12], SF6D, ICECAP, HUI2[13] and HUI3[14].
iii) Qualitative consideration of the potential impact on estimates of AS and PS of different data sources.

Where other methods have been used to estimate health state utilities for significant aspects of the economic model, such as disease specific preference-based measures, direct patient valuation or vignettes, there are unlikely to be general population estimates or mapping studies available or considered appropriate to perform.

In all cases, a base case analysis using EQ-5D-3L for the general population QALYs calculation should be presented.

5. WORKED EXAMPLE

5.1 Using EQ-5D-3L

In a hypothetical example, the cost effectiveness model output preferred by the appraisal committee estimated QALYs for the comparator arm to be 6.031 QALYs. This was based on a lifetime model horizon, discounted at 3.5% and the comparator is NHS standard of care. The mean age of patients eligible for the technology was 62 years and 33% are female. These figures were used in the economic model and considered to reflect the relevant NHS population.

Calculations of AS and PS are provided in Table 2. See Appendix 1 for more details on the calculation of QALYS for the general population.

Table 2: AS and PS for a hypothetical example (using EQ-5D-3L)

Age (years)	62		
Proportion female	0.33		
QALYs from CEA	6.031		
	Male	Female	Overall
Life expectancy	21.3	23.8	22.1
QALYs	17.07	17.91	17.35
Discounted QALYs	11.86	12.16	11.96
AS	**5.93**		
PS	**0.50**		

AS and PS require an estimate of the number of QALYs for the age and sex matched general population. At age 62, life expectancy for males is 21.3 years and for females 23.8 years based on 2017 - 2019 life tables for England. Each period of future life is assigned a quality adjustment weight based on HSE 2014 and that is discounted at 3.5%. Discounted QALYs is 11.86 for males and 12.16 for females. The weighted average QALYs for the overall population is therefore 11.96 QALYs. Note that the lower limit for AS to attract a QALY weight above 1 is 12 QALYs. Therefore, for this condition and for the age and sex of patients at this point in the treatment pathway, AS is highly unlikely to exceed 12.

AS is calculated as the difference between QALYs for the general population and the patient population: 11.96 – 6.031 **= 5.93 QALYs.**

PS is calculated as the absolute shortfall divided by the number of QALYs for the general population: 5.93/11.96 **= 0.50.**

Both AS and PS indicate that the appropriate QALY weight is 1, by some margin.

5.2 Using EQ-5D-5L

For illustrative purposes only, the same input values as for the EQ-5D-3L example above are presented in Table 3 for the EQ-5D-5L.

Table 3: AS and PS for a hypothetical example (using EQ-5D-5L)

Age (years)	62		
Proportion female	0.33		
QALYs from CEA	6.031		
	Male	Female	Overall
Life expectancy	21.3	23.8	22.1
QALYs	18.05	19.42	18.51
Discounted QALYs	12.49	13.10	12.69
AS	**6.66**		
PS	**0.52**		

6. SUMMARY

The calculation of AS and PS for NICE appraisals should be relatively straightforward in most situations because it is largely based on information and analyses already required for submissions to NICE.

This report makes recommendations for data sources and methods to calculate QALYs for the otherwise healthy population and this can be compared to the comparator arm from the cost effectiveness analysis in many cases. These proposals aim to ensure the most appropriate approaches are used and decision makers can have confidence that estimates are made in a consistent manner across appraisals.

However, no recommendations can eliminate the requirement for committees to exercise their judgement, as is the case with other analyses submitted to NICE including cost effectiveness modelling. Many of the issues that will be of concern in estimating AS and PS will be equally relevant to the consideration of cost effectiveness estimates. There will inevitably be circumstances where a case can be made that AS and PS estimates are biased in some way. When this is close to one of the thresholds for additional QALY weighting, judgements will be required.

The following recommendations should be considered:

- NICE should monitor the implementation of severity weights and identify any situations where there is a lack of clarity about suitable approaches.
- A dataset of all estimates of AS and PS that are accepted in NICE appraisals to be maintained and made publicly available (where considerations of

confidentiality permit). This case history would help committees judge the validity of estimates that are presented in the same disease areas and by position in the disease pathway. This would also help to resolve difficulties where there are claims made that the estimates are biased in some way. This may be, for example, because different sources of health utilities are used between the calculation of QALYs for the general population versus what is used in the economic model for the comparator, or because there are concerns about the baseline estimates of survival in the economic model.

- Reviews of relevant data-sources for the calculation of AS and PS should be conducted regularly.
- Whilst the analyses presented here are not complex, pre-programmed software to automate the task is available that simplifies the task further. A simple excel calculator is freely available on the DSU website[15]. McNamara et al have programmed a version based on R with a user-friendly front end[16].
- Both versions of the calculator have been updated to allow both EQ-5D-3L and EQ-5D-5L based AS/PS calculations to be obtained.

REFERENCES

[1] NICE NICE Health Technology Evaluations: The Manual. 2022 Available at https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation

[2] ONS National Life Tables: England. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandreferencetables/current

[3] ONS Period and cohort life expectancy explained: December 2019. Available at https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/methodologies/periodandcohortlifeexpectancyexplained

[4] ONS Past and projected expectations of life (ex) from the 2020-based (England) life tables available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/expectationoflifeprincipalprojectionengland

[5] Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Econ. 27(1):7-22. DOI: 10.1002/hec.3564. Epub 2017 Aug 22. PMID: 28833869; PMCID: PMC6680214.

[6] Rowen D, Mukuria C, Bray N, Carlton J, Longworth L, Meads D, Oluboyede Y, O'Neill C, Yang Y. The new EQ-5D-5L Value Set for the UK, Value in Health, 2026 DOI: 10.1016/j.jval.2026.03.008

[7] See https://www.nice.org.uk/news/articles/changes-to-nice-s-cost-effectiveness-thresholds-confirmed

[8] Hernández Alava M., Pudney S., Wailoo A. Estimating EQ-5D by Age and Sex for the UK. NICE DSU Report. 2022. Available at https://www.sheffield.ac.uk/sites/default/files/2022-02/DSU%20Age%20based%20utility%20-%20Final%20for%20website.pdf

[9] Hernandez-Alava M, Wailoo A EQ-5D-5L age adjustment estimates using the health survey for England. Policy Research Unit in Economic Evaluation of Health and Social Care Interventions. Universities of Sheffield and York. 2026 Report 081 DOI: 10.15131/shef.data.31889488

[10] Pennington R and Wong R Modelling Carer Health-Related Quality of Life In NICE Technology Appraisals and Highly Specialised Technologies, 2019 Available at https://nicedsu.sites.sheffield.ac.uk/methods-development/modelling-carer-health-related-quality-of-life

[11] Hernandez Alava M, Wailoo, A., and Pudney S. Methods For Mapping Between The EQ-5D-5L and the 3L For Technology Appraisal, NICE DSU Report, 2017: http://nicedsu.org.uk/wp-content/uploads/2017/05/Mapping-5L-to-3L-DSU-report.pdf

[12] Gray A. M., Rivero-Arias O., Clarke P. M. Estimating the association between SF-12 responses and EQ-5D utility values by response mapping. Med Decis Making. 2006 26 (1), 18-29.

[13] Hernandez Alava, M., Brazier, J., Rowen, D., & Tsuchiya, A. Common scale valuations across different preference-based measures: estimation using rank data. Med Decis Making, 2013 33(6), 839-852.

[14] Gamst-Klaussen, T., Chen, G., Lamu, A. N., & Olsen, J. A. Health state utility instruments compared: inquiring into nonlinearity across EQ-5D-5L, SF-6D, HUI-3 and 15D. Qual Life Res, 2016 25(7), 1667-1678.

[15] https://sheffield.ac.uk/nice-dsu/tsds/severity-shortfall-tsd

[16] McNamara S, Schneider P, Love-Koh J, Doran T, Gutacker N. "Quality-adjusted life expectancy norms for the English population", Value in Health, 2023 Vol. 26: 163-169.

APPENDIX 1:

Using mortality rates from life tables to calculate Quality Adjusted Life Years for the general population.

Let q_x be the mortality rate between age *x* and *(x +1)* provided in life tables, that is, the probability that a person aged *x* exact will die before reaching age *(x +1)*. l_x is the number of survivors to exact age *x* that are subject to the mortality rates q_x. Let *i* be the starting age in years.

i) Calculate average life years:

Let d_x be the proportion of deaths during each year for values of *x* = *i* to 100:

$$d_x = q_x . l_x$$

$$l_{x+1} = l_x - d_x$$

The number of years of life, L_x, at each age with half cycle correction (that is, assuming that deaths occur linearly throughout the year):

$$L_x = \frac{l_x + l_{x+1}}{2}$$

Average years of life from age *i* to 100 (the upper limit of the life table) is the sum of the proportion of individuals alive at each age:

$$\sum_{x=i}^{100} L_x$$

ii) Calculate Quality Adjusted Life Years (QALYs).

U_x is the quality adjustment weight by age:

$$\sum_{x=i}^{100} L_x . U_x$$

iii) Calculate discounted QALYs.

Let *r* be the annual discount rate:

$$\sum_{x=i}^{100} L_x . U_x . \frac{1}{(1+r)^{x-i}}$$

iv) Since q_x and U_x differ for males and females, steps 1 to 4 should be calculated separately for the two sexes and the weighted average taken