COMPARING THE EQ-5D-3L AND 5L VERSIONS. WHAT ARE THE IMPLICATIONS FOR MODEL-BASED COST EFFECTIVENESS ESTIMATES?

REPORT BY THE DECISION SUPPORT UNIT

13 August 2018

Becky Pennington[1], Monica Hernandez-Alava[1], Stephen Pudney[1], Allan Wailoo[1]

[1]School of Health and Related Research, University of Sheffield

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
Website www.nicedsu.org.uk
Twitter @NICE_DSU

ABOUT THE DECISION SUPPORT UNIT

The Decision Support Unit (DSU) is a collaboration between the Universities of Sheffield, York and Leicester. We also have members at the University of Bristol, London School of Hygiene and Tropical Medicine and Brunel University. 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 Technology Appraisal Programme. Please see our website for further information www.nicedsu.org.uk

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.

This document does not represent NICE guidance. Readers should refer to the NICE website (www.nice.org.uk) for guidance on specific technologies.

Acknowledgements

The authors wish to thank the FORWARD databank, their patient participants and directors Kaleb Michaud and Fred Wolfe. The authors wish to thank the companies who gave permission for their models to be considered as case studies in this report.

CONTENTS

TABLES

FIGURES

ABBREVIATIONS AND DEFINITIONS

3L	EuroQol-5 Dimension-3 Level
5L	EuroQol-5 Dimension-5 Level
AE	Adverse Event
CABG	Coronary Artery Bypass Graft
DSU	Decision Support Unit
EAP	Early Access Programme
ECOG	Eastern Cooperative Oncology Group
EDSS	Expanded Disability Status Score
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer core quality of life questionnaire
EQG	EuroQol Group
ERG	Evidence Review Group
FACT-O	Functional Assessment of Cancer Therapy – Ovarian
FACT-O	Functional Assessment of Cancer Therapy – Prostate
FAD	Final Appraisal Determination
HCC	Hepatocellular Carcinoma
HS	Haemorrhagic Stroke
HST	Highly Specialised Technologies
IS	Ischaemic Stroke
MI	Myocardial Infarction
MS	Multiple Sclerosis
NICE	National Institute for Health and Care Excellence
ICER	Incremental Cost-Effectiveness Ratio
PCI	Percutaneous Coronary Intervention
PTCA	Percutaneous Transluminal Coronary Angioplasty
QALY	Quality-Adjusted Life Year
RRMS	Relapsing Remitting Multiple Sclerosis
SD	Standard Deviation
SPMS	Secondary Progressive Multiple Sclerosis
SVR	Sustained Virologic Response
TA	Technology Appraisal

TIMI	Thrombolysis in Myocardial Infarction
VAS	Visual Analogue Score

EXECUTIVE SUMMARY

The National Institute for Health and Care Excellence (NICE) states that the preferred measure of health-related quality of life in adults is the EuroQol 5-Dimension instrument (EQ-5D). The EQ-5D asks a person to describe their health across five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Two descriptive systems of the EQ-5D exist: the 3-level (3L) and 5-level (5L): the 3L has 3 levels of severity for each dimension and the 5L has 5. Valuation sets exist for both the 3L and 5L which assign a utility value (measure of health-related quality of life) to each health state described by the different levels of severity for each dimension. Previous work by the Decision Support Unit (DSU) found that mapping 3L utility values to 5L in cost-effectiveness studies alongside clinical trials decreased incremental quality-adjusted life year (QALY) gains in most cases studies, with the exception of one case study where life extension was a substantial element of health gain. Most Technology Appraisals (TAs) considered by NICE use model-based economic evaluations with mean utility values, so there was a need to explore the impact of converting 3L utility values to 5L in economic models.

We developed a mapping command in Stata; EQ5Dmap; to allow patient-level or summary 3L utility values to be converted to 5L utility values using the EuroQoL group (EQG) and National Databank for Rheumatic Diseases FORWARD datasets. Where only mean 3L utility values are available, the command uses a bandwidth parameter to define the neighbourhood over which averaging is to be done and the rate at which the weight declines with increasing distance. We recommend that the size of the bandwidth depends on the 3L utility value, with lower utility values requiring larger bandwidths.

We considered comparisons from economic models from 20 NICE TAs. The comparisons covered a range of disease areas, and included 10 oncology case studies where the intervention extended life and improved quality of life, 6 non-oncology case studies where the intervention extended life and improved quality of life, and 4 non-oncology case studies where the intervention improved quality of life and did not extend life. We mapped the 3L utility values to 5L using the Stata command, and compared the incremental QALYs and incremental cost-effectiveness ratios (ICERs) using directly administered 3L and mapped 5L.

Almost all utility values increased using 5L, but the difference between best and worst states decreased, so the magnitude of change in the QALY gain depended on how much of the 3L QALY gain was attributed to life extension, and how much to improvement in quality of life.

For the 10 oncology case studies, when mapped 5L utility values were used, using either the EQG or FORWARD databanks, the incremental QALYs increased and ICERs decreased compared to 3L, with one exception. Where case studies considered utility as a function of time to death, the magnitude of the increase in QALY gain depended on the increase in the utility value in the health state furthest from death. Where case studies considered utility as a function of pre- and post-progression, both utility values increased, but the difference between the utilities decreased. The direction of change in QALY gain therefore depended on the proportion of life extension in the pre-progression state and the post-progression state (as well as the initial 3L utility scores). Where there was a treatment-specific utility benefit, this decreased using 5L and so in isolation of other utility changes, the incremental QALYs decreased.

For the 4 non-oncology case studies with no life extension, when mapped 5L utility values were used, using either the EQG or FORWARD datasets, the incremental QALYs decreased and ICERs increased compared to 3L, in some cases above the range of maximum acceptable ICERs specified in the NICE Guide to the Methods of Technology Appraisal. This is because the benefit was derived solely from the difference in utility values, and this decreased for 5L where all utility values increased.

For the 6 non-oncology case studies with life extension, the direction of change in incremental QALYs and ICERs varied between the studies.

Where interventions both avoid or delay progression and increase survival, there is a trade-off between the increase in incremental QALYs from increasing survival and the decrease in incremental QALYs from reducing the benefit of delayed progression (through decreasing the difference between best and worst states). It is therefore difficult to predict what the change in incremental QALYs and ICER will be upfront. Future changes to NICE policy need to be aware of this information in order to ensure decision making is consistent, fair and reflects scientific state of the art.

1. INTRODUCTION

Quality-adjusted life years (QALYs) combine a measure of health-related quality of life with length of life to produce a measure that allows comparison across disease areas. An intervention can generate more QALYs than its comparator by improving quality of life or improving survival, or improving both. The National Institute for Health and Care Excellence (NICE) states that health effects should be expressed in QALYs in its Guide to the Methods of Technology Appraisal[1].

Incremental cost-effectiveness ratios (ICERs) present the ratio of the expected additional total cost to expected additional QALYs for an intervention compared with alternative treatment. When the additional total cost is fixed, the ICER for an intervention decreases when the QALY gain increases, and vice-versa. NICE considers a range of maximum acceptable ICERs in deciding whether an intervention is cost-effective[1].

NICE states that the preferred measure of health-related quality of life in adults is the EuroQol 5-Dimension instrument (EQ-5D)[1]. The EQ-5D asks a person to describe their health across five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression[2]. Two descriptive systems of the EQ-5D exist: the 3-level (3L) and 5-level (5L): the 3L has 3 levels of severity for each dimension and the 5L has 5. Valuation sets exist for both the 3L and 5L which assign a utility value (measure of health-related quality of life) to each health state described by the different levels of severity for each dimension. Currently, in an interim position statement due to be revisited in 2018, NICE does not recommend the 5L valuation set and recommends that the 3L valuation set should be used for reference-case analysis[3]. There is a perception that 5L may be superior to 3L as the increase in levels increases sensitivity between health states[4-6].

The development of methods to map from 3L to 5L (and vice versa) has been described and validated in previous Decision Support Unit (DSU) reports[7,8] and peer reviewed articles[9]. We have previously performed analyses of the impact of using 5L in economic evaluations instead of 3L. Those analyses were from studies where 3L responses were reported at the individual patient level (usually within clinical trials)[7]. At the time of those analyses, the ability to map from 5L to 3L, or vice versa, required that estimates were performed on the responses given to the descriptive system of the 5L. Whilst the 9 case studies used in that report were from a wide

variety of technology and disease types, they were not specifically selected to represent the experiences of the NICE Technology Appraisal (TA) Programme.

The majority of NICE TAs use model-based economic evaluations where the analyst has no access to patient level utility value data used to populate the model, but instead must base the analysis on summary statistics reported in existing literature, from previous NICE appraisals or from other sources.

A command to map from 3L to 5L (and vice versa); EQ5Dmap; has been developed in Stata and the accompanying paper describing how this works in detail has been accepted for publication in the Stata Journal[10]. This enhanced mapping functionality now allows us to investigate the likely impact of moving from 3L to 5L in cost-effectiveness models used in the NICE TA programme.

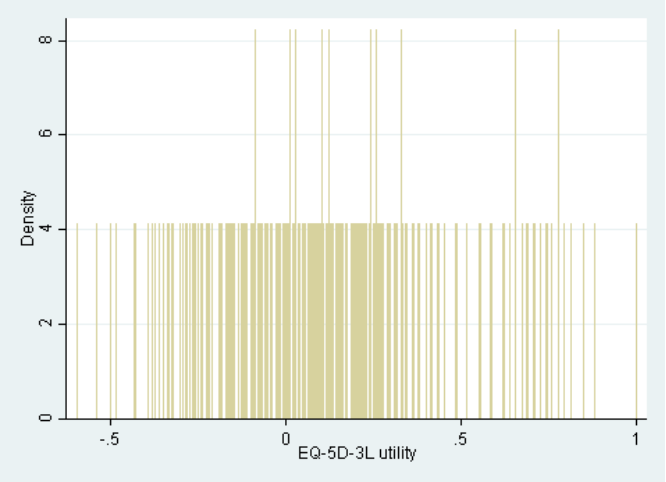
This report first describes the key features of the Stata mapping function and how we recommend it be used in cost-effectiveness analyses. We then use the mapping command in 20 NICE appraisals to estimate the impact of moving from 3L to 5L on cost-effectiveness estimates.

2. MAPPING BETWEEN 3L AND 5L

Here, we describe how the mapping command can be used for mapping from 3L to 5L in those situations where the analyst simply has an estimate of the mean 3L utility value for a health state, as is typically the case for model based cost-effectiveness analyses. In this situation, the utility value does not relate to a single health state description from the 243 health states that can be described by the 3L, but is made up from an unknown distribution of responses. The mapping command uses a distance-weighted averaging of values around the mean 3L value to estimate the 5L utility value. The bandwidth parameter simultaneously defines the neighbourhood over which averaging is to be done and the rate at which the weight declines with increasing distance from the mean 3L value. Selection of bandwidth parameters is always important, even more so in this case because of the characteristics of the 3L valuation distribution (see Figure 1). There are several gaps in the 3L distribution of values. These gaps differ in their size at different parts of the 3L distribution. For example, the gap between 1 (full health) and the next feasible state at 0.883 is the largest. Furthermore, in real 3L datasets, not

all feasible 3L states are used by respondents. This can be where certain combinations are not relevant to patients with that condition, or where sample sizes are small. This means that the same bandwidth will not necessarily be appropriate for all mean 3L values. The standard practice in these cases is to use an adaptive bandwidth, however, this is not possible here because the 3L distribution is sample specific and we rarely have information which allows us to identify it.

Figure 1:Histogram of EQ-5D-3L Utility Scores

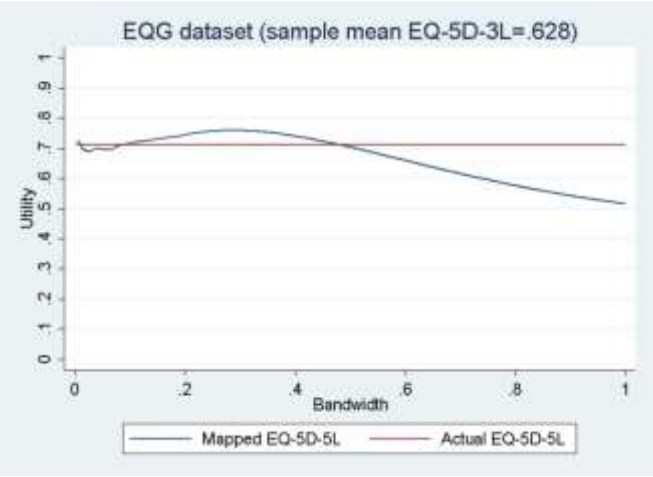


To develop guidance on the choice of bandwidth, we used two datasets that contain large numbers of respondents who completed both 3L and 5L, both of which are fully described in previous DSU reports[8]. These are the EuroQoL Group (EQG) dataset and the US-based FORWARD databank (The National Databank for Rheumatic Diseases).

The EQG sample (n=3,539)[1], has a mean 3L value of 0.628. Figure 2 shows that the true mean 5L from the same sample is 0.712 (red line). The blue line plots the estimated 5L for different bandwidth choices. This shows that, in this case, small bandwidths work relatively well, but that the predicted 5L score slightly overestimates as the bandwidth increases. At a bandwidth of approximately 0.4 the estimated and actual values are equal. Larger bandwidths lead to increasingly inaccurate estimated 5L values. This is because the use of such large bandwidths for relatively high values of 3L results in the bandwidth exceeding the feasible range for 3L values at one end (beyond values of 1). Therefore, the weighted average tends to include a much larger set of values below the mean 3L than above, distorting the mean downwards.

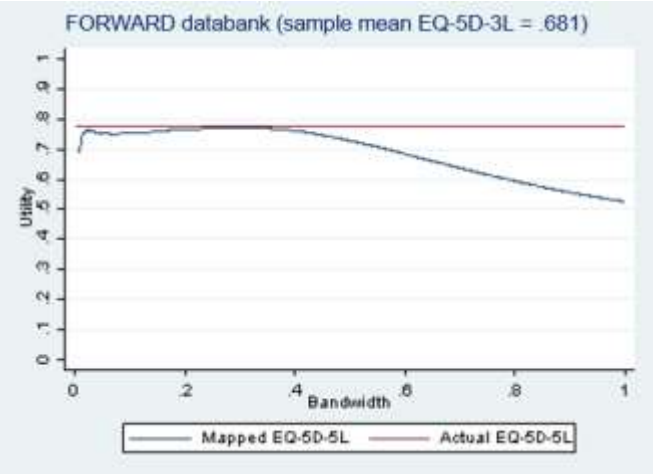
[1] The sample excludes those individuals with missing values on age, gender, 3L or 5L. It also excludes a small number of individuals under 16 years of age.

Figure 2: EQG mapped and actual score for different bandwidths



The same procedure was repeated in the FORWARD databank (Figure 3). The mean 3L value is higher at 0.681 and the corresponding 5L data also have a higher mean at 0.778 (n=5,192). Again, it can be seen that where the bandwidth is small enough that it does not lead the procedure into the non feasible utility value area (above 1) the estimated and observed values are very close. In this case a bandwidth below 0.32 is appropriate.

Figure 3: FORWARD mapped and actual score for different bandwidths



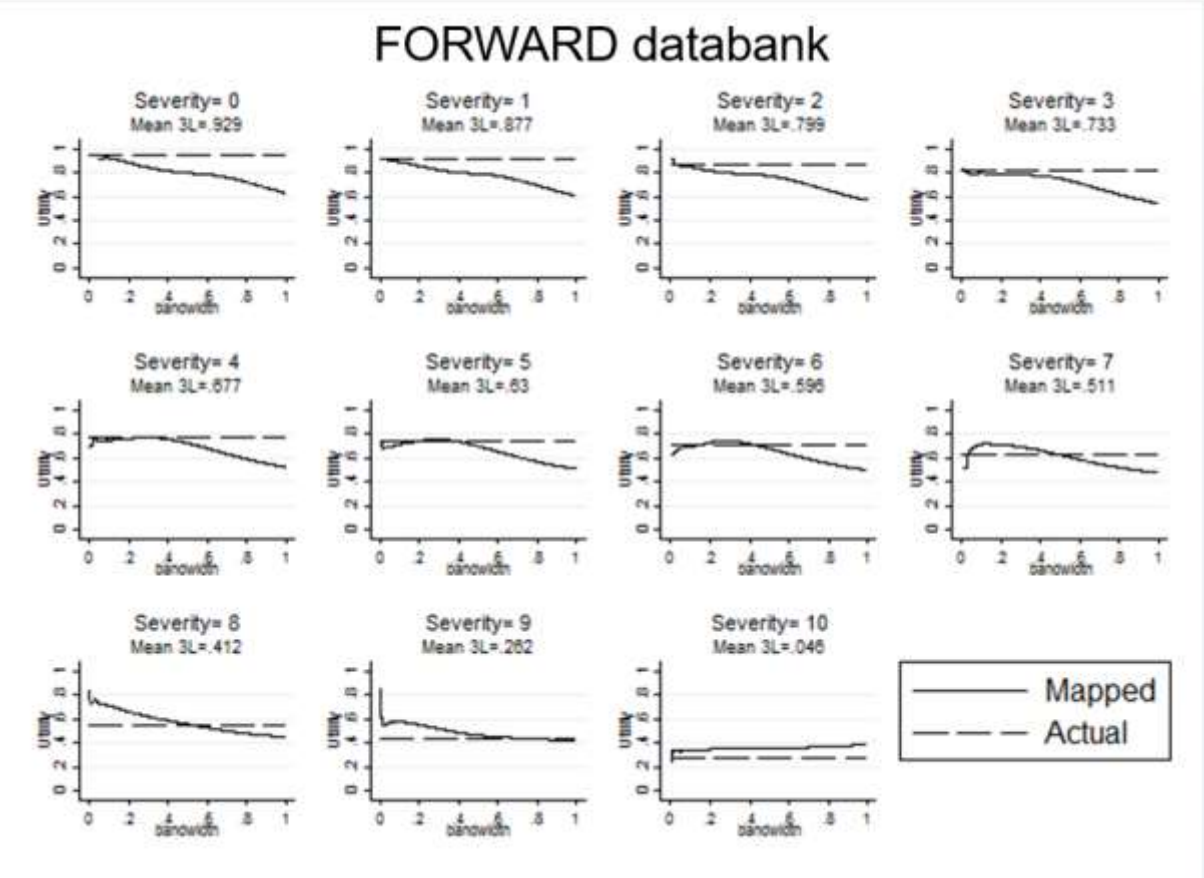
The FORWARD databank allows for further investigation by subgroups. The data were categorised by the patient global assessment of disease severity, a visual analogue scale response from 0-10.

Figure 4 presents the same plots of observed versus estimated mean 5L values for different bandwidth choices by disease severity. The following can be inferred from them:

1) For a 3L utility value in the gap between full health and the next feasible 3L value (0.883), the bandwidth should be just large enough to reach 1.

2) For 3L utility values in the range 0.883 to 0.7 (inclusive): smaller bandwidths are preferable and should not exceed 0.1.
3) For 3L utility values below 0.7 to 0.6: bandwidths of 0.2 are recommended.
4) For 3L utility values less than 0.6, larger bandwidths are required. We recommend values of 0.4.

Figure 4: FORWARD mapped and actual score for different bandwidths by disease severity



3. METHODS AND DATA

3.1. CHOICE OF ECONOMIC MODELS

We selected 20 case studies of models used in NICE TA decision making, which represented a range of disease areas reflective of those typically considered by NICE. Our case studies covered oncology, circulatory system disease, musculoskeletal diseases, infections, nervous system diseases, skin diseases, digestive system disorders, mental and behavioral disorders, and eye diseases. We considered that this reflected a pragmatic cross section of the types of interventions considered by NICE, and sufficiently large to draw general conclusions. All of the economic models used in the case studies were based predominantly (but not exclusively) on 3L utility values. The selection of case studies was discussed with the NICE project team,

such that a range of people with experience of NICE appraisals could suggest case studies for inclusion.

The case studies were drawn from TAs where final guidance had been produced, ICERs were reported in the guidance, and the role these ICERs played in developing recommendations was clear. Seventeen of the 20 case studies were from Single Technology Appraisals (STAs) in which we focused on the intervention under appraisal and the comparison reported in the final guidance. Two cases studies were from Multiple Technology Appraisals (MTAs): in one case only one comparison was considered as the treatment dominated other comparators; and in the other case we considered three interventions, drawing pairwise comparisons between the least effective and next least effective, and the latter with the most effective. Additionally, we considered one case study from the Highly Specialized Technology (HST) programme. Therefore, we considered 21 comparisons. We negotiated access to the models for via NICE and via the Assessment Group based at the University of Sheffield.

Where possible, we used the model settings which generated the ICERs reported as the most plausible scenarios or committee's preferred assumptions in the Final Appraisal Determination (FAD). However, in some cases, these ICERs were not based on EQ-5D utility values, so in these cases we used the same settings apart from 3L utility values (for example TA274). In some cases, we were unable to reproduce the exact ICERs reported in the FAD, but we used the same settings and data as far as these could be determined, to produce ICERs very similar to those considered by the committee.

3.2. MAPPING MODELS

We used the EQ5Dmap command in Stata[10]. We used both the EQG and FORWARD–based versions of the mapping algorithm. We used the copula models. Where the mean 3L score was < 0.6, we used a bandwidth of 0.4. Where the 3L score was ≥ 0.6 and < 0.7, we used a bandwidth of 0.2. Where the 3L score was ≥ 0.7, we used a bandwidth of 0.05.

4. COST EFFECTIVENESS CASE STUDIES

4.1. METHODS AND CASE STUDY DESCRIPTIONS

The models considered were all cohort models with health states – either partitioned survival analyses or state-transition models. The source of utility values in each model was primarily

3L, and these utility values could have been measured directly in trials, mapped from other measures to 3L, or sourced from published literature. Where the source of utility values was clearly 3L or another measure mapped to 3L, we mapped the utility values to 5L. Where the source was not 3L, or was not clearly reported, we did not map the utility values – this was particularly common for adverse events. We explored the impact of this in scenario analyses.

Where models considered 3L utility values for health states, we mapped the 3L utility value to 5L and used this directly in the model. Where models considered utility decrements ('disutilities'), we mapped the original 3L values with and without the disutility to recalculate the disutility. Where models considered utility multipliers, we mapped the two original 3L values to recalculate the multiplier.

To inform the age and sex variables in the mapping, we used the mean age and proportion of males reported in the source study (trial or literature) where this information was available. Where this information was not available, we used the starting age and proportion of males in the economic model.

A summary of the case study models is presented in Table 1. There are 10 partitioned-survival analysis models and 10 cohort state-transition models. Three case studies only used 3L data collected in clinical trials for the intervention (TA427, TA392 and TA274). Four case studies used 3L data collected in clinical trials for the intervention for health states and values from the literature for adverse events (TA391, TA428, TA365 and TA367). Two case studies used 3L data collected in clinical trials and adjusted for age using general population utility values from the literature (TA360 and TA401). Two case studies used 3L data collected in clinical trials, adverse event data from the literature and adjustment for age from the literature (TA428 and TA457). Three studies relied solely on utility value data reported in the literature or previous NICE appraisals (TA228, TA335 and TA363). The remaining case studies used a combination of clinical trial utility values and utility values from the literature. All case studies are referred to by their TA number, with the exception of one case study in Multiple Sclerosis (MS). In this case study, the company provided permission for the model to be used but requested that it be anonymized. To prevent identification of this case study, it is referred to as 'MS intervention for treating relapsing-remitting multiple sclerosis' and information which could be used to identify the case study has been redacted.

Table 1: Summary of case study models

TA	Model structure	Classification of utilities	Source of health state utilities	Utilities mapped to 5L?
TA391 Cabazitaxel for hormone-relapsed metastatic prostate cancer treated with docetaxel	Partitioned-survival analysis	Baseline	Clinical trial cabazitaxel (EQ-5D-3L)	Yes
		Progressive disease		Yes
		Cycles 2, 4, 6, 8, 10		Yes
		Cycles 1, 3, 5, 7, 9	Interpolated from even cycles	Interpolated from mapped values for even cycles
		Stable disease	Same as cycle 10	Same as cycle 10 mapped value
		Adverse event disutilities	Literature	Left unchanged
TA381 Olaparib for ovarian, fallopian tube and peritoneal cancer	Partitioned-survival analysis	Progression-free	Clinical trial olaparib (FACT-O mapped to EQ-5D-3L)	Yes
		progressed-disease	Clinical trial trabectedin (EQ-5D-3L), used in NICE TA222 and TA389	Yes
TA316 Enzalutamide for metastatic hormone-relapsed prostate cancer previously treated with a docetaxel-containing regimen	Partitioned-survival analysis	Stable disease	Clinical trial enzalutamide (EQ-5D-3L)	Yes
		Progression disutility	Literature (disutility for progression calculated using published utility values according to months before death)	Yes
		Utility gain for enzalutamide	Clinical trial enzalutamide (FACT-P mapped to EQ-5D-3L)	Yes
		Skeletal related event disutilities	Clinical trial enzalutamide (FACT-P mapped to EQ-5D-3L)	Yes
		Adverse event disutilities	Literature	Left unchanged

TA	Model structure	Classification of utilities	Source of health state utilities	Utilities mapped to 5L?
TA377 Enzalutamide for treating metastatic hormone-relapsed prostate cancer before chemotherapy is indicated	Partitioned-survival analysis	Stable disease	Clinical trial enzalutamide (EQ-5D-3L)	Yes
		Utility gain for enzalutamide in stable disease		Yes
		Post-progression 1	Literature (EQ-5D-3L)	Yes
		Post-progression 2	Clinical trial enzalutamide (EQ-5D-3L) – stable disease in TA316	Yes
		Utility gain for enzalutamide in post-progression 2	Clinical trial enzalutamide (FACT-P mapped to EQ-5D-3L) – utility gain in TA316	Yes
		Palliative care	Literature (EQ-5D-3L)	Yes
		Adverse events	Literature	Left unchanged
TA428 Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy	Partitioned-survival analysis.	Progression-free: >=30 days to death, <30 days to death	Clinical trial pembrolizumab (EQ-5D-3L)	Yes
		Progressed disease: >=30 days to death, <30 days to death		Yes
		Age specific disutilities	General population (EQ-5D-3L)	Yes
		Adverse events	Literature	Left unchanged
TA427 Pomalidomide for multiple myeloma previously treated with lenalidomide and bortezomib	Partitioned-survival analysis.	Regression equation. Utility values are therefore treatment- and cycle-specific.	Analysis of patient-level EQ-5D-3L data in pomalidomide clinical study.	Yes

TA	Model structure	Classification of utilities	Source of health state utilities	Utilities mapped to 5L?
TA228 Bortezomib and thalidomide for the first-line treatment of multiple myeloma	Partitioned-survival analysis	Pre-progression: on treatment	Literature (EORTC QLQ-30 mapped to EQ-5D-3L)	Yes
		Pre-progression: on treatment		Yes
		Pre-progression: post-treatment and complete response		Yes
		Pre-progression: post-treatment and not complete response		Yes
		Post-progression		Yes
TA357 Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab	Partitioned-survival analysis.	>180 days to death, 120-179 days to death, 90-119 days to death, 60-89 days to death, 30-59 days to death, <30 days to death	Clinical trial pembrolizumab (EQ-5D-3L)	Yes
		Adverse event disutilities	Literature	Left unchanged
		Age-specific disutilities	General population (EQ-5D-3L)	Yes
TA366 Pembrolizumab for advanced melanoma not previously treated with ipilimumab	Partitioned-survival analysis.	>=360 days to death, 270-359 days to death, 180-269 days to death, 90-179 days to death, 30-89 days to death, <30 days to death	Clinical trial pembrolizumab (EQ-5D-3L)	Yes
		Age-specific disutilities	General population (EQ-5D-3L)	Yes
TA401 Bosutinib for previously treated chronic myeloid leukaemia	Partitioned-survival analysis.	Chronic phase: on bosutinib	Clinical trial bosutinib (EQ-5D-3L)	Yes
		Chronic phase: on hydroxycarmabide		
		Accelerated phase		
		Blast phase		
		Age-specific utilities	General population (EQ-5D-3L)	Yes

TA	Model structure	Classification of utilities	Source of health state utilities	Utilities mapped to 5L?
TA335 Rivaroxaban for preventing adverse outcomes after acute management of acute coronary syndrome	Cohort state-transition model.	No event	Clinical trial ticagrelor (EQ-5D-3L)	Yes
		MI 1st 6 months, MI 2nd 6 months, MI later (post 12 months)	Clinical trial ticagrelor (EQ-5D-3L)	Yes
		IS 1st 6 months	Clinical trial ticagrelor (EQ-5D-3L)	Yes
		IS 2nd 6 months, IS later (post 12 months)	Clinical trial ticagrelor (EQ-5D-3L), with adjustment from literature (EQ-5D-3L)	Yes
		HS 1st 6 months	Same as IS 1st 6 months	Same as mapped IS 1st 6 months
		HS 2nd 6 months	Same as IS 2nd 6 months	Same as mapped IS 2nd 6 months
		HS later (post 12 months)	Same as IS later (post 12 months)	Same as mapped IS later (post 12 months)
		PCI/PTCA, CABG, TIMI Major bleeding, TIMI Minor bleeding, TIMI requiring medical attention	Literature (unclear if EQ-5D-3L)	Left unchanged
TA363 Ledipasvir–sofosbuvir for treating chronic hepatitis C	Cohort state-transition model.	Non-cirrhotic	Literature (EQ-5D-3L)	Yes
		Cirrhotic		Yes
		SVR increment		Yes
		Compensated cirrhosis		Yes
		Decompensated cirrhosis		Yes
		Hepatocellular carcinoma		Yes
		Liver transplant	Same as decompensated cirrhosis	Same as mapped decompensated cirrhosis
		Post-liver transplant	Literature (EQ-5D-3L)	Yes

TA	Model structure	Classification of utilities	Source of health state utilities	Utilities mapped to 5L?
MS Intervention for treating relapsing-remitting multiple sclerosis	Cohort state-transition model.	RRMS EDSS 0-9 no relapse	Information redacted	Yes
		SPMS disutility	Literature (EQ-5D-3L)	Scenario analysis
		Relapse disutility		Scenario analysis
		Carer disutility	Literature (unclear if EQ-5D-3L)	Scenario analysis
		Adverse events	Literature	Scenario analysis
TA325 Nalmefene for reducing alcohol consumption in people with alcohol dependence	Cohort state-transition model.	First year of treatment: baseline, week 12, week 24, week 36, week 52	Clinical trial nalmefene (EQ-5D-3L)	Yes
		Years 2-5: high/very high risk, medium, low/abstinent risk		Yes
		Harmful event disutilities	Health outcomes data repository (EQ-5D-3L)	Recalculated by mapping general population and disease utilities
TA279 Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral compression fractures	Cohort state-transition model.	Years 1-3: by treatment over time.	Clinical trials vertebroplasty and percutaneous balloon kyphoplasty (VAS mapped to EQ-5D-3L)	Yes
		General population utility values	General population (EQ-5D-3L)	Yes
		Hip fracture multiplier year 1 and year 2+, Verbetral fracture multiplier year 1 and 2+	Literature (EQ-5D-3L)	Recalculated by mapping utilities with and without fractures.
TA392 Adalimumab for treating moderate to severe hidradenitis suppurativa	Cohort state-transition model.	High response	Clinical trial adalimumab (EQ-5D-3L)	Yes
		Response		Yes
		Partial response		Yes
		Non-response		Yes

TA	Model structure	Classification of utilities	Source of health state utilities	Utilities mapped to 5L?
TA352 Vedolizumab for treating moderately to severely active Crohn's	Cohort state-transition model.	Remission	Clinical trial vedolizumab (EQ-5D-3L)	Yes
		Mild disease		Yes
		Moderate-to-severe disease		Yes
		Surgery	Same as Moderate-to-severe disease.	Same as mapped Moderate-to-severe disease.
		Malignancy	Clinical trial vedolizumab (EQ-5D-3L)	Yes
		Adverse events	Literature	Left unchanged
TA367 Vortioxetine for treating major depressive episodes	Cohort state-transition model.	Baseline	Clinical trial vortioxetine (EQ-5D-3L)	Yes
		Remission		Yes
		Response/Recovery		Yes
		No response		Yes
		Adverse events	Literature	Left unchanged
TA274 Ranibizumab for treating diabetic macular oedema	Cohort state-transition model.	Best corrected visual acuity: 86-100 letters, 76-85 letters, 66-75 letters, 56-65 letters, 46-55 letters, 36-45 letters, 26-35 letters, <25 letters.	Clinical trial ranibizumab (EQ-5D-3L)	Yes
HST2 Elsosulfase alfa for mucopolysaccharidosis IVA	Cohort state-transition model.	Asymptomatic	Burden of disease study (EQ-5D-5L descriptive system, valued using 3L crosswalk)	Yes
		No use of wheelchair		Yes
		Sometimes use wheelchair		Yes
		Wheelchair dependent		Yes
		Paraplegic		Yes
		End-stage		Yes
		Treatment increment		Yes
		Surgery-related disutility	Expert opinion	Left unchanged
		Caregiver disutility	Literature	Left unchanged

4.1.1. Oncology case studies

All models were partitioned-survival analyses, with health states defined as pre-progression/stable disease, progressed disease and death. In five case studies, utility was defined by progression (TA391, TA381, TA316, TA377 and TA228). In two case studies, utility was defined by time until death and not progression (TA357 and TA366). In one case study, utility was defined by progression and time to death (TA428). In one case study, utility was calculated using a regression equation which included progression (TA427). In one case study, utility was defined by disease phase (TA401). Specific details of each analysis are provided in the following subsections.

4.1.1.1. TA391 Cabazitaxel for hormone-relapsed metastatic prostate cancer treated with docetaxel

The utility values are from 3L questionnaires from the UK Early Access Programme (EAP) for cabazitaxel. The UK EAP provided utility data for the stable disease states at baseline, cycle 2, 4, 6, 8 and 10. Utility values for stable disease cycles 1, 3, 5, 7 and 9 were interpolated. The utility value for stable disease cycle 10 is used for cycles beyond 10. The UK EAP also provided a utility value for the progressive state which was the value recorded 30 days after the last cycle of treatment for patients with evidence of progression[11].

We assumed the age was 67 (mean of the UK EAP) and all observations were from males. We interpolated utility values from the even cycles to obtain utility values for the odd cycles, using the same method as the company.

4.1.1.2. TA381 Olaparib for ovarian, fallopian tube and peritoneal cancer

The utility value for the progression-free state was estimated by mapping from the FACT-0 questionnaire included in the clinical trial to 3L using a published mapping algorithm[12]. The utility value for the progressed state was taken from previous NICE appraisals which use 3L from the OVA-301 trial[13].

We mapped the 3L values for progression free and progressed-disease. We assumed that these are all for a 57-year-old person (starting age in the model), and that all are female.

4.1.1.3. TA316 Enzalutamide for metastatic hormone-relapsed prostate cancer previously treated with a docetaxel-containing regimen

The utility for stable disease is from analysis of AFFIRM 3L data[14]. The disutility for progression is from analysis of 3L data in Sandblom et al. 2004[15]. There are utility decrements for skeletal related events and an increment for treatment with enzalutamide, from analysis of FACT-P mapped to EQ-5D in AFFIRM.

We added the decrements and increments from AFFIRM to the stable disease state value and mapped these, with the stable disease state value, to 5L using a mean age of 69 (median in AFFIRM). We then recalculated the increments and decrements from the 5L data. We mapped the Sandblom et al. 2004[15] utilities using the mean age in that study, and recalculated the disutility for progression. Although there were utility decrements for adverse events, the sources of these were not necessarily 3L, and some lacked information, so we did not map these.

4.1.1.4. TA377 Enzalutamide for treating metastatic hormone-relapsed prostate cancer before chemotherapy is indicated

The progressed disease state is split into three states: post-progression 1, post-progression 2 and palliative care. The utility value for stable disease is from analysis of PREVAIL 3L data[16]. There is a utility gain from enzalutamide from analysis of PREVAIL EQ-5D data. The utility value for post-progression 1 state is from 3L in the literature[17,18] and there is a 3L utility value from Sandblom et al. 2004 for palliative care[15]. The utility value for post-progression 2 is from TA316 and there is a utility increment for enzalutamide post-chemotherapy from TA316.

We mapped the utility value for stable disease with and without the enzalutamide increment, and recalculated the increment, using the mean age of 72 (from PREVAIL[16]). We mapped the utility value for post-progression 1 using a mean age of 72[17,18] and for palliative care using a mean age of 69[15]. We used the mapped values from TA316[14] where appropriate. Although there were utility decrements for adverse events, the sources of these were not necessarily EQ-5D, and some lacked information, so we did not map these.

4.1.1.5. TA428 Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy

3L utility data is analyzed from the KEYNOTE-010 study, and is categorized by progression-free or progressed and by time to death[19]. Utility values are age-adjusted using UK population average utilities from Kind et al. 1999[20].

We mapped the utility values from KEYNOTE-010 and the UK population averages from 3L to 5L. For the KEYNOTE-010 utilities we used an age of 62 (mean age in the model), and 61.4% male. For the UK population utility values, we used the mean age in each category, and mapped male and female utility values separately, which are then combined in the model.

4.1.1.1. TA427 Pomalidomide for multiple myeloma previously treated with lenalidomide and bortezomib

A regression equation has been derived from MM-003 patient level 3L data to calculate utility values[21]. The equation has an intercept term and coefficients for disease progression, best overall response rate at week 12, hospitalization, adverse events, gender, baseline ECOG score, baseline multiple myeloma stage, whether patient is European and baseline red blood cell level. The utility values for the cohort change over time as the incidence of adverse events (AEs) and hospitalisations change, so we mapped the utility values in each cycle. We used the mean age of 68 (model starting age) and 59.93% male.

4.1.1.2. TA228 Bortezomib and thalidomide for the first-line treatment of multiple myeloma

Pre-progression is split into two sub-states of treatment and post-treatment[22]. The utility value in the treatment state is from a mapping from the European Organisation for Research and Treatment of Cancer core quality of life questionnaire (EORTC QLQ-C30) to 3L by McKenzie and van der Pol[23], using the utility value for month 1. Post-treatment, people who have a complete response have the utility value for those with complete response in MMIX RCT, when EQORTC QLQ-C30 was mapped to 3L using the same algorithm as McKenzie and van der Pol.

The utility value for post-progression, and for post-treatment for people without a complete response is based on the average of the 6 month to 36-month time points from McKenzie and van der Pol.

For the McKenzie and van der Pol utilities, we used an age of 75 (mean age in the study), 70.2% male[23]. For the MMIX utility, we used a mean age of 74, and 55.7% male (weighted average of the two arms in the study)[22].

4.1.1.3. TA357 Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab

In the manufacturer base case, utility values are defined by time to death rather than by health state. They are calculated from analysis of 3L in KN002[24]. We mapped these to 5L, using a mean age of 60 and 60.7% male, from KN002.

Adverse event disutilities are applied, extracted from a study that used standard gamble[25]. Since these are not EQ-5D, we do not map them to 5L.

There is an annual decrement for aging, calculated from Kind et al. 1999 utility values for the general population[20]. We map these by age and gender, and then recalculate the decrement.

4.1.1.4. TA366 Pembrolizumab for advanced melanoma not previously treated with ipilimumab

In the manufacturer base case, utility values are defined by time to death rather than by health state. They are calculated from analysis of 3L in KN006[26]. We mapped these to 5L, using a mean age of 60 and 59.6% male, from KN002. There is an annual decrement for aging, calculated from Kind et al. 1999 utility values for general population[20]. We mapped these by age and gender, and then recalculated the decrement.

4.1.1.5. TA401 Bosutinib for previously treated chronic myeloid leukaemia

In the manufacturer base case, utility values are determined by whether the patient is in chronic, accelerated or blast phase and by the treatment the patient receives in chronic phase. The utility values were derived from the pivotal bosutinib study (Study 200)[27]. We mapped these to 5L, using a mean age of 50 (the mean age in Study 200) and 50% male (the distribution in the model).

The model considers age-adjusted utility values, by using general population utilities from Kind et al. 1999 [20]. and applying the health state utilities as utility multipliers. We mapped the Kind et al. 1999 utilities by age and gender, and recalculated the utility multipliers.

4.1.2. Non-oncology case studies with a life year gain

These models used a range of health states appropriate to the specific indication to model progression. More detail is provided about each model in the subsections.

4.1.2.1. TA335 Rivaroxaban for preventing adverse outcomes after acute management of acute coronary syndrome

The model is a state-transition model, with states defined by acute coronary syndrome events that may occur and whether those events had long term implications[28]. The health states, and the sources of the utility values are shown in Table 2.

Table 2: TA335 utility values

Health state	Utility value	Source
No event	**0.8420**	**TA236[29]**
MI 1st 6 months	**0.7790**	**TA236[29]**
MI 2nd 6 months	**0.8210**	**TA236[29]**
MI later (post 12 months)	**0.8210**	**TA236[29]**
IS 1st 6 months	**0.7030**	**TA236[29]**
IS 2nd 6 months	**0.7476**	**TA236[29], with adjustment based on Ara and Brazier 2010[30]**
IS later (post 12 months)	**0.7921**	**TA236[29], with adjustment based on Ara and Brazier 2010[30]**
HS 1st 6 months	**0.7030**	**TA236[29]**
HS 2nd 6 months	**0.7476**	**TA236[29], with adjustment based on Ara and Brazier 2010[30]**
HS later (post 12 months)	**0.7921**	**TA236[29], with adjustment based on Ara and Brazier 2010[30]**
PCI/PTCA	**0.7920**	**Latour-Perez 2008[31]**
CABG	**0.7420**	**Latour-Perez 2008[31]**
TIMI Major bleeding	**0.7500**	**Crespin *et al.* 2011[32]**
TIMI Minor bleeding	**0.8000**	**Kazi *et al.* 2014[33]**
TIMI requiring medical attention	**0.8000**	**Sullivan *et al.* 2006[34]**

CABG, coronary artery bypass graft; HS, hemorrhagic state; IS, ischemic stroke; PTCA/PCI, Percutaneous transluminal coronary angioplasty/ Percutaneous coronary intervention; TIMI, Thrombolysis in Myocardial Infarction

The TA236 utility values were from 3L collected in the PLATO-HECON study[29]. These were used for the no event, Myocardial Infarction (MI) for 1st 6 months, 2nd 6 months and post 12 months, and Ischemic Stroke (IS) 1st 6 months states. To calculate the utility value for IS post 12 months, the manufacturer calculated the improvement in utility value 12 months after a stroke from a study by Ara and Brazier in 2010[30]. Ara and Brazier reported the mean 3L score

for "Stroke < 12 months, history of stroke + other CV condition" as 0.479 and for "No event < 12 months, history of stroke + other CV condition" as 0.641[30]. The manufacturer calculated that the improvement was (0.641 – 0.479)/0.479 = 33% and so increased the TA236 IS 1st 6 months' utility value by 33% to estimate the IS post 12 months' utility value. To calculate the utility value for the IS 2nd 6 months' state, the manufacturer took the midpoint of the IS 1st 6 months and IS post 12 months' utility values.

The manufacturer assumed that the utility values for the HS states were the same as for the IS states.

It is unclear whether the original source of the utility values for the other health states is EQ-5D[31-34].

We mapped the TA236 utility values from 3L to 5L, assuming they are for a 62 year old person (the starting age in the model) and that 75% of people were male (mean age at baseline).

We mapped the Ara and Brazier (2010) utilities, assuming they are for a 74 and 70 year old person (the mean age of the samples in the paper)[30] and that 75% of people were male.

We used the mapped 5L Ara and Brazier utilities to calculate the improvement to apply to the mapped TA236 utility values.

We did not map the other health state utility values to 5L as it was not clear if they were 3L.

4.1.2.2. TA363 Ledipasvir–sofosbuvir for treating chronic hepatitis C

The model is a state-transition model, with health states defined by sustained virologic response (SVR), treatment, cirrhosis (and level), hepatocellular carcinoma (HCC) and liver transplant[35]. Utility values are 3L, taken from published sources used in previous economic evaluations for hepatitis C[36,37].

Wright et al. 2006 report the baseline utility value without cirrhosis and without SVR for patients in the treatment arm as 0.75[36]. The mean age at baseline in the study was 41 and 64.3% were male. Wright et al. report utility values for cirrhosis as 0.55, from a sample with mean age 47, and 73% male – so we use these characteristics for cirrhosis. Wright et al. report a utility value of 0.45 for decompensated cirrhosis and HCC, and 0.67 for post-transplant, both referenced to Ratcliffe et al. 2002[38], but these appear to be subgroup analyses not reported in the paper. We therefore assume the same age and gender characteristics as for cirrhosis. Vera-Llonch et al. 2013 report the utility increment for SVR is 0.041[37]. This is calculated from multivariate regression of 3L data. To calculate the increment in our analysis, we map 0.791 (0.75 + 0.041) to 5L, and subtract the 5L value from mapping 0.75. We assume the same age and gender characteristics as for baseline without cirrhosis.

4.1.2.3. MS Intervention for treating relapsing-remitting multiple sclerosis

The model is a state-transition model, with states defined by Expanded Disability Status Score (EDSS), Relapsing Remitting Multiple Sclerosis (RRMS) or Secondary Progressive Multiple Sclerosis (SPMS), and whether the patient is in Relapse. Each health state has an associated utility value. Utility values are defined separately for RRMS and SPMS EDSS states.

To calculate the utility values for the SPMS and Relapse states, the manufacturer used the relationship between EDSS, RRMS, SPMS and Relapse states in the UK MS Survey by subtracting decrements. The UK MS Survey used regression analysis on 3L data to calculate coefficients for utility values by EDSS, relapse and RRMS or SPMS. It was therefore not appropriate to map the utility values derived from the UK MS Survey to 5L (as they were not original 3L utility values) and not possible to map the raw data to 5L and rerun the regression.

The model includes disutilities for a number of adverse events. The source of the disutilities vary.

The model additionally includes a disutility for carers, which varies by EDSS and is sourced from studies in Alzheimer's. It is unclear whether the increments are from EQ-5D, and how they are calculated.

To understand the uncertainty associated with the utility decrements associated for relapse and SPMS, adverse events, and carers we analyzed four scenarios:

1. The RRMS 3L utility values are mapped to 5L. Original utility values are used for relapse, SPMS, adverse event and carer disutilities.
2. The RRMS 3L utility values are mapped to 5L. Disutilities for SPMS and relapse are mapped to 5L by mapping the UK MS survey utility values and calculating the differences between RRMS/SPMS and Relapse/No Relapse states. Original values are used for adverse event and carer disutilities.
3. As scenario 2, but adverse event disutilities are also mapped to 5L. Disutilities are calculated by mapping the No relapse RRMS EDSS 0 states with and without each disutility and calculating the difference. Original values are used for carer disutilities.

4. As scenario 3, but carer disutilities are mapped to 5L. Disutilities are calculating by mapping a general population utility value of 0.882, and 0.882 minus each disutility, and calculating the difference.

4.1.2.4. TA325 Nalmefene for reducing alcohol consumption in people with alcohol dependence

The model is a state-transition model[39]. The manufacturer submission describes that utility values are considered in the model as follows[40]:

"The area-under-the-curve utility weights from the ESENSE1, ESENSE2, and SENSE trial are used in the base-case scenario for the first year of treatment. The area between these curves thus represents the mean effect difference between the two treatment strategies from the first year of treatment; this was considered the most sensitive approach to capture the QALY gain between compared interventions from the nalmefene trials. For years 2 to 5 of the base-case analysis, the pooled utilities from the nalmefene trials were used for high/very high, medium, and low/abstinent groups".

The utility values are from 3L questionnaires in the ESENSE1&2 and SENSE trials[39].

The model also includes utility decrements for harmful events. The probability of harmful events occurring is related to the risk groups. The utility decrements are calculated using the difference between the general population utility value and utility values for harmful events. The utility values for harmful events are taken from the alcohol policy model developed by the University of Sheffield which were derived from the Health Outcomes Data Repository which collected 3L data[41].

We assumed the age was 48 (starting age in the model, sourced from ESENSE1&2 and SENSE). We assumed that 69.0% of observations were from males (the proportion of males in the model, sourced from ESENSE1&2 and SENSE), so calculated 5L utility values for a 48-year-old male and 48-year-old female, and calculated an average weighted by the proportion of males[39].

4.1.2.5. TA279 Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral compression fractures

The model is a state-transition model, with health states defined by the initial treatment decision and subsequent hip or vertebral fractures. Utility values for the first three years after treatment are calculated from a mapping function which relates the visual analogue score (VAS) to 3L from a number of trials[42]. We mapped each 3L value to 5L, using the proportion of females

and the baseline age reported in the trial study (increasing age for values after baseline). We then re-estimated the VAS-EQ-5D mapping function and estimated 5L utility values in the first three years.

Utility values beyond three years are informed by population values, from Ara and Brazier[30]. We mapped the utility value for each age group, male and female, to 5L.

Fractures are associated with utility multipliers[43]. We estimate the 5L multipliers by converting the original 3L utility values to 5L, and then recalculating the multipliers. Where the original 3L utility values are not available, we estimate them from available data.

4.1.2.1. HST2 Elsosulfase alfa for mucopolysaccharidosis IVA

The model is a state-transition model, with health states defined by wheelchair usage, symptoms, and end-stage. The utilities for most of the health states are from a burden of disease study[44], which reports that participants completed 5L. However, at the time of publication of the burden of disease study, the valuation set for 5L was not available, and the reported scores include a value of -0.180 for a state valued by 2 patients with reported mean score 5 5 3.5 1 1, which is too low for valuation of that score on 5L[45]. It therefore appears that that 5L scores must have been converted to 3L scores using the cross-walk[46]. The utility value for the asymptomatic state is referenced to Sullivan et al[47] who use 3L, and is reported to be for 0-9 year olds[48], but the model uses a slightly different value with no clear reference. There are treatment-related utility increments within the health states linked to improvements in 6-minute walk test (6MWT) and forced vital capacity (FVC). The increment for 6MWT is referenced from a study which reported an increase in 5L, but which also is likely to have been crosswalked to 3L because of the year of publication and the references are the same as in the other 5L study[49]. The reference for the increment for FVC is unclear, and the increment appears to be 1/10 of the 6MWT increment.

We mapped the health state utility values from the burden of disease study, estimating the mean age from the distribution of ages in the study, and using the proportion of males in the study. For the other health state utility value, we assumed the age was 16 (the minimum age in the mapping command), and used the proportion of males from the burden of disease study. For the 6MWT utility increment, we mapped the mean score of the study (0.552) and the mean plus the increment (0.752) and mapped these to 5L, assuming the mean age from the study (estimated from the distribution of ages) and the proportion of males[49]. We assumed the 5L FVC increment is 1/10 of the 5L 6MWT increment.

We did not map the surgery utility decrements as these were elicited from experts, or the caregiver disutilities as these were not EQ-5D.

4.1.1. Non-oncology case studies with no life year gain

4.1.1.1. TA392 Adalimumab for treating moderate to severe hidradenitis suppurativa

The utility values are from 3L questionnaires in the M11-810 clinical trial. These have been analyzed to produce values for high response, response, partial response and non-response[50]. We assumed the age was 35 (starting age in the model and the median of the pooled M11-810 and M11-313 populations). We assumed that 65.9% of observations were from females (the proportion of females in the intention-to-treat population in the pooled M11-810 and M11-313 trials), so calculated 5L utility values for a 35-year-old male and 35-year-old female, and calculated an average weighted by the proportion of females.

4.1.1.2. TA352 Vedolizumab for treating moderately to severely active Crohn's

Health states are defined as remission, mild disease, moderate-to-severe disease and surgery. Each health state has an associated utility value. In the base case the utility values for remission, mild disease and moderate-to-severe disease are from analysis of 3L data collected in the GEMINI II and GEMINI III trials. The model assumes that the utility value for surgery is equivalent to that of moderate-to-severe disease[51].

The model includes disutilities for several adverse events, from multiple sources. On following up these sources, we discovered that the source of disutility estimate for malignancy is 3L[52]. The disutility estimates for serious infection, acute hypersensitivity reactions and skin site reactions are not 3L[51].

The starting age in the model is 36.57 and 43.9% are male. This is based on pooled data from clinical trials included in the mixed treatment comparison. We mapped the 3L utility values for remission, mild disease, moderate-to-severe disease and malignancy. We assumed that these are all for a 37-year-old person, and that 43.9% are male (by mapping for a male and a female and taking the average according to the proportion male).

4.1.1.3. TA367 Vortioxetine for treating major depressive episodes

The model is a hybrid decision tree – state-transition model, with health states defined by response, remission, and treatment continuation. Each health state has an associated utility value. In the revised analysis, the manufacturer used utility values from the REVIVE trial that

used 3L. Adverse events were also associated with utility decrements, but these were derived from analyses of 3L scores in other studies, so it was not appropriate to directly map the disutilities, and not possible to map the 3L scores from the original analyses to 5L and analyses the new results[53].

We mapped the health state utility values assuming they are for a 43-year-old, and that 74.7% are female (the baseline characteristics in the REVIVE trial).

4.1.1.4. TA274 Ranibizumab for treating diabetic macular oedema

The model is a Markov state-transition mode, with health states defined by visual acuity in the treated eye. Each health state has an associated utility value. The manufacturer used utility values from a published study in which participants valued quality of life using time trade off. Using this utility data, the ICER for treating both eyes in a subgroup of patients with a central retinal thickness of 400micrometres or more was £13,322 (this ICER was calculated by multiplying the ICER for the better-seeing eye model by a factor of 1.5)[54]. However, the manufacturer also included 3Ldata from RESTORE, which gives an ICER for treating both eyes in this population of £30,929 using unadjusted scores. The mean age in RESTORE was 63, and 58% were male[55]. We mapped the RESTORE health state utility vales to 5L using these characteristics.

4.2. RESULTS

The incremental QALYs for intervention versus comparator for 3L, mapped 5L EQG and mapped 5L FORWARD are shown in Figure 5, and the percentage change in incremental QALY is shown in Figure 6. The HST case study is presented on a separate scale. The results for the MS intervention are for Scenario 4. Generally, we see that:

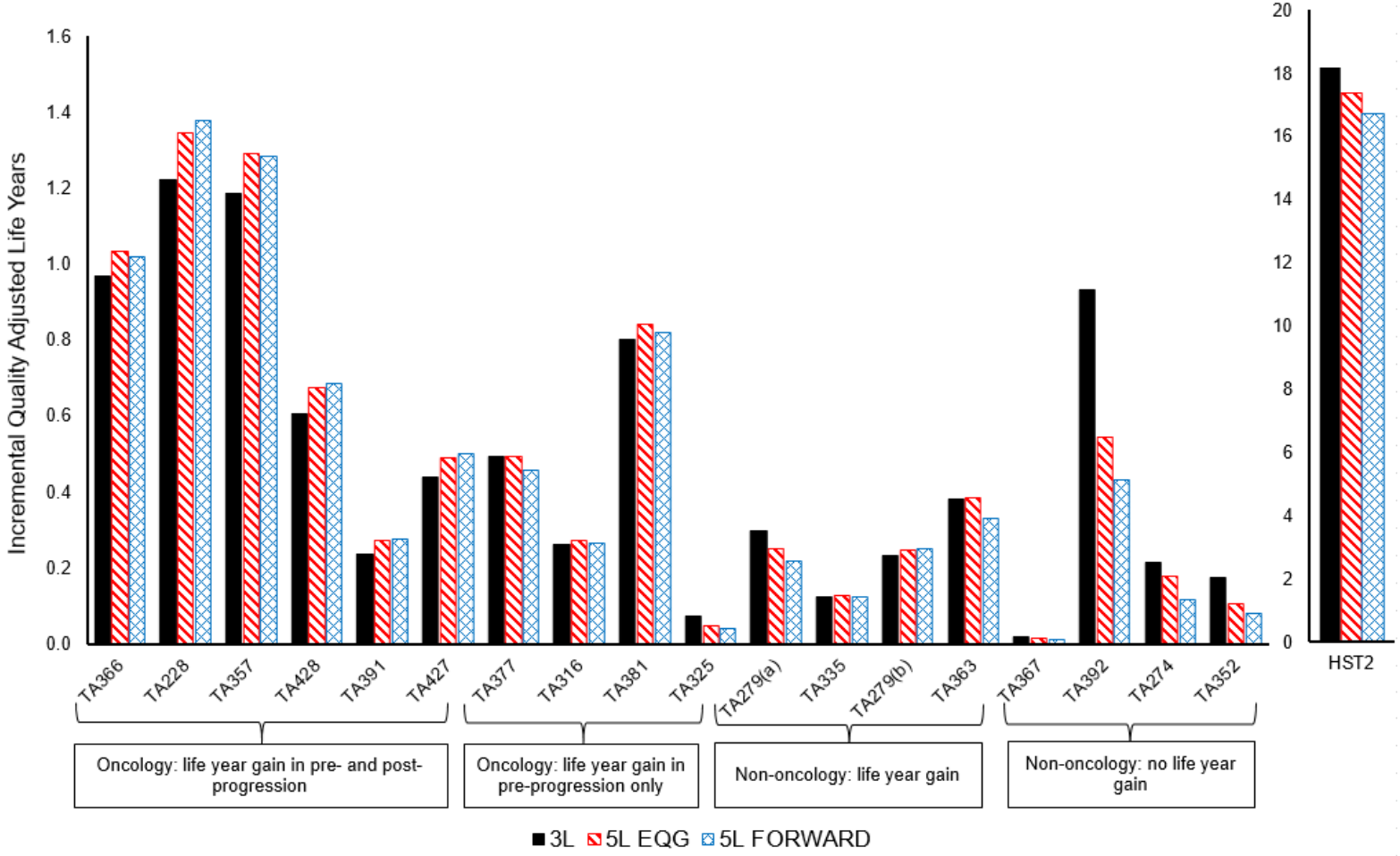
- When 5L EQG or 5L FORWARD are used, the incremental QALYs for the oncology case studies increase compared to 3L, with the exception of TA377
- When 5L EQG or 5L FORWARD are used, the incremental QALYs for the non-oncology case studies with no life year gain decrease compared to 3L
- When 5L EQG or 5L FORWARD are used, the incremental QALYs for the non-oncology case studies with life year gain increase for some case studies, and decrease for others

Figure 7 presents the ICERs and Figure 8 presents the percentage change in ICERs– these change as expected, with ICERs generally decreasing for oncology case studies, increasing for

non-oncology case studies without life year gain, and no clear pattern for non-oncology case studies with life year gain. ICERs are generally higher for FORWARD than EQG. The ICERs for TA325 and TA367 are not presented as the intervention is dominant (cost saving and more effective) for both of these case studies, using 3L and 5L.

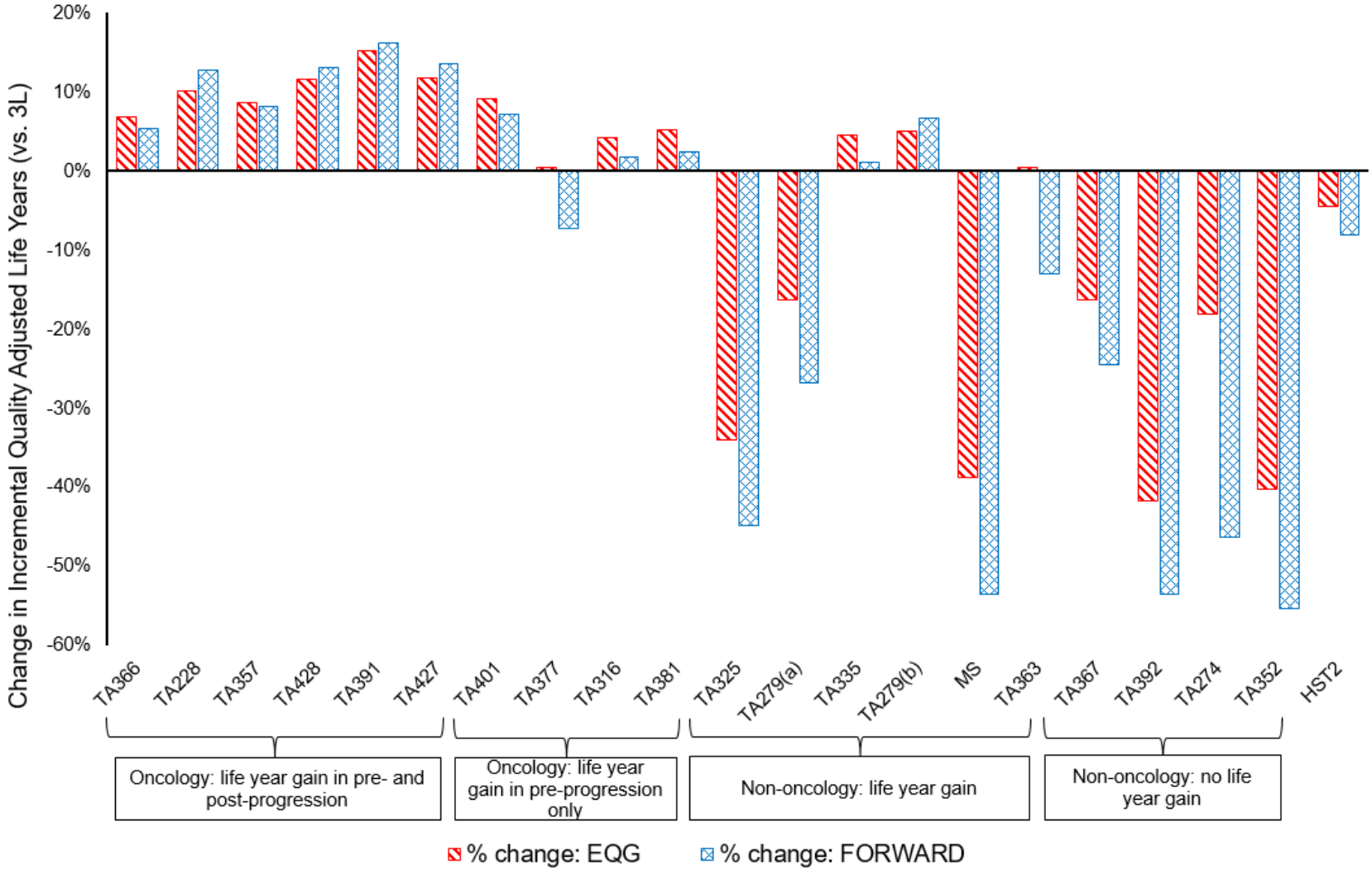
The three categories of case study are explored in further detail in subsequent sections

Figure 5: Summary results: incremental QALYs



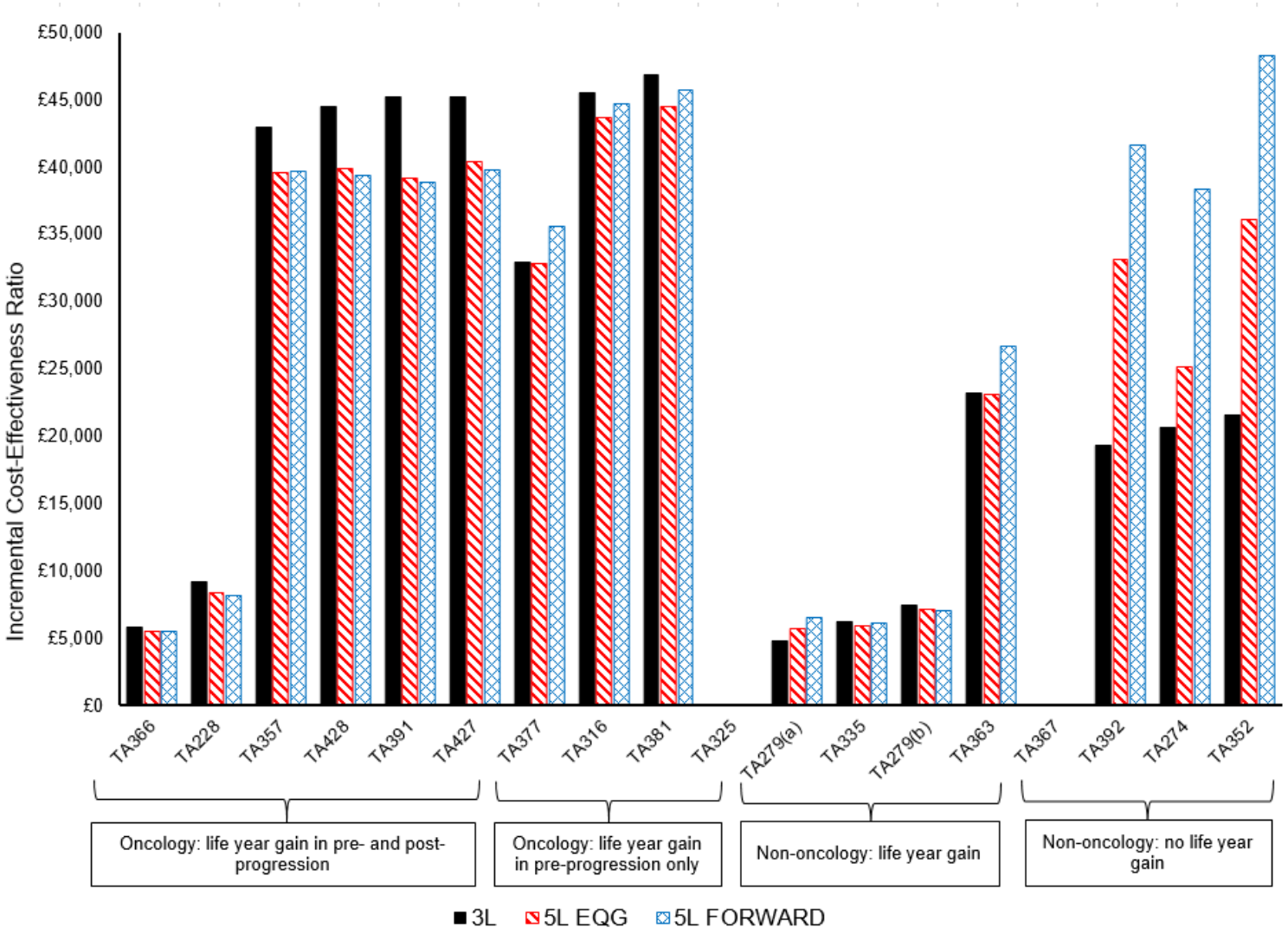
TA401 and MS are not presented in this figure as their incremental QALYs are confidential

Figure 6: Summary results: percentage change in incremental QALYs



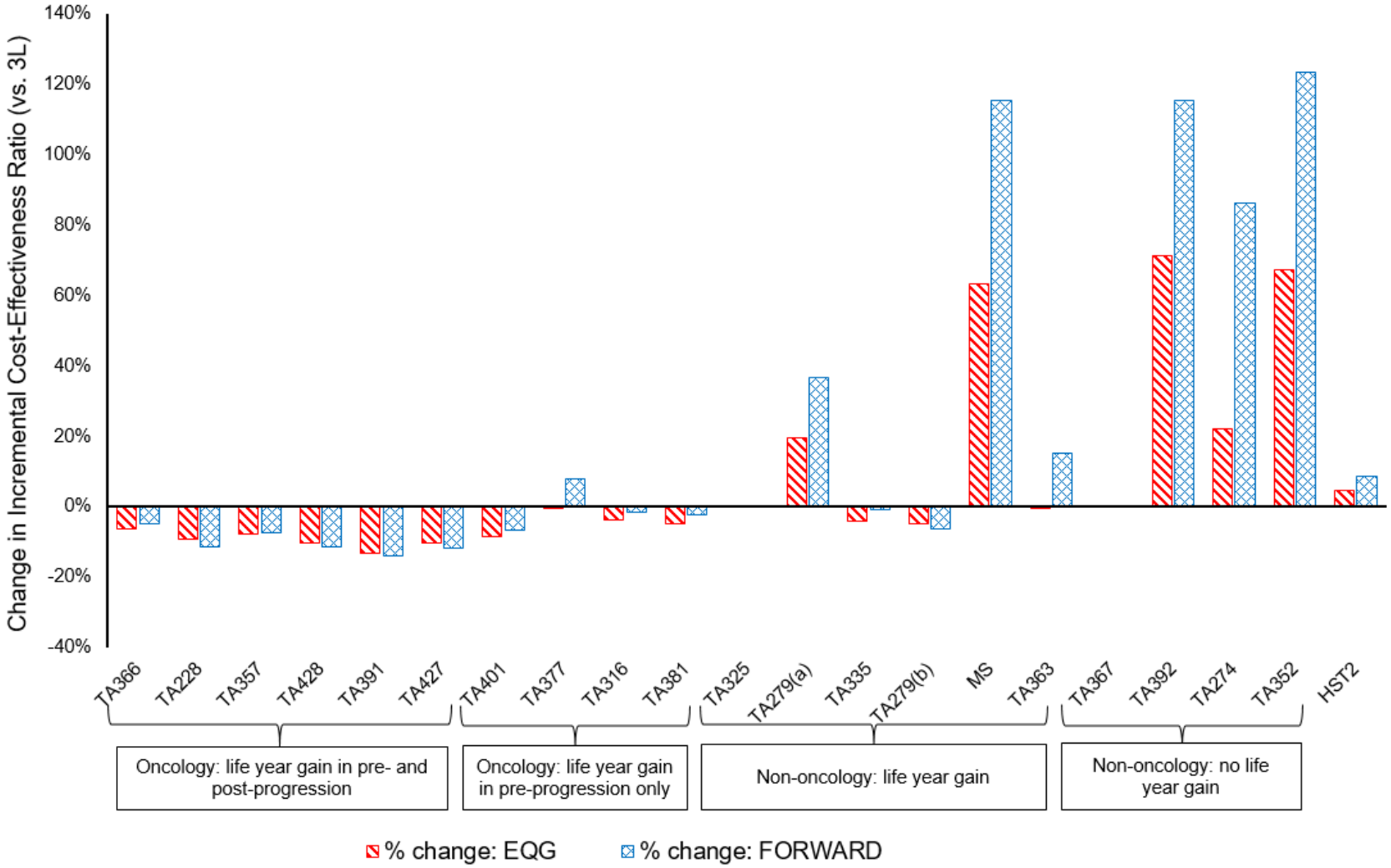
Scenario 4 is presented for the MS intervention.

Figure 7: Summary results: ICERs



TA401, MS and HST2 are not presented in this figure as their ICERs are confidential

Figure 8: Summary results: percentage change in ICERs



Scenario 4 is presented for the MS intervention.

To understand what may predict the direction and magnitude of change in the incremental QALYs and ICER after mapping from 3L to 5L, we plotted the change in incremental QALYs with the 5L compared to 3L against the 3L incremental QALYs and against the incremental life years (Figure 9**Error! Reference source not found.**), and the ICER change for 5L against the 3L incremental QALYs and against the incremental life years (Figure 10**Error! Reference source not found.**). The R^2 values denote that the 3L incremental QALYs explain a lower percentage of the QALY gain and ICER increase than the incremental life years do. This is unsurprising given that there is a bigger incremental QALY increase (and ICER decrease) for interventions that increase life years than for those that do not, whereas all interventions increase QALYs.

Figure 9: Change in incremental QALY correlation with incremental QALYs and incremental life years

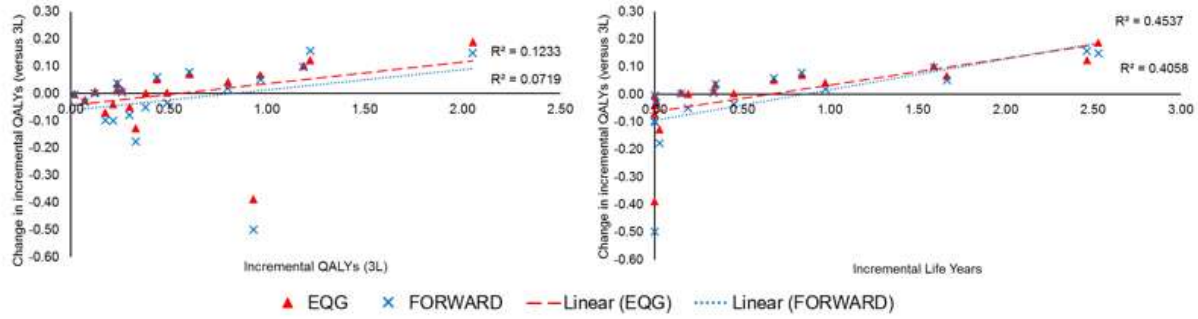
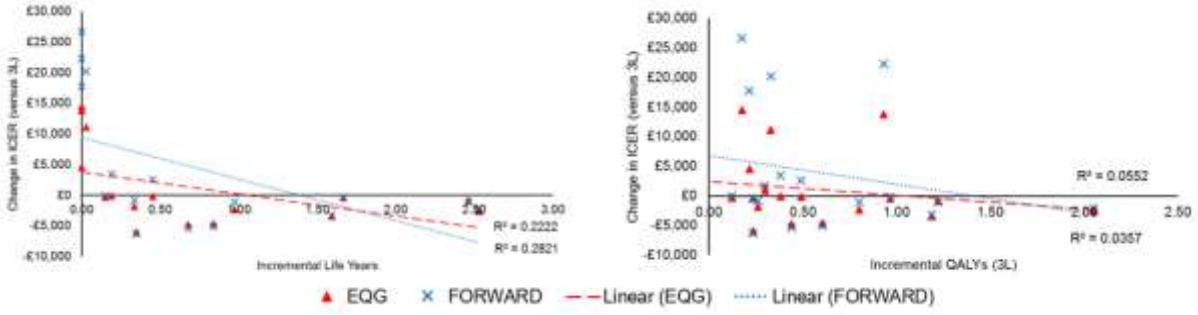


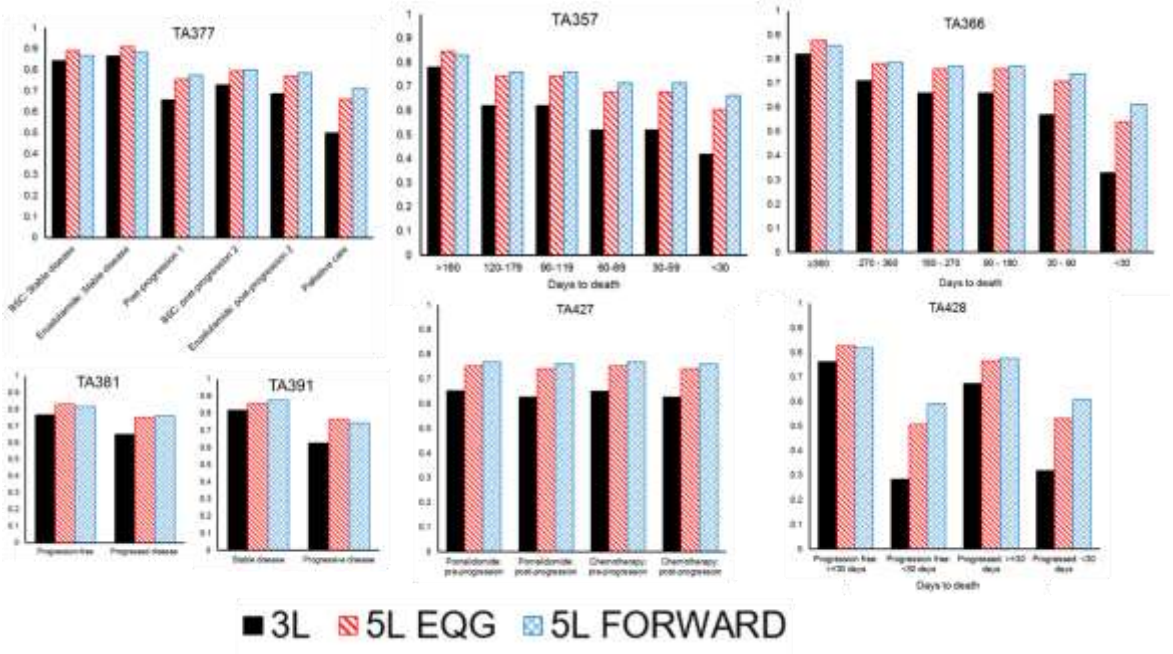
Figure 10: Change in ICER correlation with incremental QALYs and incremental life years



4.2.1. Oncology case study results

In each oncology case study, the utility value for all health states increased using 5L (EQG or FORWARD), shown in Figure 11. The total QALYs for intervention and comparator in all case studies therefore increased.

Figure 11: Utility values for oncology case studies



In order for the incremental QALYs to increase, the increase in QALYs for the intervention must be greater than the increase in QALYs for the comparator. The incremental QALYs increased for almost all oncology case studies, with the exception of TA377 using FORWARD. We explore why this is.

4.2.1.1. Utility as a function of time to death

Two case studies (TA357 and TA366) defined utility by time to death, with higher utility values for health states further from death. The utility values for all states increase using 5L compared to 3L. In both of these cases, the intervention increases the time spent in the health state furthest from death, and patients in the intervention and comparator arms spend approximately the same amount of time in the other health states. The increase in incremental QALYs using 5L therefore depends on the increase in utility of the health state furthest from death, and the difference in life years between intervention and comparator. The difference in life years between intervention and comparator is greater for TA357 than for TA366, and the increase in utility using 5L rather than 3L is greater for TA357 than TA366, so the increase in incremental QALYs is greater for TA357 than TA366.

4.2.1.2. Utility as a function of pre- and post-progression

Seven of the case studies (TA228, TA316, TA381, TA377, TA391, TA427, TA428) defined utility by health states for pre- and post-progression – with a higher utility for pre- than post-progression. Using 5L, the utility value for pre-progression and post-progression increases for all cases, but the difference between the pre-progression and post-progression utility value decreases using 5L. This means that a (hypothetical) intervention which increases progression-free survival but does not affect overall survival will result in a smaller incremental QALY using 5L than 3L, as the difference between the pre- and post-progression utility is less. A (hypothetical) intervention which does not affect progression-free survival and increases overall survival will result in a greater incremental QALY using 5L than 3L, as it is not affected by the difference between pre- and post-progression utilities, and increasing the post-progression utility increases the intervention QALYs.

In both TA381 and TA391, using 5L, the utility values for pre-progression and post-progression increase, and the difference between the utility values for the states decreases. In TA381, the QALY gain increase slightly using 5L, whereas the QALY gain in TA391 increases much more using 5L. This is partly because the intervention in TA381 increases the life years in pre-progression but decreases the life years in post-progression, whereas the intervention in TA391 increases the life years in pre-progression and post-progression. (The exact utility values also influence the size of the QALY gain increases).

4.2.1.1. Utility as a function of treatment

Three of the seven case studies (TA316, TA377, TA427) that defined utility by progression status also differentiated between utility in pre- and/or post-progression by intervention. In these cases, the incremental utility benefit for receiving intervention instead of comparator is reduced when using 5L compared to 3L. This decreases the incremental QALY somewhat.

In TA377, using 3L, the intervention resulted in more QALYs than the comparator because patients receiving the intervention spent longer in the pre-progression (stable) state, and the utility for pre-progression was higher for intervention than comparator. Patients receiving the intervention spent less time in each of the post-progression states than patients receiving the comparator. Using 5L, the utilities for the pre- and post-progression states all increases, but this increase is greater for the post-progression states than for the pre-progression state. Although the QALY gain from spending more time in pre-progression instead of being dead increases, the QALY gain from keeping people in pre-progression instead of post-progression

decreases. Additionally, the incremental utility for receiving the intervention instead of the comparator decreases, and so the QALY gain is decreased. When using FORWARD, the decreases in QALY gain outweigh the increases such that the overall QALY gain decreases.

TA401 compared bosutinib and hydroxycarbamide. There are health states for chronic phase on bosutinib, chronic phase on hydroxycarbamide, accelerated phase and blast phase. In the model, the undiscounted life years for chronic phase on hydroxycarbamide, accelerated phase and blast phase are identical for intervention and comparator. Patients in the intervention arm spend an additional period of time in the chronic phase on bosutinib state, and therefore the incremental QALYs for intervention are only influenced by the utility value of the chronic phase on bosutinib state.

4.2.1.2. Utility decrements

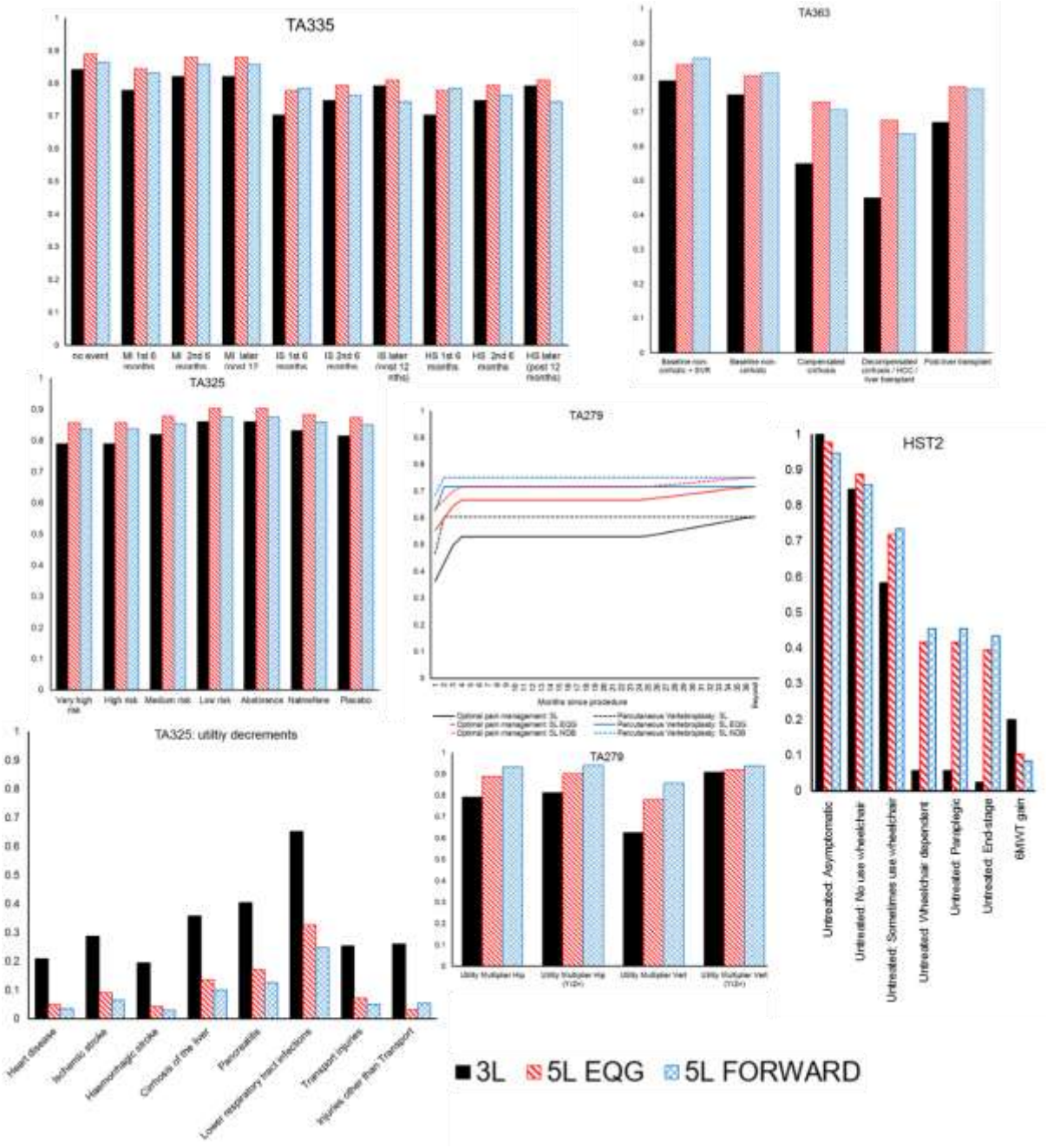
In most cases, we did not map the utility decrements for adverse events as the source of the decrement was not EQ-5D, or was not clear. The exceptions to this were TA427, where the regression included adverse events and we mapped the health state utilities (including decrements) rather than individual coefficients, and TA316 where there was a coefficient for skeletal related events. In TA316, the rate of skeletal related events was higher for comparator than intervention. The decrement for skeletal related events decreased when using 5L, so the QALYs for the comparator increased, which decreased the incremental QALY.

4.2.1. Non-oncology case studies with a life year gain

The 3L and 5L utility values for the health states in TA363, TA335, TA325, TA279 and HST2 are shown in Figure 12. The utility values increase using 5L for almost all health states – there are some exceptions in TA335 as the utility values for some health states are calculated using the relationships between other mapped values. Like the oncology case studies with health states for pre- and post-progression, there are some increases in the incremental QALYs from increasing the utility value in health states and some decreases in the incremental QALYs from reducing the difference in utility between health states. In TA325, there are utility decrements for serious events that are associated with higher risk levels. Since the intervention reduces risk levels, these serious events are less frequent for the intervention than comparator. These decrements are lower for 5L than 3L, which further reduces the incremental QALYs for intervention.

HST has the biggest gain in QALYs and life years of all the case studies. We may have expected therefore that the QALY gain would increase using 5L. The utility values for all health states except asymptomatic increase using 5L and so the QALYs for both elosulfase alfa and no treatment increase. The 6MWT gain increment decreases using 5L, so the benefit of being on elosulfase alfa rather than no treatment within the same health state decreases. Elosulfase alfa increases life years substantially compared to no treatment, but also delays progression and has a treatment-related utility increment. Using 5L, more value is given to the increase in life years through higher utility values, but less value is given to delaying progression as the difference between health state utilities decreases, and less value is given to the benefit of treatment within the same health state. These three factors combined mean that there is relatively little difference in incremental QALYs and ICERs using 5L.

Figure 12: Utility values for non-oncology case studies with a life year gain



4.2.1.1. Scenario analyses

The change in utility values between 3L and 5L is particularly noteworthy for the MS case study, where values that were negative or very low using 3L are now much higher. This means the incremental QALYs for 5L are much lower than for 3L, and the ICERs therefore increase substantially.

For the MS case study, there are also utility decrements for relapse, transition to SPMS, adverse events, and carers – these are explored further.

In this analysis, we explore the impact of mapping the disutilities for relapse, SPMS, adverse events and caregivers in the MS case study. We find that the QALY gain decreases and ICER increases when each additional element is mapped to 5L.

Scenario 1

When only the utility values for the RRMS states are mapped to 5L (and the original values are used for relapse, SPMS, adverse events and carer disutilities), the results are shown in Table 3, where the incremental QALY gain decreases and ICER increase. Since the decrease in QALYs is greater using FORWARD than EQG, the increase in ICER is greater using FORWARD.

Table 3: MS scenario 1

	Change in incremental QALYs (compared to 3L)	**Change in ICER (compared to 3L)**
EQ-5D-5L: EQG	-27.75%	38.41%
EQ-5D-5L: FORWARD	-36.05%	56.37%

RRMS state utility values mapped to 5L; original values used for relapse, SPMS, adverse events and carer disutilities

Scenario 2

The disutilities for SPMS and relapse decrease when 5L is used instead of 3L. The intervention reduces progression to SPMS and relapse. As expected, when the 5L disutilities for SPMS and relapse are used in addition to the 5L RRMS state utility values, and the original values are used for adverse events and carer disutilities, the incremental QALY gain decreases and ICER increases (Table 4).

Table 4: MS scenario 2

	Change in incremental QALYs (compared to 3L)	**Change in ICER (compared to 3L)**
EQ-5D-5L: EQG	-28.68%	40.20%
EQ-5D-5L: FORWARD	-43.09%	75.72%

RRMS state utility values, relapse, and SPMS mapped to 5L; original values used for adverse events and carer disutilities

Scenario 3

The disutilities for adverse events are smaller using 5L than 3L. The comparator has a higher incidence of adverse events than the intervention, so decreasing the utility decrements for the adverse events decreases the incremental QALYs and increases the ICER (Table 5).

Table 5: MS scenario 3

	Change in incremental QALYs (compared to 3L)	Change in ICER (compared to 3L)
EQ-5D-5L: EQG	-30.53%	43.94%
EQ-5D-5L: FORWARD	-44.86%	81.35%

RRMS state utility values, relapse, SPMS and adverse events mapped to 5L; original values used for carer disutilities

Scenario 4

Disutilities for caregivers are shown in Figure 13. The disutilities are smaller using 5L than 3L. Therefore the incremental QALYs decrease and ICER increases using 5L, as shown in Table 6.

Figure 13: MS disutilities for caregivers

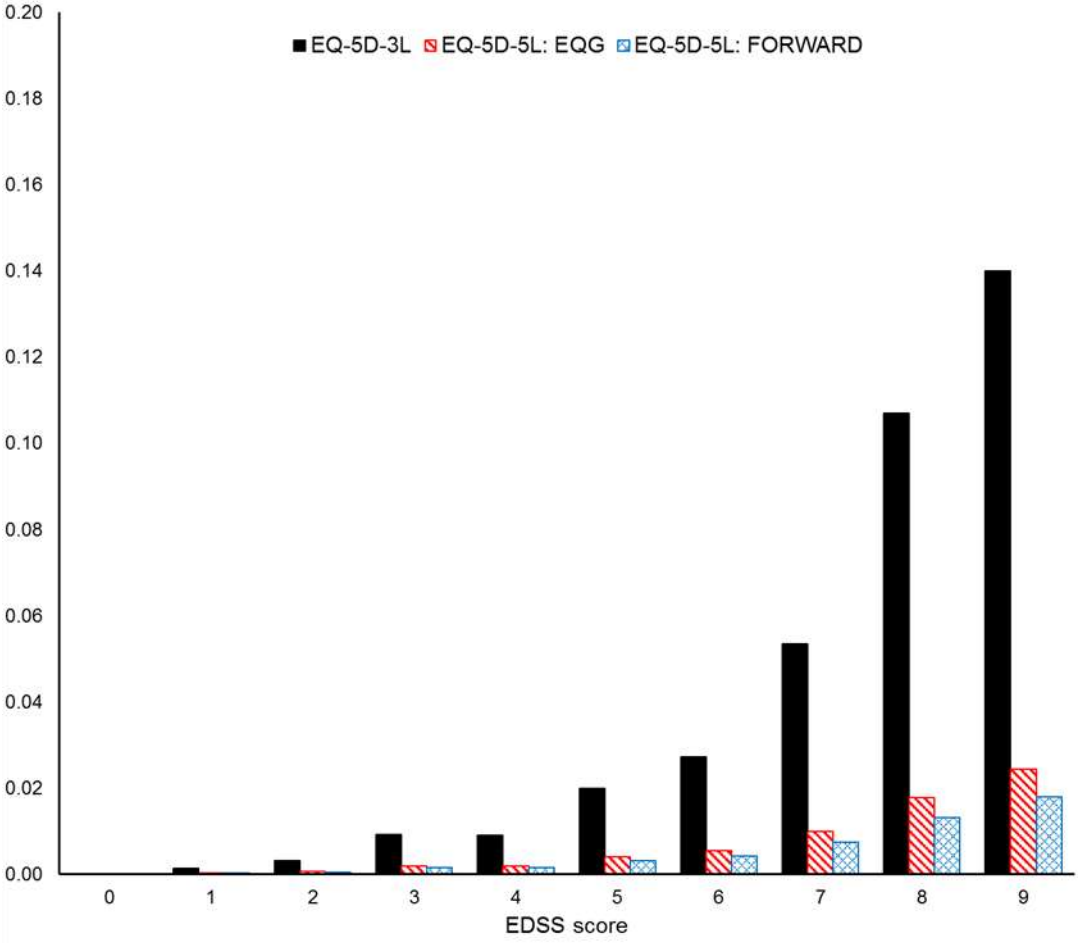


Table 6: MS scenario 4

	Change in incremental QALYs (compared to 3L)	**Change in ICER (compared to 3L)**
EQ-5D-5L: EQG	-38.79%	63.37%
EQ-5D-5L: FORWARD	-53.52%	115.16%

RRMS state utilities, relapse, SPMS, adverse events and carer disutilities mapped to 5L

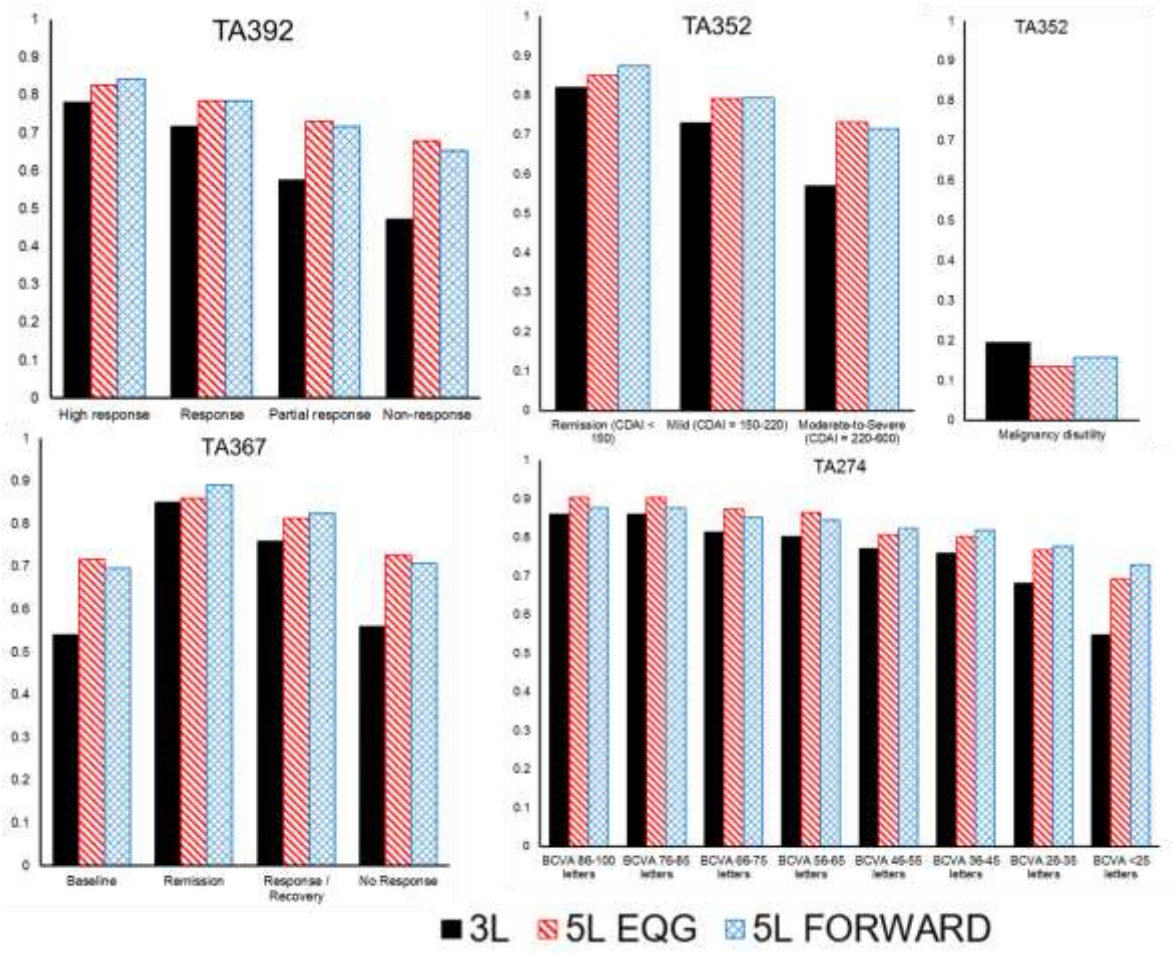
4.2.2. Non-oncology case studies with no life year gain

The 3L and 5L utility values for the health states in TA392, TA352, TA367 and TA274 are shown in Figure 14. The utility values for all health states increase when 5L (EQG or FORWARD) is used, so the total QALYs for intervention and comparator increase. The

difference in utility values between the health states decreases using 5L, so the benefit of avoiding disease progression or improving response is decreased. The incremental QALY for the intervention in all cases therefore decreases. This increases the ICERs, in some cases to above the £20,000-£30,000 range considered by NICE (TA367 remains dominant).

Additionally, the disutility for malignancy in TA352 is also shown – this decreases using 5L. There are more malignancies for comparator than intervention, so decreasing the disutility increases the QALYs for comparator and therefore decreases the incremental QALYs.

Figure 14: Utility values for non-oncology case studies with no life year gain



5. DISCUSSION

Mapping from 3L to 5L increased almost all of the utility values, (with the exception of the asymptomatic in HST2 and some calculated utilities in TA335). Generally, the utility value for any given health state increased, and the difference between best and worst health states decreased. This meant that the QALY gain for interventions that provided a substantial extension to life within one health state increased, whereas the QALY gain for interventions that derived most of their benefit from the difference in utility values between health states

decreased. In the case studies considered here, the oncology interventions saw the biggest increases in QALY gains, especially those interventions which improved survival in both the pre- and post-progression states. Conversely, the non-oncology interventions with little or no improvement in survival saw large decreases in the QALY gain.

The magnitude of change in the QALY gain from mapping from 3L to 5L was generally greater when the FORWARD databank was used than when the EQG dataset was used. Previous research has found that using EQG or FORWARD in the mapping generates different 5L scores[7] and discussed that this is because of differences in the characteristics of the datasets. The FORWARD databank includes only patients with rheumatic and musculoskeletal diseases whereas the EQG dataset includes patients from a range of different diseases and a healthy population. The FORWARD databank is further restricted to a US/Canadian population and respondents were primarily female and in better health than the respondents in the EQG dataset[8]. It has been suggested therefore that the EQG data may be preferable for use in NICE appraisals[8], although the more substantial separation between the 3L and 5L questions in the FORWARD databank may make it more likely that the responses were truly independent.

Our findings are consistent with previous research exploring the impact of mapping from 3L to 5L in case studies in economic evaluations alongside clinical trials[7] and with predictions from studies comparing the 3L and 5L value sets[56]. Our research adds to the existing literature by considering economic modelling case studies and including studies which improved survival in addition to quality of life. Our analysis used 20 different economic models across a broad range of disease areas, and which used several different approaches to model utility. We therefore believe that our results will be generalizable to other economic evaluations, including those that use other model structures. Although all of our models were cohort-level state transition or partitioned survival analysis models, some models used health state utility values, some used disutilities, and some used utility multipliers. The direction and magnitude of the results were not influenced by the model structure or approach to modelling utility, indicating that our findings would apply across other economic evaluations.

Since the magnitude and direction of change in QALY gain was not consistent across case studies, the magnitude and direction of change in ICER was also not consistent. ICERs generally decreased for oncology interventions, and increased for non-oncology interventions. While the change in ICERs for oncology interventions were relatively small (mean -7.54% for

EQG and -6.43% for FORWARD) and consistent (range: -13.23% to 0.49% for EQG and -13.94% to 7.89%) across case studies, there was much more variation in the change in ICERs for non-oncology interventions. In two non-oncology case studies the ICERs decreased by less than 10%, whereas the ICERs in others increased by over 100%. Clearly, therefore, there is no simple way to re-estimate how the ICER for a non-specified intervention would change using 5L. Our results suggest that a switch to 5L would result in oncology and other substantially life-extending interventions becoming more cost-effective and interventions which primarily improve quality of life becoming less cost-effective. This could potentially change the range of interventions reimbursed in the UK, although in practice the pricing of technologies may adapt such that NICE's decisions would not change.

The impact of 5L on NICE's decision making is further complicated by the additional considerations given to a 'life-extending treatment at the end of life'. Where a treatment meets specific criteria, the NICE Methods Guide states that the appraisal committee will consider the impact of assuming that the extended survival period is experienced at the full quality of life expected for a healthy individual of the same age[1]. If all utility values increase for 5L and the range of utility values decreases for 5L, as they do in the case studies, then the difference between disease-specific and healthy population utility values will decrease. The impact of using utility values for a healthy population will then be less than it would be using 3L and the reduction in the ICER under this scenario would be smaller.

Our case studies were limited to those which considered primarily 3L. Although EQ-5D is NICE's preferred measure of quality of life, it is not used exclusively in economic models considered by NICE. In our analyses, we did not map utilities which were not clearly EQ-5D scores, and so in many cases utility decrements for adverse events were left unchanged. In scenario analysis, we explored the impact of mapping utility increments and decrements that were not reported as 3L scores, and found that the QALY gain decreased as each additional input was mapped to 5L, because modelled patients in the intervention arm spent less time in the health states associated with disutilities. (If an intervention is associated with more adverse events than the comparator, and the utility decrements decrease using 5L, then the QALY gain would increase). Mapping only the utility inputs which were clearly 3L increased the ICER by 38-56%, whereas mapping all utility inputs increased the ICER by 63-115%. Hypothetically, it is possible that an ICER increase of 38-56% may not change NICE's recommendation, whereas an ICER increase of 63-115% may do, demonstrating the potential importance of

mapping all utility inputs to 5L regardless of their origin. It could be argued that utility values from other sources should not be mapped to 5L as they were not 3L and so their valuation system has not changed, but it could also be argued that they should be mapped to 5L as they are being used in place of 3L. The decision as to whether non-EQ-5D values should be mapped has important consequences. For example, in metastatic breast cancer, utility values for pre-progression may be from 3L included in trials, whereas utility values for post-progression are often taken from Lloyd et al. 2006[57-63], which derived utility values from a vignette study using standard gamble[64]. If a post-progression value of 0.443, were assumed to be 3L and mapped to 5L, it would increase to approximately 0.627-0.659. This is important when considering the difference in pre- and post-progression utility values, as discussed earlier.

Our analysis is further limited by the approach to selecting case studies, which was pragmatic and drew on the experience of the project team to ensure a broad range of disease areas was covered. However, a more systematic approach to selecting case studies may have been more thorough and less prone to potential bias.

6. CONCLUSION

Mapping from 3L to 5L increases utility scores, and reduces the difference in utility between best and worst states. This means that for an intervention that only increases survival and does not increase quality of life, the incremental QALYs will increase and the ICER will decrease. The increase in QALYs and decrease in ICER will generally be larger using FORWARD than using EQG. For an intervention that only increases quality of life by delaying or avoiding disease progression and does not affect survival, the incremental QALYs will decrease and the ICER will increase. The decrease in QALYS and increase in ICER will generally be larger using FORWARD than using EQG. Many interventions both avoid or delay progression and increase survival, so there is a trade-off between the gain in incremental QALYs from increasing survival and the decrease in incremental QALYs from reducing the benefit of delayed progression.

The use of 3L and 5L within economic models leads to different estimates of cost-effectiveness. The choice of 3L or 5L within one appraisal could lead to different reimbursement decisions, and the use of 3L in some appraisals and 5L in others could lead to inconsistencies in decision

making. Future changes to NICE policy need to be aware of this information in order to ensure decision making is consistent, fair and reflects scientific state of the art.

7. BIBLIOGRAPHY

1. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. 2013. https://www.nice.org.uk/process/pmg9/chapter/the-reference-case (accessed 13 December 2017).
2. Foundation ER. About EQ-5D. 28 April 2017 2017. https://euroqol.org/eq-5d-instruments/ (accessed 26 February 2018).
3. National Institute for Health and Care Excellence. Position statement on use of the EQ-5D-5L valuation set. 2017. https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf.
4. Brazier J, Briggs A, Bryan S. EQ-5D-5L: Smaller steps but a major step change? *Health Economics* 2018; **27**(1): 4-6.
5. Devlin N, Brazier J, Pickard AS, Stolk E. 3L, 5L, What the L? A NICE Conundrum. *PharmacoEconomics* 2018.
6. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L Better Than EQ-5D-3L? A Head-to-Head Comparison of Descriptive Systems and Value Sets from Seven Countries. *Pharmacoeconomics* 2018.
7. Wailoo A, Hernandez Alava M, Grimm S, et al. Comparing the EQ-5D-3L and 5L versions. What are the implications for cost effectiveness estimates? 2017. http://scharr.dept.shef.ac.uk/nicedsu/methods-development/eq-5d-5l/.
8. Hernandez Alava M, Wailoo A, Pudney S. Methods for mapping between the EQ-5D-5L and the 3L for technology appraisal. . 2017. https://scharr.dept.shef.ac.uk/nicedsu/methods-development/eq-5d-5l/.
9. Hernandez-Alava M, Pudney S. Econometric modelling of multiple self-reports of health states: The switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis. *J Health Econ* 2017; **55**: 139-52.
10. Hernandez-Alava M PS. eq5dmap: a command for mapping from 3-level to 5-level EQ-5D. *The STATA Journal*.
11. National Institute for Health and Care Excellence. Cabazitaxel for hormone-relapsed metastatic prostate cancer treated with docetaxel. 2016. https://www.nice.org.uk/guidance/ta391
12. National Institute for Health and Care Excellence. Olaparib for maintenance treatment of relapsed, platinum-sensitive, BRCA mutation-positive ovarian, fallopian tube and peritoneal cancer after response to second-line or subsequent platinum-based chemotherapy. 2016. https://www.nice.org.uk/guidance/ta381
13. National Institute for Health and Care Excellence. Topotecan, pegylated liposomal doxorubicin hydrochloride, paclitaxel, trabectedin and gemcitabine for treating recurrent ovarian cancer. 2016. https://www.nice.org.uk/guidance/ta389
14. National Institute for Health and Care Excellence. Enzalutamide for metastatic hormone-relapsed prostate cancer previously treated with a docetaxel-containing regimen. 2014. https://www.nice.org.uk/guidance/ta316.

15. Sandblom G, Carlsson P, Sennfalt K, Varenhorst E. A population-based study of pain and quality of life during the year before death in men with prostate cancer. *Br J Cancer* 2004; **90**(6): 1163-8.
16. National Institute for Health and Care Excellence. Enzalutamide for treating metastatic hormone-relapsed prostate cancer before chemotherapy is indicated. 2016. https://www.nice.org.uk/guidance/ta377.
17. Wolff JM DV, Klier J, Erhardt W, Dr. Dass RN, Geiges G. Quality of life among German patients with metastatic castration-resistant prostate cancer *Value Health* 2012; **15**: A431.
18. Diels J, Hamberg P, Ford D, Price PW, Spencer M, Dass RN. Mapping FACT-P to EQ-5D in a large cross-sectional study of metastatic castration-resistant prostate cancer patients. *Qual Life Res* 2015; **24**(3): 591-8.
19. National Institute for Health and Care Exclellence. Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy. 2017. https://www.nice.org.uk/guidance/ta428.
20. Kind P HG, Macran S. UK Population Norms for EQ-5D, 1999.
21. National Institute for Health and Care Excellence. Pomalidomide for multiple myeloma previously treated with lenalidomide and bortezomib. 2017. https://www.nice.org.uk/guidance/ta427.
22. National Institute for Health and Care Excellence. Bortezomib and thalidomide for the first-line treatment of multiple myeloma. 2011. https://www.nice.org.uk/guidance/ta228.
23. McKenzie L, van der Pol M. Mapping the EORTC QLQ C-30 onto the EQ-5D instrument: the potential to estimate QALYs without generic preference data. *Value Health* 2009; **12**(1): 167-71.
24. National Institute for Health and Care Excellence. Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab. 2015. https://www.nice.org.uk/guidance/ta357.
25. Beusterien KM, Szabo SM, Kotapati S, et al. Societal preference values for advanced melanoma health states in the United Kingdom and Australia. *Br J Cancer* 2009; **101**(3): 387-9.
26. National Institute for Health and Care Excellence. Pembrolizumab for advanced melanoma not previously treated with ipilimumab. 2015.
27. National Institute for Health and Care Excellence. Bosutinib for previously treated chronic myeloid leukaemia. 24 August 2016 2016. https://www.nice.org.uk/guidance/ta401/chapter/4-Committee-discussion (accessed 07 March 2018).
28. National Institute for Health and Care Excellence. Rivaroxaban for preventing adverse outcomes after acute management of acute coronary syndrome. 2015. https://www.nice.org.uk/guidance/ta335.
29. National Institute for Health and Care Excellence. Ticagrelor for the treatment of acute coronary syndromes. 2011. https://www.nice.org.uk/guidance/ta236
30. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010; **13**(5): 509-18.
31. Latour-Perez J, Balsa, E.D., Betegon, L., Badia, X. Using triple antiplatelet therapy in patients with non-ST elevation acute coronary syndrome managed invasively: A cost-effectiveness analysis. *Value Health* 2008; **11**(5): 853-61.
32. Crespin DJ, Federspiel JJ, Biddle AK, Jonas DE, Rossi JS. Ticagrelor versus genotype-driven antiplatelet therapy for secondary prevention after acute coronary syndrome: a cost-effectiveness analysis. *Value Health* 2011; **14**(4): 483-91.

33. Kazi DS, Garber, A.M., Shah, R.U., Dudley, R.A., Mell, M.W., Rhee, C. et al. Cost-Effectiveness of Genotype-Guided and Dual Antiplatelet Therapies in Acute Coronary Syndrome. *Ann Intern Med* 2014; **160**(4): 221-32.
34. Sullivan PW, Arant, T.W., Ellis, S.L., Ulrich, H. The cost-effectiveness of anticoagulation management services for patients with atrial fibrillation and at high risk of stroke in the US. . *Pharmacoeconomics* 2006; **24**(10): 1021-33.
35. National Institute for Health and Care Excellence. Ledipasvir–sofosbuvir for treating chronic hepatitis C. 2015. https://www.nice.org.uk/guidance/ta363.
36. Wright M GR, Roberts R, Main J, Thomas HC. Health benefits of antiviral therapy for mild chronic hepatitis C: randomised controlled trial and economic evaluation. *Health Technology Assessment* 2006; **10**(21).
37. Vera-Llonch M, Martin M, Aggarwal J, et al. Health-related quality of life in genotype 1 treatment-naive chronic hepatitis C patients receiving telaprevir combination treatment in the ADVANCE study. *Aliment Pharmacol Ther* 2013; **38**(2): 124-33.
38. Ratcliffe J, Longworth L, Young T, et al. Assessing health-related quality of life pre- and post-liver transplantation: a prospective multicenter study. *Liver Transpl* 2002; **8**(3): 263-70.
39. National Institute for Health and Care Excellence. Nalmefene for reducing alcohol consumption in people with alcohol dependence. 2014. https://www.nice.org.uk/guidance/ta325.
40. Lundbeck Limited. Nalmefene for reducing alcohol consumption in people with alcohol dependence, 2014.
41. Purshouse R BA, Latimer N, Meng Y, Rafia R, Jackson R, Meier P. Modelling to assess the effectiveness and cost-effectiveness of public health related strategies and interventions to reduce alcohol attributable harm in England using the Sheffield Alcohol Policy Model version 2.0.2009. https://www.sheffield.ac.uk/polopoly_fs/1.107991!/file/Appendices.pdf (accessed.
42. National Institute for Health and Care Excellence. Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral compression fractures. 2013. https://www.nice.org.uk/guidance/ta279.
43. Kanis JA, Johnell O, Oden A, et al. The risk and burden of vertebral fractures in Sweden. *Osteoporos Int* 2004; **15**(1): 20-6.
44. Christian J Hendriksz CL, Mahmut Coker, Sema Kalkan Ucar, Mohit Jain, Lisa Bell, Christina Lampe. Burden of disease in patients with Morquio A syndrome: results from an international patient-reported outcomes survey. *Orphanet Journal of Rare Diseases* 2014; **9**(32): 1-8.
45. 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* 2017.
46. Mandy van Reened BJ. EQ-5D-5L User Guide, 2015.
47. Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. *Med Decis Making* 2011; **31**(6): 800-4.
48. BioMarin Europe Limited. Elosulfase alfa for the treatment of mucopolysaccharidosis type IVA, 2015.
49. Lampe C, Jain M, Olaye A, Meesen B, Decker C, Mengel E. Relationship Between Patient-Reported Outcomes and Clinical Outcomes in Patients With Morquio A Syndrome. *Journal of Inborn Errors of Metabolism and Screening* 2015; **3**.
50. National Institute for Health and Care Excellence. Adalimumab for treating moderate to severe hidradenitis suppurativa. 2016.

51. National Institute for Health and Care Excellence. Vedolizumab for treating moderately to severely active Crohn's disease after prior therapy. 2015. https://www.nice.org.uk/guidance/ta352.
52. Wild D WM, Pettengell R, Lewis G. Utility elicitation in patients with follicular lymphoma. *Value Health* 2006; **9**(6): A294.
53. National Institute for Health and Care Excellence. Vortioxetine for treating major depressive episodes. 2015. https://www.nice.org.uk/guidance/ta367
54. National Institute for Health and Care Excellence. Ranibizumab for treating diabetic macular oedema. 2013. https://www.nice.org.uk/guidance/ta274/chapter/4-Consideration-of-the-evidence.
55. Mitchell P, Bandello F, Schmidt-Erfurth U, et al. The RESTORE study: ranibizumab monotherapy or combined with laser versus laser monotherapy for diabetic macular edema. *Ophthalmology* 2011; **118**(4): 615-25.
56. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. *Appl Health Econ Health Policy* 2017; **15**(2): 127-37.
57. National Institute for Health and Care Excellence. Bevacizumab in combination with capecitabine for the first-line treatment of metastatic breast cancer. 2012. https://www.nice.org.uk/guidance/ta263 (accessed 13 December 2017).
58. National Institute for Health and Care Excellence. Eribulin for treating locally advanced or metastatic breast cancer after 2 or more chemotherapy regimens. 2016. https://www.nice.org.uk/guidance/ta423 (accessed 13 December 2017).
59. National Institute for Health and Care Excellence. Everolimus with exemestane for treating advanced breast cancer after endocrine therapy. 2016. https://www.nice.org.uk/guidance/ta421 (accessed 13 December 2017).
60. National Institute for Health and Care Excellence. Fulvestrant for the treatment of locally advanced or metastatic breast cancer. 2011. https://www.nice.org.uk/guidance/ta239 (accessed 13 December 2017).
61. National Institute for Health and Care Excellence. Lapatinib or trastuzumab in combination with an aromatase inhibitor for the first-line treatment of metastatic hormone-receptor-positive breast cancer that overexpresses HER2. 2012. https://www.nice.org.uk/guidance/ta257 (accessed 13 December 2017).
62. National Institute for Health and Care Excellence. Trastuzumab emtansine for treating HER2-positive advanced breast cancer after trastuzumab and a taxane. 2017. https://www.nice.org.uk/guidance/ta458 (accessed 13 December 2017).
63. National Institute for Health and Care Excellence. Ribociclib for breast cancer. 2017. https://www.nice.org.uk/guidance/ta496 (accessed 13 December 2017).
64. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. *Br J Cancer* 2006; **95**(6): 683-90.