FRAMEWORK FOR ANALYSING RISK IN HEALTH TECHNOLOGY ASSESSMENTS AND ITS APPLICATION TO MANAGED ENTRY AGREEMENTS

REPORT BY THE DECISION SUPPORT UNIT

January 2016

Sabine Grimm, Mark Strong, Alan Brennan, Allan Wailoo
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.

Mark Strong is also funded by a National Institute for Health Research Post-Doctoral Fellowship (PDF-2012-05-258). This report is independent research supported by the National Institute for Health Research. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

Acknowledgements

The authors would like to thank the following for their contribution to the production of this report; Meindert Boysen, Simon Dixon, Jenny Dunn, Greg Fell, Elisabeth George, Melinda Goodall, Susan Griffin, Pall Jonsson, Bijal Joshi, Nina Lathia, Leeza Osipenko, Nina Pinwill, Claire Rothery, Eldon Spackman, Andrew Stevens, Matt Stevenson and Paul Tappenden.

EXECUTIVE SUMMARY

Background: Recent changes to the regulatory landscape of pharmaceuticals mean that decision-making bodies such as NICE will often be required to issue guidance on technologies with an evidence base that is smaller or earlier in its evolution than has previously been the case. For example, the European Medical Agency's new licensing scheme called "Adaptive Pathways" is designed to bring new technologies to patients sooner. In such situations, there will often be greater uncertainty regarding the clinical and cost-effectiveness of new technologies at the point of decision making. Over a series of decisions, this uncertainty leads to increased costs to the health system and foregone health in the population served.

Managed Entry Agreements (MEA), often termed "risk-sharing agreements", are schemes designed to reduce these costs. MEAs have been used by decision makers to recommend technologies under two broad conditions: (1) that the price of the technology be reduced and/or (2) that further research be conducted. Both conditions have the aim of reducing risk and decision uncertainty. MEA schemes have been used in the UK and other countries. They are not currently considered systematically within NICE technology appraisals. Taxonomies that categorise MEAs do exist, but at present there is no analysis framework to assess the need for an MEA scheme or to facilitate the choice between different possible MEA schemes.

Objectives: The main aim of this work was to develop a framework for analyzing risk in health technology assessments and applying this framework to proposed MEA schemes, to assess the value of alternative MEA schemes.

Methods: A review of existing taxonomies of MEA schemes was synthesized and updated. For each of the identified schemes we set out the rationale for its use and circumstances when it might be appropriate. We designed a questionnaire for use in technology appraisals to clarify issues surrounding the uncertainties in the appraisal. Finally, we developed the MEA Risk Analysis Framework to support a quantified analysis of the value of alternative MEAs in terms of expected health and cost outcomes and the reduction in decision risk.

Case Studies: To illustrate the practical use of the MEA Risk Analysis Framework, we applied it in-depth to a case study on the appraisal of pazopanib for treatment of advanced

renal cell carcinoma in which a real MEA has been agreed. This analysis entailed quantifying uncertainty associated with different decision options, designing the most suitable MEA schemes and assessing their value in terms of the reduction in the Payer Strategy and Uncertainty Burden (see below) likely to result from their implementation. We also undertook seven other case studies on previous NICE appraisals. In four of these we quantified the uncertainty associated with different decision options (without analysing specific MEA proposals), using an extended version of the Sheffield Accelerated Value of Information (SAVI) tool. In the other three of these seven case studies we examined the decision context in detail, described the rationale for, and the potential design of, possible MEA schemes, and set out in broad terms how the MEA Risk Analysis Framework would be used to assess the value of these different schemes.

Results – An Adapted Taxonomy of Possible Managed Entry Agreements

The new taxonomy of possible MEAs adapts recent work in other jurisdictions and sets out options on two dimensions. The first is price adjustments of various kinds, from straight discounts through to prices which are conditional on health outcomes for patients. The second is further evidence collection, through research vehicles ranging from further randomised trials through to observational 'real world evidence' studies and further analysis of existing data. Both of these dimensions can be used in a proposed MEA, creating a matrix of possible options.

Results – The MEA design guidance questionnaire

The MEA design guidance questionnaire consists of a set of questions that are useful to answer to identify the potentially appropriate schemes from the taxonomy of possible managed entry agreements. Because the pivotal factors influencing the design of MEAs are quantitative in nature, the MEA design guidance questionnaire can only fully be answered once the MEA risk analysis framework has been applied to an appraisal.

Results - Summary of the MEA Risk Analysis Framework

The MEA Risk Analysis Framework extends the content of the current NICE methods guide. It considers (i) costs to the NHS and social care system (more internationally described here as 'payer costs'), (ii) 'payer health benefits' expressed in quality adjusted life years (QALYs) though any outcome can be used, and (iii) decision uncertainty quantified using probabilistic sensitivity analysis, cost-effectiveness planes and acceptability curves. The key extension to

this standard process is to use Value of Information calculations to quantify the risk associated with current decision uncertainty and how that risk would change under any MEA. This enables an assessment of the expected value of the risk reduction that can occur when a specific proposed managed entry agreement is implemented. The key elements of the framework are as follows.

(1) Expected Costs, QALYs and Net Benefit of Each Strategy – This is the standard health economic analysis undertaken in NICE technology appraisals. The ICER being below £20,000 to £30,000 per QALY would suggest a recommendation by NICE in standard appraisals, and if there are several options, then the strategy with the highest expected net benefit (e.g. £20,000*QALYs – Cost) would be recommended.

(2) Uncertainty Analysis - Probabilistic sensitivity analysis (PSA) assesses the uncertainty in the costs and effects of each possible intervention in the decision problem based on the existing evidence. The cost-effectiveness plane and the CEAC already used in NICE appraisals quantify the probability that the intervention is cost-effective i.e. that the proposed decision to adopt the new technology would be correct. For example, imagine a new technology with an ICER of £19,000 per QALY, which would be recommended by NICE because it is below the maximum acceptable ICER, but also with substantial uncertainty - the PSA showing 55% probability of being cost-effective at £20,000 per QALY and the ICER reaching as high as £80,000 per QALY.

(3) Payer Uncertainty Burden (PUB) – is the monetary value of the risk of making a particular decision due to unresolved decision uncertainty, given the proposed price and evidence (expressed in either monetary or health units). It combines two key concepts: first, the probability that the chosen strategy is not the true optimal strategy, and second, the consequences of a 'wrong' decision in terms of QALYs and NHS costs that could have been saved if the truly optimal strategy had been selected instead.

(3a) Payer Uncertainty Burden (PUB) relates to the overall decision. The PUB is the expected value of the risk that the expected optimal strategy option (i.e. the one with the highest expected net benefit given current evidence and price) is not in truth the most cost-effective. For our imagined new technology, if the true ICER is £80,000 per QALY then NICE would be making the 'wrong' decision if it recommended approval, and the NHS could

have invested that money on other interventions to provide more QALYs for the same money - the 'opportunity cost' or 'opportunity loss'. An interpretation of the maximum acceptable ICER is that it is the rate at which the NHS uses resources to achieve additional health gain (i.e. it costs around £20,000 to purchase an additional QALY in the NHS) and thus the NHS could have purchased 4 QALYs with the £80,000 rather than the one QALY it would achieve with the new technology. Of course we do not know in advance that the true ICER is £80,000 per QALY but the PSA quantifies the chances that it is £80k, or £60k, or £40k etc. For each possibility we can also quantify the opportunity cost, and hence we can calculate the overall average expected opportunity cost. This averaged opportunity cost we call the Payer Uncertainty Burden (PUB).

The PUB per patient can be estimated using financial value (QALYs*λ – Costs) or on a QALY scale (Financial Value / λ). This can then be scaled up to an annual Payer Uncertainty Burden for the jurisdiction (e.g. England) by multiplying by the number of patients per year for whom the decision is made. It can also be quantified for a longer decision relevant time horizon e.g. for the next five or ten years.

The Payer Uncertainty Burden for the optimal strategy is mathematically the same as both the overall Expected Value of Perfect Information and the overall Expected Opportunity Loss in the health economics literature.

(3b) The Payer Strategy and Uncertainty Burden (P-SUB)

In addition to the Payer Uncertainty Burden, which relates to the risk associated with the decision, there also is risk associated with each strategy (interchangeably used with technology) in the decision problem. If the decision maker were to recommend a technology with an ICER above the threshold then there are two important consequences. First, the expected net benefit accrued would be lower than that for the optimal strategy. We define the difference between the expected net benefit of a technology and the expected net benefit of the optimal strategy as the Payer Strategy Burden (PSB) of that technology, which is the strategy-specific burden given the expected price and the research evidence. Second, there remains decision uncertainty, which applies to each technology in the decision problem. Mathematically, the expected opportunity loss of adopting a technology which is not cost effective and about which there is uncertainty given current evidence is the sum of the Payer

Strategy Burden of that technology and the Payer Uncertainty Burden for the decision problem. We define this as the Payer Strategy and Uncertainty Burden (P-SUB). Others may favour different terminology and this requires further consideration.

(3c) The Payer Optimality Gain (POG)

The Payer Optimality Gain is the reversed interpretation of the Payer Strategy Burden. It reflects how much money or health benefit can be saved over a comparator strategy by recommending a more cost-effective technology. The POG therefore quantifies the resources freed in the NHS by recommending the optimal strategy. The POG, in the same fashion as the PUB and PSB, can be presented over the population and for any time horizon.

(4) Analysing the Impact of Proposed MEAs

The taxonomy of MEAs has two main dimensions, effective price changes and further evidence collection.

To analyse the impact of an effective price change simply requires the proposed price change to be entered into the original cost-effectiveness model so that a revised PSA can be run. The results of the revised PSA generate revised estimates of the cost-effectiveness of the technologies, perhaps changing which is most cost-effective, and also the uncertainty around their costs and benefits. Crucially this also instantly generates a revised expected net benefit for each decision option, a revised Payer Uncertainty Burden for the decision, and P-SUBs of all strategies.

To analyse the impact of further evidence collection requires one additional step compared to that for effective price changes – simulation of the data from the proposed study. Technically, we need to specify a statistical model for the data that would be collected, e.g. a normal, or Weibull distribution, say, for a study with a specified sample size, follow-up period etc. For each ‘row’ of the PSA, we simulate a dataset from the proposed study, conditional on the parameters of that row. Given that simulated data, we calculate the implied model parameters for the decision model if only the data from the new study were used (e.g. the mean improvement in a disability score with the new technology, or the estimated parameters of a proportional hazards survival model). The extended Sheffield Accelerated Value of Information (SAVI) tool then estimates the impact on the decision given the new evidence, and generates revised P-SUBs of all strategies.

When these new values for decision risk under the MEA are compared with the original values from the decision problem without the MEA in place, the differences represent the value of the proposed MEA. Again this can be expressed in terms of a financial or QALY value per patient, annually for the prevalent population in England, and for a longer decision relevant time horizon.

Results – Feasibility of the Framework

The case study work undertaken in this report confirms that the adapted taxonomy of possible MEAs can be used to consider potential options in a structured way.

The case study work on previous NICE appraisals (pazopanib for renal cell carcinoma, biologic therapies for ulcerative colitis, imatinib for adjuvant treatment of GIST, drug powder inhalers in cystic fibrosis, chemotherapy options for metastatic colorectal cancer, trabectedin for soft tissue sarcoma, lenalidomide for treatment of MDS and dasatinib for CML) indicates that the MEA Risk Analysis Framework developed to quantify the decision risk in terms of the PUB and P-SUB of each technology in the decision problem is feasible. The MEA Risk Analysis Framework is relatively easy to implement for MEAs which affect prices using the uncertainty analysis methods and the SAVI tool. MEAs which involve further evidence collection can also be analysed successfully, provided a statistical model for the data that will be generated by the proposed evidence collection study can be described.

Analysis of a combined price adjustment and evidence collection MEA for the pazopanib case study was successfully undertaken. We found that the features of the pazopanib MEA, the price discount component and the further evidence collection via an additional trial, were impactful in two ways. Firstly, the price discount increased the expected net benefit of a decision to adopt pazopanib. Secondly, the combination of the price discount and the evidence collection substantially reduced the decision risk associated with recommending pazopanib, generating a much lower expected Payer Uncertainty Burden and lower Payer Strategy Burden for the decision options concerned.

In a hypothetical example that was based on the pazopanib appraisal data, but in which End of Life valuation did not apply, it was shown that the timing of research crucially determined

whether recommending a cost-ineffective technology with the condition of undertaking further research could be a worthwhile strategy.

Conclusions and Implications of MEA Risk Analysis Framework for NICE & Manufacturers

This report sets out a framework to evaluate systematically and routinely the decision risk in terms of Payer Uncertainty Burden and Payer Strategy Burden in technology appraisals. This evaluation requires *only* the outputs from a probabilistic sensitivity analysis, which is already part of standard NICE processes. The probabilistic sensitivity analysis is entered directly into the Sheffield Accelerated Value of Information (SAVI) framework, which identifies the key drivers of uncertainty and allows calculation of the Payer Uncertainty Burden and Payer Strategy Burden.

The taxonomy of possible MEAs can be used to consider what kind of MEA scheme might be most suitable in a particular context, together with our MEA design guidance questionnaire. Proposed MEA schemes have then been successfully evaluated in terms of increased expected net benefit and in terms of reductions in decision risk using the Payer Strategy and Uncertainty Burden measures.

According to our findings, recommending technologies that are not cost-effective based on available evidence is associated with a great risk of taking the wrong decision and consequently a cost to the payer that could reach much larger magnitudes than the cost of decision uncertainty alone. Effective price reductions can be a good option, when a strategy is not cost-effective. When there is large uncertainty, research schemes may be more appropriate although price reductions may also reduce uncertainty. Price reductions can reduce uncertainty, when the strategy for which the price reduction will be put in place is expected to be cost-effective. When a technology is not expected to be cost-effective and price is reduced, decision uncertainty may increase but if that technology consequently becomes cost-effective, the PUB may reduce again and the PSB reduces to zero, thus reducing the overall cost to the payer. Price reductions and research recommendations can be complementary in reducing the P-SUB. It is important to consider the timing of research when considering RwR schemes in an appraisal in which the recommended strategy is not cost-effective.

We also demonstrated that it can be desirable to conduct further research when a technology appeared to be cost-effective but had the caveat of large decision uncertainty. We consequently think that MEA schemes should be considered routinely, as a way of reducing uncertainty and the expected opportunity loss to the payer.

Because much of the MEA Risk Analysis Framework is already part of standard NICE processes, it would be straightforward to operationalise. Manufacturers or ERGs in the STA process, or AGs in the MTA process could provide estimates of PUB and P-SUB alongside existing analyses. An alternative that may be feasible is for these calculations to be performed in committee once there is agreement on the preferred set of parameter values for the base case analysis. Provided this is one of the scenarios that have already been performed, and the PSA results have been stored, the additional analysis can be undertaken "live". Once the committee is convinced that the exploration of MEAs is relevant then such analyses could be undertaken.

Scope for further research lies in applying the framework to further real world appraisals to gain experience, explore the practical considerations associated with the different schemes and provide the opportunity to review our framework at the time of re-appraisal.

We conclude that coherent, consistent and transparent assessments of proposed MEA schemes are critical. The MEA Risk Analysis Framework developed here provides a feasible and valuable means for such assessment. Implementation of the framework would permit NICE and its stakeholders to confidently appraise a wider range of decision options through MEAs.

CONTENTS

TABLES AND FIGURES

ABBREVIATIONS AND DEFINITIONS

CMA	Conditional Marketing Authorisation
CEAC	Cost-Effectiveness Acceptability Curve
EAMS	Early Access to Medicines Scheme
EoL	End of Life
ENBS	Expected Net Benefit of Sampling
EVI	Expected Value of Information
EVPI	Expected Value of Perfect Information
EVPPI	Partial EVPI
EVSI	Expected Value of Sample Information
GIST	GastroIntestinal Stromal Tumours
HTA	Health Technology Assessment
HTAi	Health Technology Assessment international
ICER	Incremental Cost Effectiveness Ratio
EMA	European Medicines Agency
MEA	Managed entry agreements
MHRA	Medicines and Healthcare products Regulatory Agency
OIR	Only in Research
PSB	Payer Strategy Burden
P-SUB	Payer Strategy and Uncertainty Burden
PUB	Payer Uncertainty Burden
PASs	Patient Access Schemes
PBRSAs	Performance-Based Risk Sharing Agreements
PIM	Promising Innovative Medicine
PSA	Probabilistic Sensitivity Analysis
QALYs	Quality Adjusted Life Years
RCT	Randomised Controlled Trial
RwR	Recommended with Research
SAVI	Sheffield Accelerated Value of Information

1. INTRODUCTION

1.1. PURPOSE

The purpose of this document is to set out methods for the assessment of Managed Entry Agreements that could be used within the NICE Technology Appraisals Programme.

1.2. RECENT DEVELOPMENTS BY REIMBURSEMENT AND REGULATORY AUTHORITIES

Recent changes to the regulation of pharmaceuticals are likely to lead to the licensing of medicines at an earlier stage of evidence development. At a European level, the European Medicines Agency (EMA) recently announced a new licensing scheme called "Adaptive Pathways". This ties in with the existing Conditional Marketing Authorisation (CMA) under which a one year license can be issued for a medicine with a positive risk-benefit balance that addresses a life-threatening condition, conditional on further research on their effectiveness or safety being conducted.[1] CMAs can be renewed subject to the manufacturer providing further study data and monitoring effectiveness and safety.

Adaptive pathways add more flexibility to the existing CMAs by allowing approval in a well-defined patient subgroup with a high medical need first[2] rather than issuing the license for the whole population at once. This approval can subsequently be extended to include a larger patient population or could lead to conditional licensing with more evidence collection required.

In the UK, the Medicines and Healthcare products Regulatory Agency (MHRA) can review technologies and potentially issue Promising Innovative Medicine (PIM) designations. Technologies with PIM designations can then become subject to the UK Early Access to Medicines Scheme (EAMS) under which patients may gain access to medicines before marketing authorisation.

In light of these developments, which are ultimately intended to make medicines available to patients sooner, it is highly likely that NICE will be required to issue guidance on technologies with an evidence base that is smaller or earlier in its evolution than has previously been the case. With greater uncertainty in both clinical and economic evidence,

schemes that allow the development of further evidence or that entail a risk-sharing component may be of particular value.

General decision options for a Health Technology Assessment (HTA) agency in relation to the use of a specific health technology include (a) Recommended, (b) Not recommended, (c) Recommended only for the use in a defined study population: Only in Research (OIR), or (d) Recommended only in conjunction with a Managed Entry Agreement (MEA): Options include Recommended with Research (RwR) schemes, price reduction schemes or a combination of these.[3, 4] The focus of this report is on (d) recommendation of a technology with a MEA scheme.

MEAs have been used in the past, particularly when clinical evidence was inconclusive. The term MEA was used in the Health Technology Assessment international (HTAi) Policy Forum 2010 and defined as *"an arrangement between a [pharmaceutical] manufacturer and payer/ provider that enables access to (coverage or reimbursement of) a health technology subject to specific conditions. These arrangements can use a variety of mechanisms to address uncertainty about the performance of technologies or to manage the adoption of technologies in order to maximise their effective use, or limit their budget impact"*.[5] Other terms describing the same are performance-based risk sharing agreements (PBRSAs)[6] and patient access schemes (PASs), the latter being commonly used in the UK to describe schemes entailing a reduction in the effective price.

In a review of the use of MEAs in different countries (that included PASs), 29 out of the 148 identified MEAs up to including 2013 were identified in the UK, which made the UK, in terms of MEA use, the second-most active country after Italy.[7] In the UK, MEAs have risen in popularity until 2012[8] after which their use has become slightly less frequent. Most of the UK schemes were made up of oncology arrangements and PASs are much more frequently used than coverage with evidence development schemes.[9]

1.3. OBJECTIVES OF THIS REPORT

This project aims to develop a framework to evaluate the desirability of different MEA schemes. Section 2 will describe our MEA Risk Analysis Framework which consists of a taxonomy of MEA schemes and a MEA design guidance questionnaire that helps make an

informed decision on the choice of MEA schemes. In Section 5, we demonstrate the feasibility of assessing risk in technology appraisals using the Payer Uncertainty Burden of the decision problem and the Payer Strategy Burdens associated with each strategy in the decision problem in four past appraisals using the online Sheffield Accelerated Value of Information (SAVI) tool. In Section 4, we assess the risk by calculating the P-SUBs in three further past appraisals and, in addition, consider the MEA scheme designs that could have potentially been used based on the P-SUB results and the MEA design guidance questionnaire. In Section 5, we describe how the full MEA Risk Analysis Framework was applied to an example past appraisal, pazopanib for the treatment of advanced renal cell carcinoma, and evaluate the reduction of risk associated with different MEA schemes. Section 6 will provide a discussion of the findings.

2. THE MEA RISK ANALYSIS FRAMEWORK

To develop a framework for choosing MEA schemes, we undertook four steps:

1. We conducted a brief literature review to find existing taxonomies of MEA schemes and synthesized and updated these to arrive at the taxonomy used as the foundation for this project.
2. We outlined the rationale for each of the MEA schemes.
3. We developed the MEA design guidance questionnaire that can assist with the choice of MEA scheme designs.
4. We developed the MEA Risk Analysis Framework to assess a) the need for and b) the value of MEA schemes.

2.1. AN UPDATED TAXONOMY OF MEAS

To develop an updated taxonomy of the different decision options within MEAs and their characteristics, we performed a brief review of the relevant literature. Knowing that MEAs have been reviewed and summarised in a comprehensive taxonomy (Walker *et al.*, 2012[10] and Walker *et al.*, 2012[3]) and another summary developed by Garrison *et al.*[6], which was further adapted by Bruegger,[9] we started our search from these publications. Further articles were found through "pearl-growing", that is the search of publications in which the aforementioned articles were referenced, as well as a reference search of the above-mentioned articles.

Apart from the above-mentioned work, eight other publications were included for review. These publications included an earlier review of MEAs with more detail provided on them (Stafinski *et al.*[11]), proposals of a leasing scheme (Edlin *et al.*, 2014,[12] Edlin *et al.*, 2013[13]) that had not been covered in detail by the four key publications, the use of MEAs since the three key publications (Ferrario & Kanavos,[8] Carlson *et al.*[7]) and two publications on good practice and principles (Klemp *et al.*,[5] Franken *et al.*[14]) complementing the ones set out in Garrison *et al.*[6] We also included the report by Claxton *et al.*[15] which furthermore provided an algorithm on the choice of OIR and RwR schemes.

The taxonomy provided by Walker *et al.*[10] lays out the following decision options on coverage. First, a distinction is made between reductions in effective price schemes (Figure 1) and evidence generation schemes (Figure 2).

Price reduction schemes can be outcome based or non-outcome based. Among the outcome based decisions, there are the following options:[3, 10]

A. Money back guarantee: Such a scheme can be offered by manufacturers for cases in which individual patients do not reach a specific target. The refund may not be at 100% and may be financial or non-financial (eg. replacing stock).

B. Conditional treatment continuation: This refers to continued payment of medicine only in patients who reached a specified target. These schemes can also be paired with money back guarantees.

C. Price linked to outcomes: Better outcomes for patients could be reflected in a higher price. Alternatively, the technology could be reimbursed at the time when the health outcome occurs rather than at the time of treatment. This latter option includes the Technology Leasing Reimbursement Strategy (TLRS) proposed by Edlin[13] where the payer pays for health outcomes when they occur rather than for the treatment when treatment is provided. It has to be noted that these TLRSs are mainly relevant to technologies that cause positive or negative health effects long after treatment has stopped, for instance in technologies such as implants without biologically active components and prostheses as well as certain cancer treatments.[12] They might become more

relevant with the emerging of other treatment with long-term benefits and often a short duration of treatment administration, such as gene therapy or tumor suppression therapy in certain conditions.[16]

The non-outcome based coverage decisions comprise individual and population level arrangements. Individual level non-outcome MEAs are:[10]

D. Discounted treatment initiation: Patients receive treatment cheaper at first and the price reverts to the list price after a time.

E. Patient utilisation cap: The cost of treatment is reduced (often to zero) after an agreed length of treatment.

F. Fixed cost per patient: These arrangements describe agreements on a fixed cost per patient regardless of the number of treatments actually received.

At the population level MEA options comprise:[10]

G. Discount: The negotiated price differs from the list price. Due to global reference pricing these price changes usually need to be kept confidential.

H. Expenditure cap: The total expenditure is limited without limiting the total quantity.

I. Price volume agreement: Such agreements could include nonlinear pricing schemes, where the price per unit may be reduced after reaching a certain volume of units.

Figure 1: MEA options: Reductions in effective price (adapted from Walker *et al.*,[10] Garrison *et al.*[6] and Bruegger[9])

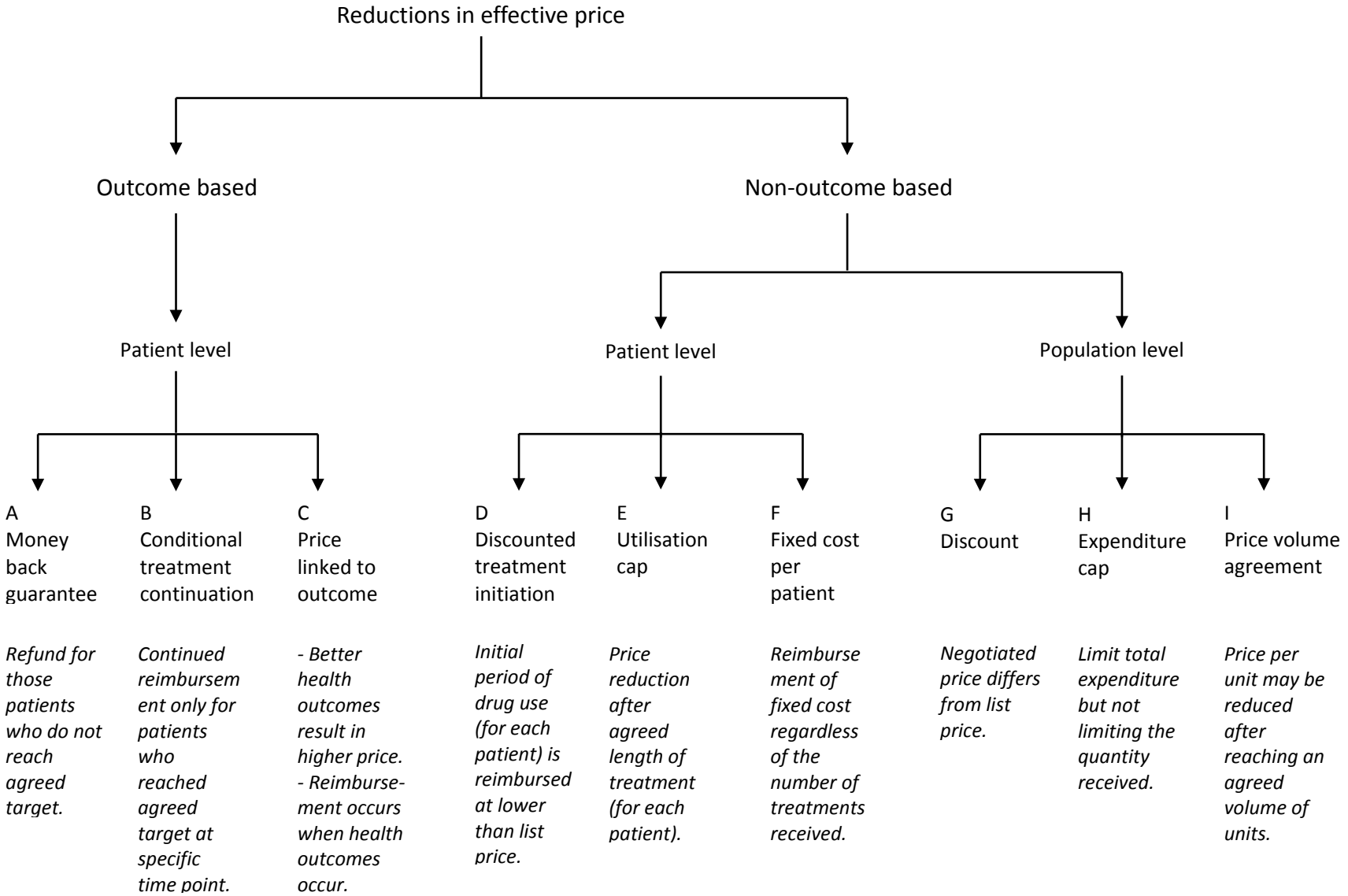
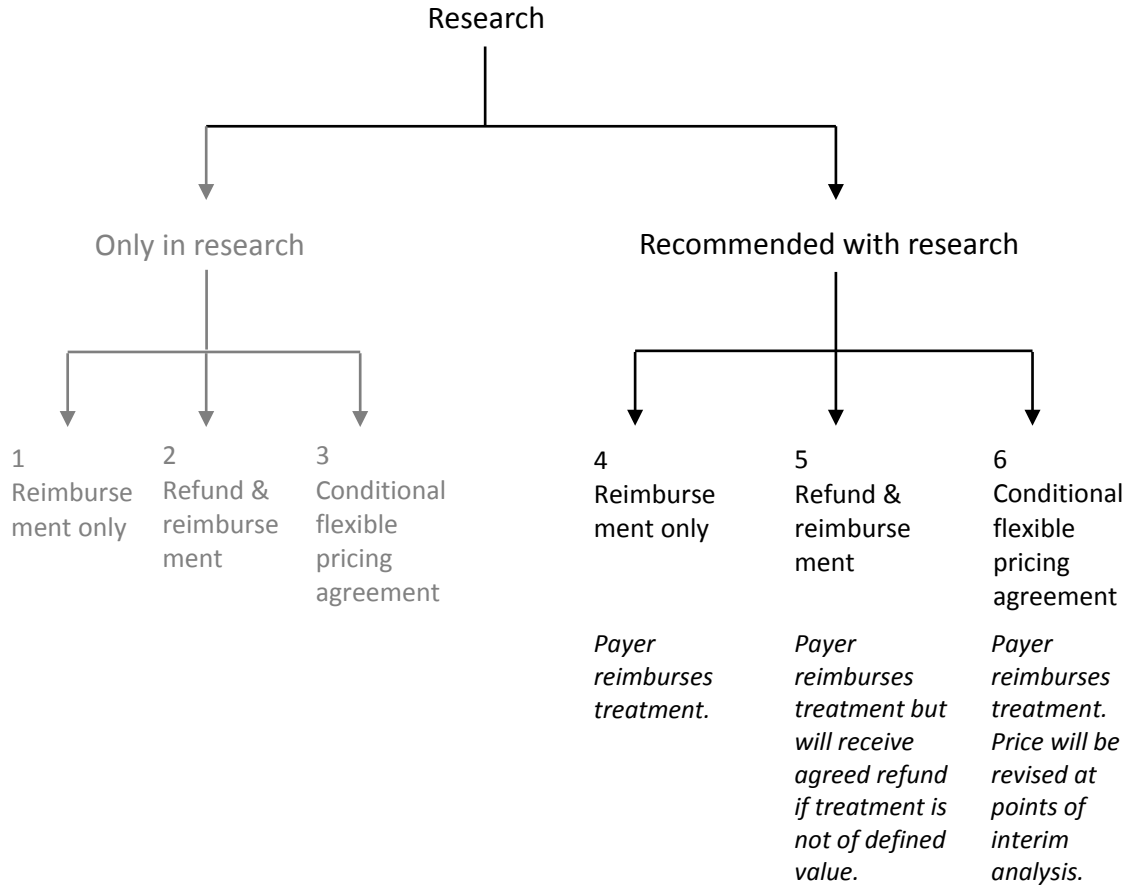


Figure 2: MEA options: Research (adapted from Walker *et al.*,[10] Garrison *et al.*[6] and Bruegger[9])



There are two types of evidence generation schemes: Only in Research (OIR) which is effectively not recommending a new technology but reimbursing the technology in a study population to generate more evidence for review at a later stage. OIR decisions do not facilitate widespread access to the technology and are thus excluded from further analysis (and printed in grey). Recommended with Research (RWR) decisions are favourable funding decisions but with the condition of more research being conducted and a review scheduled for a later stage. Both types of evidence generation schemes have the following options:

1., 4. Reimbursement only: In a reimbursement only decision, the payer agrees to reimburse the manufacturer for treatment of patients enrolled in the trial (OIR) or all patients (RwR). The research itself may be funded by the payer, the manufacturer or another stakeholder.[9]

2., 5. Refund and reimbursement: In a refund and reimbursement decision, the payer will receive an agreed refund or rebate in the case that research showed that the technology is not of value.

3., 6. Conditional flexible pricing agreement: A conditional flexible pricing agreement between the payer and the manufacturer gives the possibility to directly link price to the results of research. This may include a refund agreement. An example of this scheme is the use of interferon-β or glatiramer acetate for the treatment of multiple sclerosis in the UK NHS. Both treatments were recommended on the condition that their effectiveness was monitored for ten years. Price adjustments could be made every two years in order to ensure the ICER would not exceed £36,000 per QALY.[3]

Historically speaking, evidence generation schemes were the most frequently used type of MEAs in all countries included in the previously mentioned review by Carlson *et al.*[7] In the UK, however, a trend was noted to simpler and less administratively burdensome schemes, such as discounts.[7] In fact, most schemes were simple discounts[17] and the focus was on improving cost-effectiveness rather than addressing uncertainty.[8]

With this trend potentially changing towards a greater use of evidence generation schemes, it is important to consider how a choice between different MEA schemes can be made. Some discussions on the desirability of MEA schemes are available from different sources. For instance, Franken *et al.*[14] state that RwR schemes have been faced with the criticism of being inefficient, considering that the value of information was often lower than the cost associated with them. To avoid such inefficiencies, assessing the expected value of information and the cost of collecting evidence was considered vital by Garrison *et al.*[6]

Claxton *et al.*[15] developed a much more detailed algorithm to help with the decision on whether to take a RwR decision or not. The main points to consider when assessing the desirability of a RwR scheme were the expected cost-effectiveness, the need for evidence and whether or not the type of research required can be conducted once a technology is approved for widespread use, whether or not there are sources of uncertainty that cannot be resolved by research but only over time and whether or not there are significant (opportunity) costs that will be committed and cannot be recovered once the technology is approved.

The algorithm developed by Claxton *et al.*[15] helps clarify the process to go through when faced with an uncertain decision. In particular, it is very useful in deciding whether more research could and should be conducted. However, we are not aware of any formal method of assessing the suitability of different MEA schemes in circumstances in which the guidance of Claxton *et al.*[15] would conclude that a RwR scheme is recommended.

Essentially, we do know which criteria to apply to consider the use of MEA schemes for a particular technology. But what we do not know is which MEA to choose. This decision is much more likely to be influenced by quantitative assessments of the situation the decision-maker is faced with. For example, Claxton *et al.*[15] state that in some cases a RwR decision may be appropriate even in technologies that are not expected to be cost-effective. What is required in this instance is a way of assessing the burden of such a decision with the present evidence and whether this burden can be offset by the benefit of future research.

More generally, the value of further evidence needs to outweigh the associated costs but also some evidence generation schemes may offer more value than others. For this reason, we outline rationales for each MEA option. This is not to duplicate previous work but to complement it and make the decision for MEA schemes more transparent.

2.2. RATIONALES FOR DIFFERENT MEA SCHEMES

The rationales for each MEA option are similar across their purposes. RwR schemes should reduce uncertainty and the payer's risk of taking the wrong decision. Price reduction schemes should reduce the payer's risk by making the technology more affordable, thereby reducing the cost associated with taking a, potentially wrong, decision. Slight differences are however observable and summarised in Table 1 and Table 2, together with examples of possible uses.

Table 1: Rationales for price reduction schemes

Price reduction schemes	Rationale	Possible use
Money back guarantee	Bring costs down and reduce payer risk surrounding the success or failure of treatment	Decision uncertainty is mainly associated with treatment success and failure.
Conditional treatment continuation	Bring costs down and reduce payer risk surrounding the success or failure of treatment	There is a decrement to the likelihood that a treatment results in success after a certain length of time.
Price linked to outcome	Bring costs down and reduce risk surrounding a set of health outcomes	There are different possible health states resulting from treatment that are associated with different probabilities, utilities and costs, which are the cause of decision uncertainty. Alternatively, there may be a long lag between treatment and health outcomes.
Discounted treatment initiation	Bring costs down	Treatment is too expensive and utility gain occurs after a certain period of treatment.
Patient utilisation cap	Bring costs down and avoid excessive treatment	There is no further benefit after certain length of treatment or dose.
Patient cost cap	Bring costs down and enable patients to benefit from treatment after reimbursement has stopped	The length of treatment until response is achieved is highly uncertain.
Discount	Bring costs down	Treatment is simply too expensive.
Expenditure cap	Control budget impact	Treatment is prohibitively expensive for the health system.
Price volume agreement	Control budget impact	There are economies of scale.

Table 2: Rationales for RwR schemes

Recommendation with Research schemes	Rationale	Possible use
Reimbursement	Reduce payer risk caused by a group of uncertain parameters that can be studied in a feasible number of studies	Technology is expected to be cost-effective or close to cost-effective but there is great uncertainty.
Refund & reimbursement	Reduce payer risk caused by a group of uncertain parameters that can be studied in a feasible number of studies and reduce payer risk of reimbursing treatment that is not cost-effective	Technology is not expected to be cost-effective and there is great uncertainty.
Conditional flexible pricing agreement	Reduce payer risk caused by a group of uncertain parameters that can be studied in a feasible number of studies and set price according to health outcomes	Technology is just cost-effective or not cost-effective and there is great uncertainty and there is long-term research with quantifiable health outcomes at interim analyses.

In principle, evidence generation schemes can be combined with price reduction schemes. Combinations of schemes may provide even greater value, with potential synergies to be had when effective price reduction schemes complement RwR schemes. Price reductions tend to reduce some uncertainty if the technology is expected to be cost-effective[15] but may increase uncertainty if the technology in question was not expected to be cost-effective. If price reductions are considered, it is therefore vital to assess uncertainty with and without such a price reduction and then evaluate the value of further research for both.

2.3. ISSUES AND GUIDANCE ON THE CHOICE OF MEA SCHEMES

While we have provided possible uses for each MEA, there may be a wide variety of scenarios that warrant the choice of different MEA schemes. There cannot be a prescribed way of choosing MEAs as the decision does not only depend on binary factors such as whether some defined condition is fulfilled or not, but also on the scale of the different factors that may influence the decision. A quantification of risk is therefore unavoidable.

It may be useful to consider the following questions to help clarify issues and obtain some guidance on the choice of MEAs (Table 3) first. Some of the questions in the MEA design guidance questionnaire can only be fully answered once the MEA risk analysis framework proposed in Sections 2.4 and 2.5 has been undertaken.

Table 3: MEA design guidance questionnaire

1. What are the (number and characteristics of) treatment options?
2. What is the base-case cost-effectiveness?
3. What is the nature and scale of risk in this appraisal?
3.a What is the nature and scale of risk captured by the PSA?
3.b What is the nature of uncertainty not captured by the PSA?
3.c What is the temporal nature of uncertainty, e.g. is there more uncertainty beyond the trial period or is it resolvable with open-label follow up?
4. What is the uncertainty caused by individual / groups of parameters?
5. What alternative treatment strategies might be available?
6. What measures of patient-based outcomes are available and measurable?
7. Is price a substantial part of overall costs associated with treatment?
8. Are there any precedent PASs in place?
9. Could price agreements be national or local?

2.4. THE MEA RISK ANALYSIS FRAMEWORK PART I: ASSESSING THE NEED FOR AND BROAD DESIGN OF MANAGED ENTRY AGREEMENTS

2.4.1. Overview of Approach - 5 Questions Decision Makers & Stakeholders Ask

In this section we set out

- an analytical framework to estimate three useful quantified measures of risk
    - the Payer Uncertainty Burden (PUB) and
    - the Payer Strategy Burden (PSB) and
    - how they combine to estimate Payer Strategy and Uncertainty Burden (P-SUB – pronounced 'Pee-SUB')
- how these measures are simple to calculate because they are extensions of the existing cost-effectiveness analyses already used in NICE technology appraisals
- how these measures can be interpreted by stakeholders to enable a structured comparison of the need (or otherwise) for an MEA from one appraisal to the next,
- and how they can support decision makers, analysts and manufacturers' assessment of whether a proposed MEA should entail effective price reductions, or further evidence collection, or both.

In order to explain the conceptual framework, we will take a route through an illustrative example technology appraisal and we will explain the PUB, PSB and P-SUB measures in terms of concepts already well understood in the current NICE appraisal process. A more concise overview of the concepts used and mathematical expressions are in the appendix.

As a starting point, we assume that we have a manufacturer submitted or Assessment Group cost-effectiveness model of the technology of interest and appropriate comparators. For the purposes of describing the analytical framework, we will also assume, for now, that the model is 'accepted as reasonably representing the decision problem'. (Later in this report we discuss what can be done in circumstances when the NICE committee or ERG have concerns about the model structure and/or parameter values.) For now, we assume there is a reasonable health economic decision model, in which the appropriate comparators are analysed and for which a reliable probabilistic sensitivity analysis has been run.

In considering the need for a managed entry agreement, we will imagine the ERG, NICE committee and other stakeholders asking a sequence of 5 questions as shown in Table 4.

Table 4: Five Key Questions in Establishing the Potential Need for a MEA

Five Key Questions in Establishing the Potential Need for a MEA in a Technology Appraisal
Q1) Which intervention do we expect to be most cost-effective given proposed prices and current evidence?
Q2) How uncertain are we?
Q3) How useful would it be to eliminate uncertainty?
Q4) Given current evidence and proposed prices, what is the strategy-specific risk to the NHS?
Q5) How much would the NHS expect to gain by eliminating the risks associated with both uncertainty and the strategy?

It is worth going through these questions a little more slowly with our simple illustrative example because this will help to explain both the concepts and the rationale for their use.

2.4.2. Q1. Which of the interventions is estimated by the modelling to be the option we expect to be most cost-effective given proposed prices and current evidence?

This first question is part of standard practice in NICE appraisals. It is answered by estimating the expected costs and QALYs of each option using the cost-effectiveness model and taking the average costs and QALYs from the probabilistic sensitivity analysis (PSA)

results. The incremental cost-effectiveness ratio of each intervention versus each of the other comparator interventions can be calculated. If the resulting ICER for a new intervention versus current care is lower than the maximum acceptable ICER (for example if the ICER for a new intervention is £10,000 and the maximum acceptable ICER (MA.ICER), denoted by the Greek letter lambda, is λ=£20,000 per QALY) then a decision maker would conclude that the new intervention is expected to be more cost-effective than current care given proposed prices and existing evidence, and that it would be likely to be recommended for use in the NHS.

Our simplified hypothetical model is illustrated in Figure 3. This shows we have run a PSA with just 10 samples – of course normally we would run say 10,000 samples but the small numbers will help us to follow the logic of all the calculations through the steps in Q1 to Q5. The QALYs obtained from the new treatment vary from 8 to 8.4 in the different PSA samples and the average QALYs for the new treatment is 8.2. The QALYs for current care are 8.1 in each PSA run (this simple model assumes there is no uncertainty for current care). Therefore, the average incremental difference in QALYs is 8.2 minus 8.1 = +0.1 QALYs in favour of the new treatment versus current care. The costs of the new intervention (£8000) and current care (£7000) are also assumed to have no uncertainty, so the mean difference in costs is £1,000. The ICER is the ratio of £1,000 incremental costs divided by 0.1 QALYs gained, which is £10,000 per QALY gained.

Figure 3: Simplified hypothetical model – PSA Results and ICER calculation for Q1.

MA.ICER	£20,000			
	QALYs		Costs	
PSA Run	New Treatment	Current Care	New Treatment	Current Care
1	8.4	8.1	£ 8,000	£ 7,000
2	8.4	8.1	£ 8,000	£ 7,000
3	8.3	8.1	£ 8,000	£ 7,000
4	8.2	8.1	£ 8,000	£ 7,000
5	8.2	8.1	£ 8,000	£ 7,000
6	8.2	8.1	£ 8,000	£ 7,000
7	8.2	8.1	£ 8,000	£ 7,000
8	8.1	8.1	£ 8,000	£ 7,000
9	8	8.1	£ 8,000	£ 7,000
10	8	8.1	£ 8,000	£ 7,000
Average	8.2	8.1	£ 8,000	£ 7,000
	Incremental Costs - New v Current			£1,000
	Incremental QALYs - New v Current			0.1
	ICER - New v Current			£10,000

The net monetary benefit measure provides another simple way of summarising which intervention is expected to be most cost-effective, and we present a discussion of net monetary benefit here because it is used to calculate the PUB, PSB and P-SUB measures. The net monetary benefit measure converts the QALYs gained by an intervention into a financial valuation, assuming the value of 1 QALY is say £20,000 (or whatever MA.ICER is appropriate), and then nets off the costs of the intervention. So, the net monetary benefit for the new treatment in the first PSA run is calculated as 8.4 QALYs * £20,000 (which is £168,000) minus the £8,000 cost, which equals £160,000. Figure 4 calculates net monetary benefit for each PSA run and for both treatments. We can see that the average (also called 'expected') net monetary benefit for the new treatment is £156,000. The expected net monetary benefit for current care is £155,000. So, there is a difference in favour of the new intervention of £1,000. We can see intuitively that this calculation is correct by looking at the average incremental costs and QALYs. Essentially we are financially valuing the expected 0.1 QALYs difference gained by using the new intervention (0.1 QALYs * £20,000 =

£2,000) and netting off the expected incremental costs (£1,000) so that the expected *incremental* net monetary benefit of the new treatment is £1,000.

The expected net monetary benefit measure is particularly useful when comparing across several interventions, because it is much quicker to interpret than doing all of the ICERs of each treatment against all of the others. The interpretation is simple. The intervention which has the highest expected net benefit is the one which we expect to be most cost-effective. In our example, it is the new treatment which has the highest *expected* (i.e. averaging over the PSA runs) net monetary benefit. Any other intervention than this would not be expected to be cost-effective.

Figure 4: Simplified hypothetical model – Expected Net Monetary Benefit for Q1.

MA.ICER	£20,000					
	QALYs		Costs		Net Benefit = QALYs * MA.ICER - Cost	
PSA Run	New Treatment	Current Care	New Treatment	Current Care	New Treatment	Current Care
1	8.4	8.1	£ 8,000	£ 7,000	£160,000	£155,000
2	8.4	8.1	£ 8,000	£ 7,000	£160,000	£155,000
3	8.3	8.1	£ 8,000	£ 7,000	£158,000	£155,000
4	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000
5	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000
6	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000
7	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000
8	8.1	8.1	£ 8,000	£ 7,000	£154,000	£155,000
9	8	8.1	£ 8,000	£ 7,000	£152,000	£155,000
10	8	8.1	£ 8,000	£ 7,000	£152,000	£155,000
Average	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000
					Highest expected net monetary benefit	Expected to be Suboptimal
	Incremental Costs - New v Current			£1,000		
	Incremental QALYs - New v Current			0.1		
	ICER - New v Current			£10,000		

It turns out that the two methods of working out which intervention is expected to be cost-effective given proposed prices and current evidence, i.e. net monetary benefit calculations and ICER calculations, are mathematically equivalent, so either or both can be used.

Note also that the net *monetary* benefit can be converted into a net *health* benefit measure. This is done by dividing the net monetary benefit by the maximum acceptable ICER (£20,000), so that for example the expected net monetary benefit for the new intervention of £156,000 is equivalent to an expected net health benefit of 7.80 QALYs. The expected net monetary benefit for current care of £155,000 is equivalent to an expected net health benefit of 7.75 QALYs. And so, the expected difference between the two, the incremental net monetary benefit of £1,000 is equivalent to an expected incremental net health benefit of 0.05 QALYs in favour of the new intervention.

Next, we consider measures of uncertainty as expressed in the PSA in more depth.

2.4.3. Q2. How uncertain are we? What is the probability that one of the other interventions is actually more cost-effective than the 'option we expect to be most cost effective given proposed prices and current evidence'?

Answering this question is also standard in NICE appraisals, using *Probabilistic Sensitivity Analysis (PSA)*. The existing evidence base is examined and the modelling characterises the uncertainty in model inputs e.g. the uncertainty in efficacy and effectiveness of each intervention, costs of interventions, longer-term disease progression and natural history, the occurrence of complications, co-morbidities, adverse events, effects on health related quality of life and downstream costs to the health service etc. The Monte Carlo sampling from the statistical distribution used in the PSA enables an understanding of how allowing all of these uncertain inputs to vary at the same time within the model affects results. Each PSA result is a single sample and produces the costs and QALYs for each intervention if the model inputs were exactly at the sampled values. The two main visual illustrations of uncertainty provided to NICE are the cost-effectiveness plane and the cost-effectiveness acceptability curve.

The C-E plane, as shown in Figure 5 for our model, gives a visual understanding of the uncertainty in incremental QALYs and incremental costs. Each PSA run is a 'dot'. In our example, the incremental QALYs are uncertain and range from -0.1 (i.e. current care would be better than the new intervention) to +0.3 QALYs. Incremental costs are not uncertain and each PSA run has an incremental cost of +£1,000. The average (over all PSA runs) incremental cost is £1,000 and average incremental QALYs are +0.1, and this point is shown as a diamond. The 45 degree line is the maximum acceptable ICER, in this case drawn with a slope of £20,000 per incremental QALY gained. The fact that the diamond for the expected

incremental costs and QALYs is below and to the right of the MA.ICER line indicates that the new intervention is expected to provide additional QALYs at a cheaper incremental rate than the MA.ICER (i.e. the new intervention costs £10,000 per additional QALY, which is below λ=£20,000). So, the interpretation is that, given proposed prices and current evidence, the new intervention is expected to represent a cost-effective use of NHS resources. A PSA run dot below and to the right of the MA.ICER 45 degree line indicates that, if the true values of the uncertain model inputs were at the levels sampled in that PSA run, then the new intervention would be cost-effective. In our case 7 of the 10 dots are below the MA.ICER line. But also, three dots are above and to the left of the MA.ICER line. A PSA run dot above and to the left of the MA.ICER 45 degree line indicates that, if the true values of the model inputs were at the levels sampled in that PSA run, then the new intervention would not be cost-effective. That is, the decision maker would be better off sticking with current care and not recommending the new intervention. Indeed, in our example, for some PSA runs the new intervention is not even effective because it would provide fewer QALYs than current care.

Figure 5: Simplified hypothetical model – Cost-Effectiveness Plane for Q2.

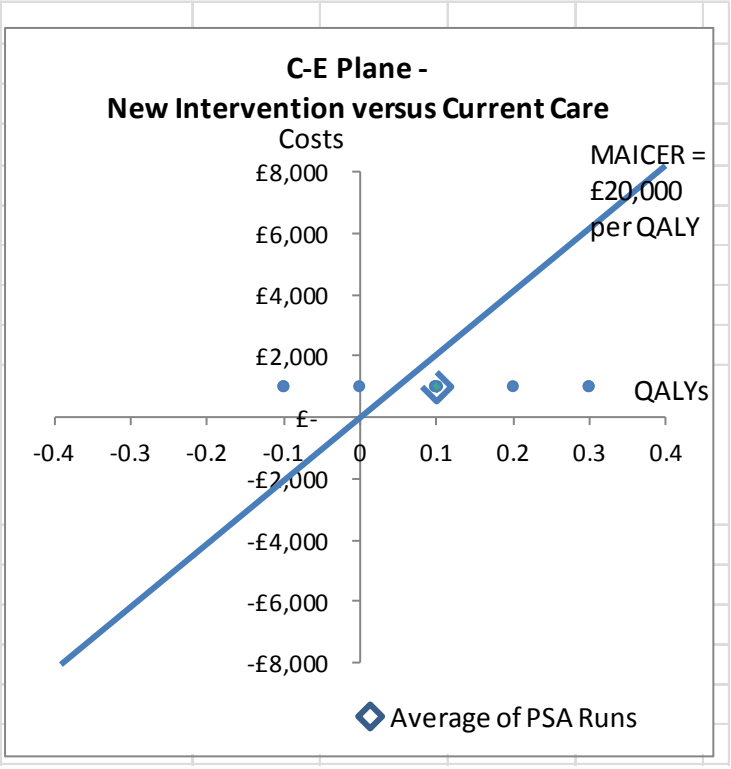


Figure 6 shows the cost-effectiveness acceptability curve for our example model. The cost-effectiveness acceptability curve (CEAC) plots (on the y-axis) the proportion of PSA runs for

which each intervention would be considered the most cost-effective, against different values of the MA.ICER on the x-axis[18]. In a comparison of two strategies, it simply shows the proportion of PSA run dots that are below and to the right of the MA.ICER line. In our example, 70% (7 of the 10 PSA runs) show the new intervention being cost-effective at a MA.ICER of £20,000 per QALY. The x-axis on a CEAC allows for varying the MA.ICER and investigates how the probability that an intervention is cost-effective responds to different values of λ. The CEAC is easy to extend to comparison of several strategies.

Figure 6: Simplified hypothetical model – CEAC for Q2.

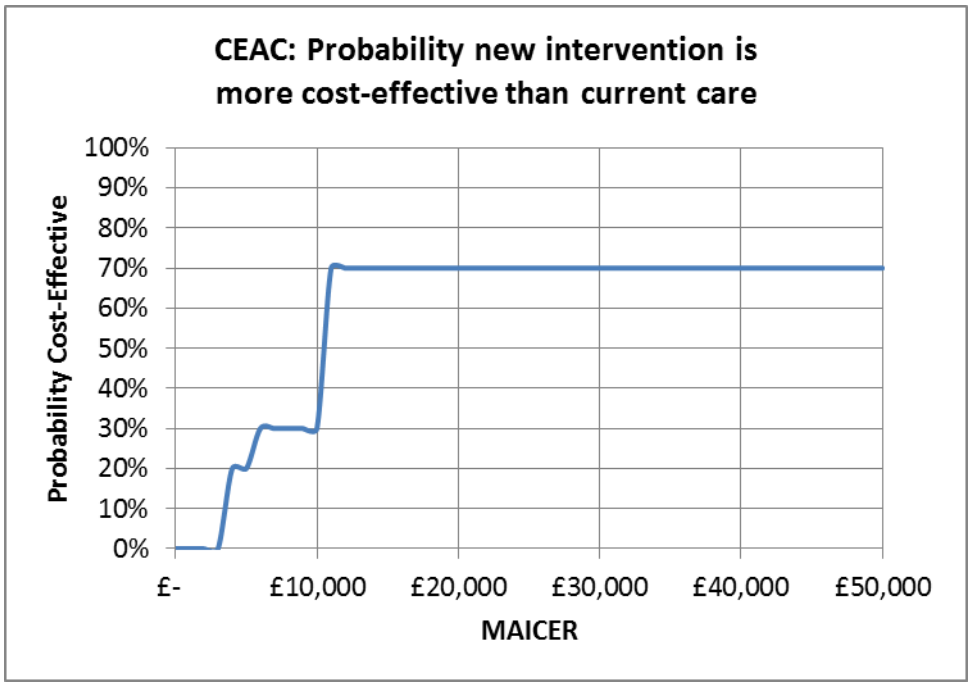


Figure 7 shows the spreadsheet model extended slightly to clarify which PSA runs show either the new intervention or current care as most cost-effective. Again, we can see 7 out of 10 PSA runs show the new intervention providing higher net monetary benefit than current care. If the true values of the uncertain model inputs were at the levels sampled in these PSA runs, then the new intervention would be the most cost-effective. For the last 3 PSA runs, it is the other way around, the new intervention provides lower net monetary benefit than current care, and so current care would be more cost-effective.

Overall, averaging over all the PSA runs, given proposed process and current evidence, we *expect* the new intervention to provide more than current care on the net monetary benefit measure (and equivalently on the net health benefit measure), but there is a minority 30%

chance that we are wrong about this and current care could actually be the more cost-effective strategy providing more net monetary benefit.

Figure 7: Simplified hypothetical model – PSA Runs for Q2.

MA.ICER	£20,000				
	Net Benefit = QALYs * MA.ICER - Cost			Probability Intervention is most cost-effective	
PSA Run	New Treatment	Current Care	Difference in Net Monetary Benefit	New Treatment most cost-effective?	Current Care most cost-effective?
1	£160,000	£155,000	£5,000	Yes	No
2	£160,000	£155,000	£5,000	Yes	No
3	£158,000	£155,000	£3,000	Yes	No
4	£156,000	£155,000	£1,000	Yes	No
5	£156,000	£155,000	£1,000	Yes	No
6	£156,000	£155,000	£1,000	Yes	No
7	£156,000	£155,000	£1,000	Yes	No
8	£154,000	£155,000	-£1,000	No	Yes
9	£152,000	£155,000	-£3,000	No	Yes
10	£152,000	£155,000	-£3,000	No	Yes
Average	£156,000	£155,000	£1,000	70%	30%
	↑	↑		↗	↗
	Highest expected net monetary	Expected to be Suboptimal		More likely to be cost-effective	Less likely to be cost-effective

In the next sub-section, we *imagine*, having answered Q1 and Q2, that the decision maker (e.g. the NICE Technology Appraisal Committee) recommends the 'option we *expect* to be most cost effective given current evidence' and that the NHS follows the recommendation. We then ask the third question concerning how important is the uncertainty in the evidence base.

2.4.4. Q3. Given proposed prices, how much more health benefit or cost savings could the NHS expect to get if we were able to eliminate all of the current uncertainty in the evidence? This we call the Payer Uncertainty Burden – PUB

The rationale here is that, if the current uncertainty in the evidence base were able to be eliminated, then the decision maker would be able to choose definitively the best option rather than choosing the 'option we *expect* to be most cost effective given current evidence'. It is a simple extension of the PSA results to calculate the expected value of eliminating uncertainty in the current evidence.

Figure 8 demonstrates how this calculation of the value of eliminating all uncertainty in the current evidence works. In our hypothetical model, as we have seen, the decision maker's expected most cost-effective option is the new intervention. The current uncertainty in evidence is reflected in the different model inputs in the PSA. That is, the true underlying value of the costs and QALYs could be any one of the PSA rows. If all uncertainty were eliminated for the decision maker and it was discovered that the true underlying values were as row 1 of the PSA, then the decision maker would choose the new intervention (because it would be known for certain that it has higher incremental net monetary benefit). And, this decision would be no different to the decision made with current evidence. So, in this case, there would be no additional gain to the NHS if uncertainty were eliminated and values were found to be as per row 1 in the PSA, because the decision the decision maker would stick with is the new intervention, and the NHS would get the QALYs and costs associated with the new intervention. The same is the case if uncertainty were eliminated and it was discovered that the true underlying values were as per rows 2, 3 etc. through to 7 in Figure 8.

In row 8 however, things are different because the decision 'switches'. If all uncertainty were eliminated and it was discovered that the true underlying values of the model inputs were as per row 8, then the decision maker would switch to choosing current care. Why? – because it would now be known that current care is truly more cost-effective than the new intervention – indeed, current care would provide £155,000 versus £154,000 i.e. £1,000 more net monetary benefit (or in net health benefit terms 0.05 more QALYs). In rows 9 and row 10, a similar situation occurs. If the uncertainty were eliminated and it turned out that the true underlying values of the model inputs were as row 9 (or row 10 which is identical), and if the decision maker knew that for certain, then the decision maker would switch to choosing current care, this time with £3,000 more net benefit.

So, overall, we have a 70% chance that the decision made using current evidence is truly the most cost-effective, and in these cases the decision maker would not switch decision, but also a 30% chance of finding out the current care is truly the most cost-effective, and then switching the decision and getting £1,000 (if row 8 is true), £3,000 (row 9) or £3,000 (row 10) more net monetary benefit than with the new intervention. Finally, for our calculation, we average over all 10 PSA runs. This enables us to calculate the *expected* gain in monetary net benefit if all uncertainty were eliminated, and in our hypothetical model this is £700 per person affected by the decision (equivalent to an expected gain in QALYs of 0.035 per person). We call this the Payer Uncertainty Burden (PUB).

Figure 8: Simplified hypothetical model – Payer Uncertainty Burden Calculation for Q3.

MA.ICER	£20,000					
	Net Benefit = QALYs * MA.ICER - Cost			Probability Intervention is most cost-effective	Gain in net monetary benefit if	
PSA Run	New Treatment	Current Care	Difference in Net Monetary Benefit	New Treatment most cost-effective?	all uncertainty were eliminated and decision maker could switch to true optimal	Gain in Net health benefit
1	£160,000	£155,000	£5,000	Yes	£0	-
2	£160,000	£155,000	£5,000	Yes	£0	-
3	£158,000	£155,000	£3,000	Yes	£0	-
4	£156,000	£155,000	£1,000	Yes	£0	-
5	£156,000	£155,000	£1,000	Yes	£0	-
6	£156,000	£155,000	£1,000	Yes	£0	-
7	£156,000	£155,000	£1,000	Yes	£0	-
8	£154,000	£155,000	-£1,000	No	£1,000	0.050
9	£152,000	£155,000	-£3,000	No	£3,000	0.150
10	£152,000	£155,000	-£3,000	No	£3,000	0.150
Average	£156,000	£155,000	£1,000	70%	£700 PUB	0.035
	Highest expected net monetary benefit	Expected to be Suboptimal		More likely to be cost-effective	Expected (average) gain in monetary net benefit if all uncertainty were eliminated	

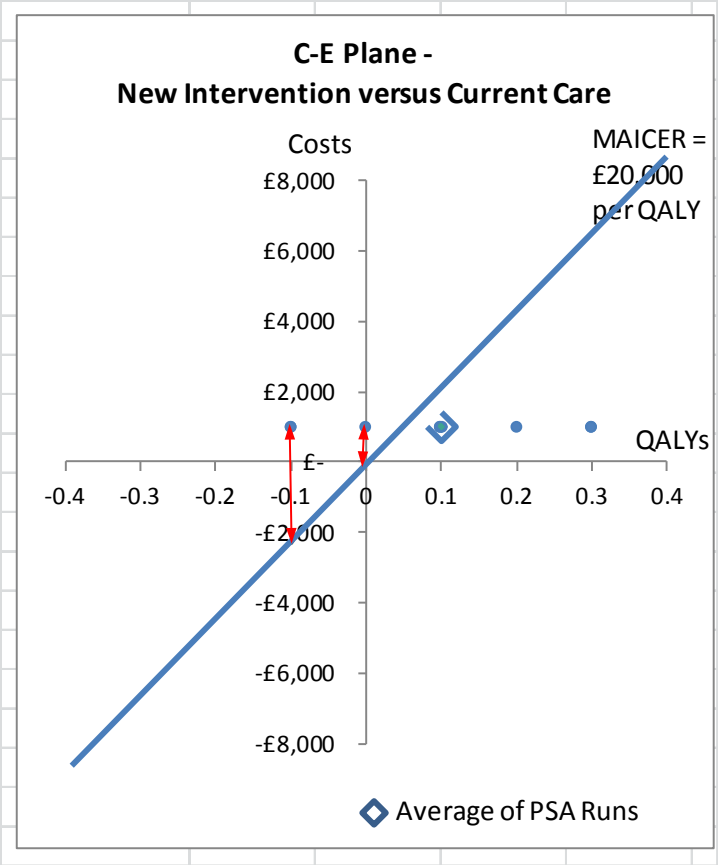
The scale of the Payer Uncertainty Burden is determined by a combination of two factors, and thinking about each factor in terms of the cost-effectiveness plane helps to develop an

instinctive understanding of the PUB. The first factor is the probability that one of the other interventions is truly more cost-effective than the one we expect given current evidence, in other words the probability we have got the decision, based on current evidence, wrong. If the probability of another intervention being cost-effective is very low then the Payer Uncertainty Burden will be low. So, visually, in a technology appraisal where there are only a very small number of 'dots' on the wrong side (i.e. above and to the left) of the MA.ICER line in a cost-effectiveness plane diagram, then the PUB will be small. Or similarly, if the CEAC shows the probability that the new intervention is cost-effective at say 99.9%, then the PUB will be small. In contrast, in a technology appraisal where the decision between two interventions is borderline and, say 49% of the dots are above the line, then all else being equal, the PUB would be much larger.

The second factor determining the scale of the Payer Uncertainty Burden is how much cost and QALY would be gained in those circumstances where uncertainty is eliminated and the decision would switch. One could imagine an appraisal similar to our hypothetical model but with say 10 times more difference in net monetary benefit in those PSA runs 8 to 10 where current care is more cost-effective. There would still be a 30% chance of current care being most cost-effective. But the PUB when calculated as the average over the PSA runs would be £7,000 per person (equivalent to 0.35 QALYs per person) i.e. 10 times more than the £700 PUB in our example. Again, we can get a visual instinct for this by looking at the cost-effectiveness plane. A dot with a greater vertical distance from the 45 degree MA.ICER line indicates that a larger net monetary benefit would be obtained by switching the decision.

Indeed, the vertical distance to the MA.ICER line is *exactly* the net monetary benefit difference. This can be seen in Figure 9. PSA run 8 is the dot with zero QALY difference and a £1,000 cost difference. We saw in Figure 8 that the gain in net monetary benefit of eliminating uncertainty and finding the true values of the model inputs are the same as in row 8 of the PSA is £1,000, which is exactly the vertical distance from that dot to the MA.ICER line as shown by the red double ended arrow. Similarly for rows 9 and 10 in the PSA, the net monetary benefit gain of eliminating uncertainty and finding that the true values of the model inputs are as in row 9 or 10 of the PSA is £3,000, which is the vertical distance from MA.ICER line to the relevant dot on the plane (situated at -0.1 QALYs and +£1,000 incremental costs).

Figure 9: Simplified hypothetical model – The Vertical Distance of a 'dot' in Cost-Effectiveness Plane provides a visual understanding of the Payer Uncertainty Burden for Q3.



So, a visually derived indication of the scale of Payer Uncertainty Burden can be obtained by looking at these two factors on the cost-effectiveness plane: (i) the proportion of dots on the wrong side of the line (probability our decision based on current evidence is wrong), and (ii) the vertical distance from each of those dots on the wrong side to the MA.ICER line. The dots below the MA.ICER line have zero value in terms of expected gain when eliminating uncertainty because finding out exactly where you are at below the line does not cause the decision maker to switch decision from the option expected to be most cost-effective.

Clearly the PUB will be different for different appraisals. Figure 10 shows four examples of different C-E planes on the same scale to illustrate how the two factors determine the scale of the PUB. Two of the examples have a probability of cost-effectiveness of just 14% (at MA.ICER = £20,000) suggesting small levels of uncertainty, but because example c) has a wider spread than example a), and larger vertical distance from the MA.ICER line, the value of eliminating uncertainty is substantially larger (£227 v £59 per person). Example b) has a similarly scaled PUB to example c) because although the spread of uncertainty is smaller,

there are many more dots on the wrong side of the line (49%). Example d) has a high number of dots a large distance away from the line, resulting in a very large PUB per person (£1794). Note that at a MA.ICER of £20,000, the net health benefit measure of the PUB in these four examples are a) 0.003 QALYs per person affected by the decision, b) 0.011, c) 0.011, d) 0.090 or in terms of days of perfect health a) 1, b) 4, c) 4, and d) 33.

This same calculation of the value of eliminating uncertainty in the current evidence base can of course be undertaken when there are several interventions compared rather than just two. The PUB is this expected value per person affected by the decision of eliminating all uncertainty in the current evidence base. It is also called overall expected value of perfect information in the health economics methods literature.

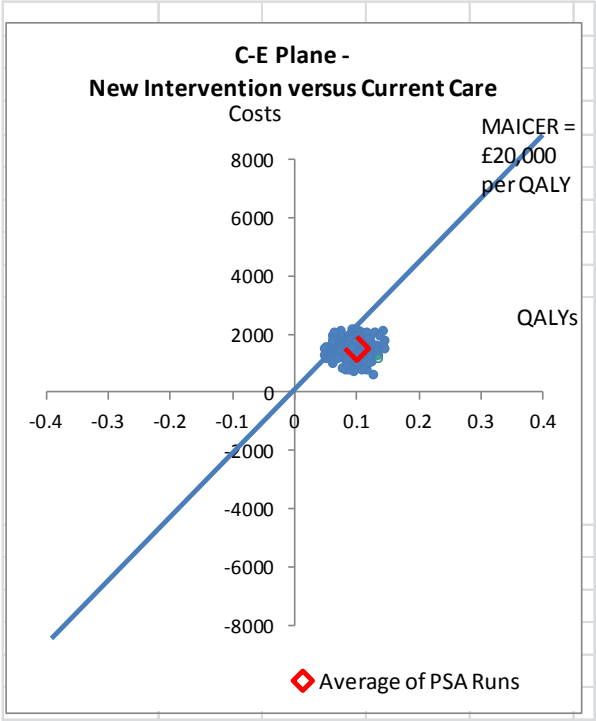
The overall burden of uncertainty also depends on the number of people affected by the decision per annum. If there is only a very small number of people affected then the value of eliminating the uncertainty for say the UK or England as a whole would be lower than if many people are affected.

So why do we call it a burden? Most decision makers would agree that a low PUB is better than a high PUB – because a high PUB indicates a greater risk of the decision based on current evidence and proposed process being wrong; and greater consequences in terms of health or cost savings forgone by not resolving or mitigating against the risk. Others may favour different terminology and this requires further consideration. In a sense, the PUB represents an opportunity to remove uncertainty. If the decision maker could, free of any cost, eliminate all of the uncertainty, then the monetary net benefit measure of the Payer Uncertainty Burden quantifies the expected net gains that the decision maker could achieve.

Figure 10: Four examples of how the two factors (i) probability of decision given current evidence being wrong, and (ii) distance from the MA.ICER if decision is wrong determine the scale of the PUB per person affected by the decision

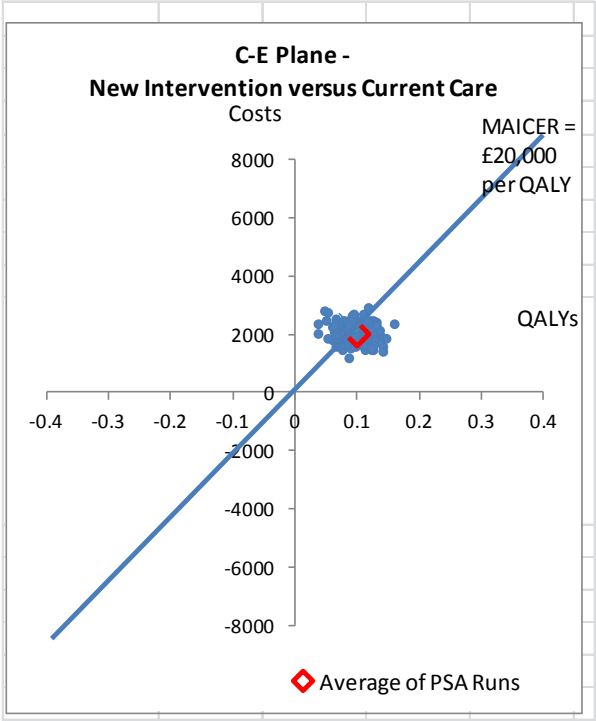
a) Small probability (14%), small distance

Very small PUB - £59



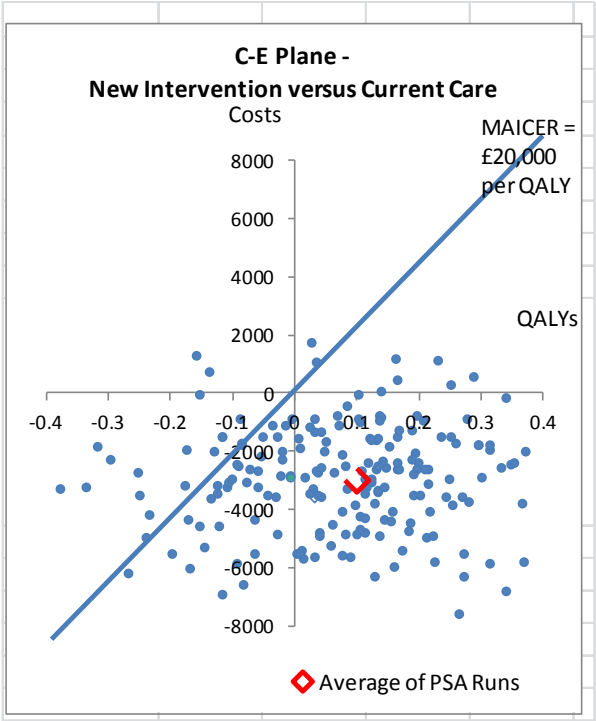
b) Large probability (49%), small distance

Medium PUB – £213



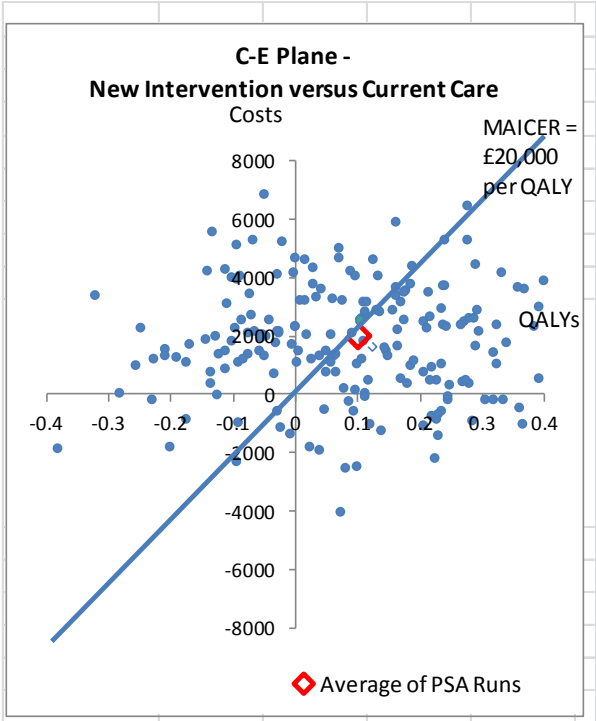
c) Small probability (14%) , large distance

Medium PUB - £227



d) Large probability (49%), large distance

Very large PUB - £1794



Most importantly, the PUB changes if elements of the decision problem or evidence base are changed. The PUB changes if further evidence is collected. If as the decision maker, one could obtain additional evidence, then that would reduce uncertainty in the model input parameters, and this would immediately change the shape of the C-E plane, and as a result the PUB measure for the technology appraisal would change.

The PUB also changes if proposed prices for the interventions are altered. If one of the manufacturers were to alter the proposed or actual prices of the intervention, then of course the incremental cost component of the C-E plane would change, and this would change the proportion of dots on each side of the line and the vertical distance from the line, and as a result the PUB measure for the technology appraisal would change.

In summary for Q3, we can use the standard PSA results from a technology appraisal to quantify the value of eliminating all uncertainty – the PUB – and can measure this in monetary or health terms and compare across different appraisals both in terms of burden per person affected by the decision and burden for the country as whole.

In the next two sub-sections, we *imagine*, that the decision maker (e.g. the NICE Technology Appraisal Committee) recommends one of the interventions that is NOT the 'option we *expect* to be most cost-effective given current evidence and price that is considered in the PSA'. It is important to mention here that a PSA does not always reflect all of the evidence. When we are using the term 'given current evidence and price' in the following, it is assuming that all of the existing evidence is reflected in the PSA. When this is not the case, that is what we call structural uncertainty and we discuss this problem later on in the discussion section. We then ask the last two questions.

2.4.5. Q4. Given current evidence and proposed prices, what is the strategy-specific risk to the NHS, that is the loss in terms of QALYs and costs? This, we call the expected Payer Strategy Burden – PSB

If the decision-maker considers recommending one of the interventions other than the 'option we *expect* to be most cost-effective given current evidence', then we would *expect* to get fewer QALYs or higher costs. Figure 11 shows that, in our hypothetical example, choosing current care rather than the new intervention would be expected to result in fewer QALYs (8.1 versus 8.2 for the new intervention) although for a slightly lower cost. In terms of net

benefit, 7 out of the 10 PSA runs would show a lower net monetary benefit and we would expect (on average over the PSA runs) to obtain £155,000 which is £1,000 lower than the expected net monetary benefit for the new intervention. This, we call the Payer Strategy Burden (PSB). This can also be measured in the net health benefit measure and the PSB for our example would be 0.05 QALYs per person affected by the decision.

Figure 11: Simplified hypothetical model – Payer Strategy Burden Calculation if the Decision Maker were to select "current care" rather than "new intervention" for Q4.

MA.ICER	£20,000						
	QALYs		Costs		Net Benefit = MA.ICER * QALYs - Cost		
PSA Run	New Treatment	Current Care	New Treatment	Current Care	New Treatment	Current Care	Net Monetary Benefit Difference if Choose Suboptimal Intervention
1	8.4	8.1	£ 8,000	£ 7,000	£160,000	£155,000	-£5,000
2	8.4	8.1	£ 8,000	£ 7,000	£160,000	£155,000	-£5,000
3	8.3	8.1	£ 8,000	£ 7,000	£158,000	£155,000	-£3,000
4	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000	-£1,000
5	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000	-£1,000
6	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000	-£1,000
7	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000	-£1,000
8	8.1	8.1	£ 8,000	£ 7,000	£154,000	£155,000	£1,000
9	8	8.1	£ 8,000	£ 7,000	£152,000	£155,000	£3,000
10	8	8.1	£ 8,000	£ 7,000	£152,000	£155,000	£3,000
Average	8.2	8.1	£ 8,000	£ 7,000	£156,000	£155,000	-£1,000
					Highest expected net monetary benefit	Expected to be Suboptimal	Expected Payer Suboptimality Burden is
	Incremental Costs - New v Current			£1,000			PSB = £1,000
	Incremental QALYs - New v Current			0.1			
	ICER - New v Current			£10,000			PSB = 0.05 QALYs

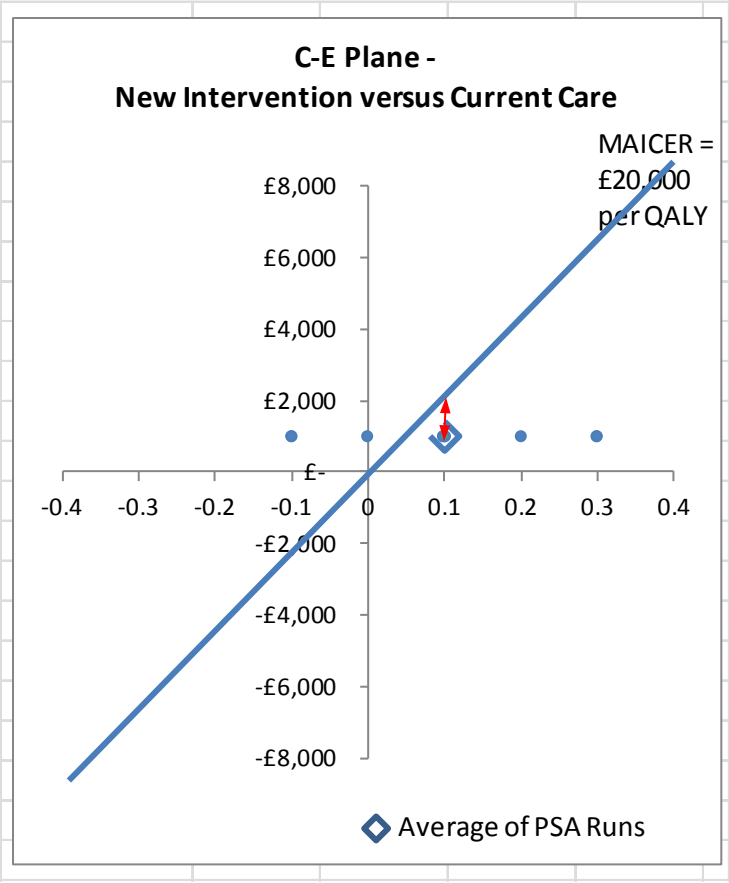
If there are several interventions, then each will have a different Payer Strategy Burden when compared against the option we expect to be most cost-effective.

We can also calculate a Payer Optimality Gain (POG) measure for the option we expect to be most cost-effective versus each of the other options. This is exactly the opposite measure to the PSB, in a sense the other side of the same coin. So, in our example, the Payer Optimality Gain of the new intervention versus current care would be £1,000, because the new

intervention provides an expected £156,000 net monetary benefit compared to £155,000. In the health economic literature, this Payer Optimality Gain also gets termed the incremental net monetary benefit of the most cost-effective option versus other options.

In a two strategy comparison we get a visual instinct for the Payer Strategy Burden by looking at the C-E plane. The vertical distance from MA.ICER line to the centre of the cloud of dots, i.e. the Diamond which is at the point (mean incremental QALYs, mean incremental cost) is exactly the value of the Payer Strategy Burden for current care. Or, to consider the other side of the coin, this vertical distance is the expected Payer Optimality Gain of the new intervention versus current care. Figure 12 shows that for our hypothetical example, the PSB of current care is £1,000.

Figure 12: Simplified hypothetical model – The Vertical Distance of the 'Diamond' central estimate provides a visual understanding of the Payer Strategy Burden of current care for Q4.



2.4.6. Q5. How much would the NHS expect to gain by eliminating the risks associated with both uncertainty and the strategy?

If the decision maker were to switch back to the most cost effective option rather than one of the alternative options and also eliminate all uncertainty, how much health benefit and / or NHS cost saving could we expect the system to gain? Given proposed prices and current evidence, we call this the Payer Strategy and Uncertainty Burden (P-SUB) associated with an option we do not expect to be the most cost effective.

Figure 13: Simplified hypothetical model – Payer Strategy & Uncertainty Burden Calculation - Q5.

MA.ICER	£20,000				
	Net Benefit = QALYs * MA.ICER - Cost			Gain versus optimal option in net monetary benefit if	Gain versus suboptimal option in net monetary benefit if
PSA Run	New Treatment	Current Care	Net Monetary Benefit if all uncertainty were eliminated	all uncertainty were eliminated and decision maker could switch to true optimal strategy	all uncertainty were eliminated and decision maker could switch to true optimal strategy
1	£160,000	£155,000	£160,000	£0	£5,000
2	£160,000	£155,000	£160,000	£0	£5,000
3	£158,000	£155,000	£158,000	£0	£3,000
4	£156,000	£155,000	£156,000	£0	£1,000
5	£156,000	£155,000	£156,000	£0	£1,000
6	£156,000	£155,000	£156,000	£0	£1,000
7	£156,000	£155,000	£156,000	£0	£1,000
8	£154,000	£155,000	£155,000	£1,000	£0
9	£152,000	£155,000	£155,000	£3,000	£0
10	£152,000	£155,000	£155,000	£3,000	£0
Average	£156,000	£155,000	£156,700	£700 PUB	£1,700 PSUB
	Highest expected net monetary benefit	Expected to be Suboptimal PSB =£1,000		Expected (average) gain in monetary net benefit if all uncertainty were eliminated versus optimal	Expected (average) gain in monetary net benefit if all uncertainty were eliminated versus suboptimal

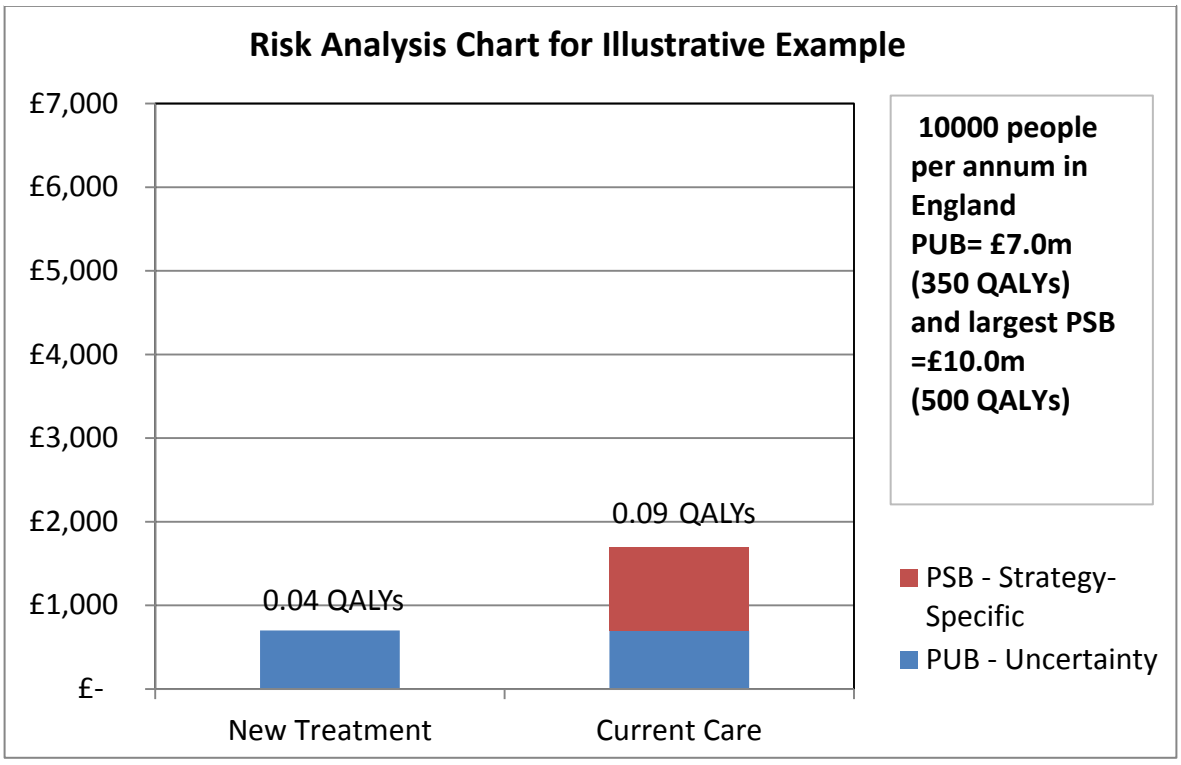
Figure 13 shows how this is calculated for our hypothetical example. If all uncertainty were eliminated, then the decision maker would be able to choose the best strategy in each PSA row, and this is shown in the third column of Figure 13 and with an average net monetary benefit of £156,700. As before, we define the PUB as the difference in monetary net benefit between eliminating all uncertainty (column 3) and the option we expect to be most cost-effective if we make the decision now based on current evidence (column 1). The two blue

arrows compare these figures and lead to the PUB calculation of £700 per person affected by the decision. We now define the Payer Strategy & Uncertainty Burden as the difference in monetary net benefit between eliminating all uncertainty (column 3) and the option we do NOT expect to be most cost-effective if we make the decision now based on current evidence (column 2). The two red arrows in Figure 13 compare these two columns and lead to the P-SUB calculation of £1,700 per person affected by the decision.

It turns out mathematically that PSUB = PUB + PSB. That is, for each strategy given proposed prices and current evidence, the Payer Strategy & Uncertainty Burden is simply obtained by adding together the Payer Uncertainty Burden and the Payer Strategy Burden. In our example the PUB of £700 plus the PSB of £1,000 add together to give a P-SUB of £1,700 for current care. For the new intervention, the PSB is £0 and the P-SUB therefore equals the PUB.

If there are several interventions in a technology appraisal, then we repeat the calculations for Q4 and Q5 for each strategy. There is a different value for the P-SUB for each of the different strategies. We propose presenting the P-SUBs associated with each strategy in our Risk Analysis Chart (Figure 14).

Figure 14: Risk analysis chart for illustrative example



The risk analysis chart in Figure 14 shows the P-SUBs of both, the new treatment and current care. The blue bar represents the PUB, which is the same for both strategies. The red bar represents the PSB, which is zero for the new treatment. We present the QALY value of the P-SUB, so the sum of PUB and PSB, on top of each stacked bar. On the right hand side, we show the PUB and the largest PSB in this appraisal per annum for England which is calculated by multiplying up by the number of people expected to be affected by the technology appraisal decision per year.

In summary, an existing cost-effectiveness model with probabilistic sensitivity analysis results can be used to instantaneously calculate the PUB for the technology appraisal concerned as well as the PSB of deciding to adopt any option which we do not expect to be the most cost-effective. These two measures can be combined to quantify the P-SUB.

Most importantly, the PUB, PSB and the P-SUB will change if those two crucial elements of the decision problem, the proposed price or the evidence available, are changed. A technology appraisal with a large PUB indicates that a Managed Entry Agreement relating to collecting further evidence could be valuable because it could reduce decision uncertainty and enable the decision maker to have greater knowledge for future decisions. A technology appraisal with a large P-SUB indicates that an MEA in which proposed prices change could be valuable.

2.5. THE MEA RISK ANALYSIS FRAMEWORK PART II: ASSESSING THE VALUE OF DIFFERENT SPECIFIC DESIGNS FOR MANAGED ENTRY AGREEMENTS

In this next section we present

- how the measures of PUB, PSB and P-SUB per person are affected by a particular proposed Managed Entry Agreement
- the process for calculating the effect of changes to proposed pricing of interventions on PUB, PSB, and P-SUB per person
- the process for calculating the effect of collecting additional evidence on PUB, PSB, and P-SUB per person
- the process for calculating the effect of a combined price reduction and RwR MEA scheme

- consideration of the number of people affected, the time horizon of any proposed RwR study, and the decision relevance horizon (the number of years which the decision maker considers the decision between the treatment options will be relevant e.g. that time in the future when one expects a completely new class of therapies to overtake the current set of treatment options)

2.5.1. Processes for Calculating the Effects of an MEA on PUB, PSB and P-SUB per person affected by the decision

If a new intervention is not expected to be the most cost-effective given proposed prices and current evidence, we might want to assess the value of a MEA in terms of its reduction of the P-SUB for that intervention.

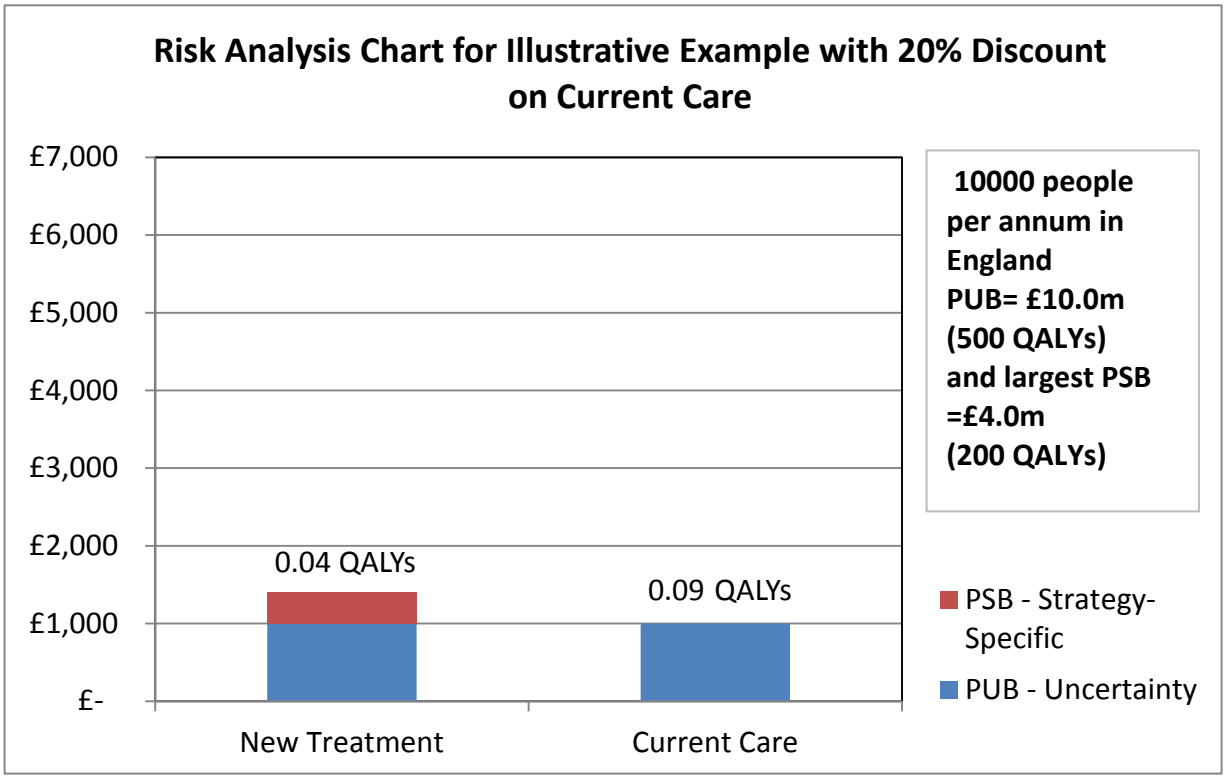
For this, we introduce the residual P-SUB, which is the Payer Strategy & Uncertainty Burden that remains when a MEA is adopted. The residual P-SUB will still have two components – the PUB measuring the remaining uncertainty and the PSB measuring the remaining Payer Strategy Burden if the new intervention is still not expected to be the most cost-effective option even after the MEA scheme. The process for calculating the residual P-SUB is different for price reduction and RwR schemes.

To calculate the P-SUB for price reduction schemes, all that is necessary is to perform another Probability Sensitivity Analysis (PSA) (varying all uncertain parameters in the model and recording results for a large number of iterations) with the new proposed price rule in place. The P-SUB resulting from those revised PSA results will be the residual P-SUB with that price reduction MEA.

Using the results of this revised PSA with the MEA price reduction in place, the decision maker can investigate the effects of the MEA. A revised risk analysis chart with the new price rule in place can then be produced and compared with the original risk analysis chart. An example of a revised risk analysis chart using a 20% discount on the price of current care is presented in Figure 15. The chart shows that current care is now not anymore associated with a PSB and has therefore become the cost-effective strategy. The new treatment, however, is now not cost-effective any longer and is associated with a PSB. It is also

noteworthy that the discount has increased the PUB. That means that we are now less certain about which strategy to choose than we were before the discount.

Figure 15: Revised risk analysis chart for illustrative example with 20% discount on price of current care



If the expected net monetary benefit analysis indicates that the previously cost-ineffective strategy is now expected to be cost-effective, then the decision maker would conclude that the MEA scheme is valuable in that it has removed the Payer Strategy Burden (PSB) component associated with the new intervention. Even if the MEA price reduction means that the new intervention would now be expected to be the most cost-effective and the PSB is now zero, there may of course still be uncertainty as measured by the revised PSA's Payer Uncertainty Burden calculation. So, the decision maker will want to examine whether the residual P-SUB has reduced to a level that seems acceptable. Different price reduction schemes can be simulated, each using a revised PSA, to test how sensitive the measures of the PUB and PSB, and therefore the P-SUB, are to the levels of proposed pricing for the new intervention.

If the expected net monetary benefit analysis indicates that the previously cost-ineffective strategy is still not expected to be cost-effective then the decision maker may still conclude that the MEA scheme has some value in that it may have substantially reduced the P-SUB.

In some appraisals, for some MEA price reduction schemes, the residual P-SUB may still be large due to uncertainties in the evidence that informs the model. In such cases, the decision-maker may want to examine the potential value of MEA schemes which might resolve some of this uncertainty by collecting additional research evidence.

To calculate the P-SUB for RwR schemes, the analytical approach required is slightly more complex (than the simple revision of the PSA model for a price reduction scheme) because it requires specifying the proposed data collection (e.g. what outcome measures, for what sample size, over what follow-up period) and simulating the data which might be obtained. For this analysis it is first of all useful to perform expected value of perfect parameter information *(EVPPI)* analysis on the existing cost-effectiveness model in order to identify the parameters contributing the most decision uncertainty.

Essentially, the EVPPI calculation estimates how much of the Payer Uncertainty Burden is associated with a particular individual or group of the uncertain model parameters. So for example, if the overall decision uncertainty as measured by the PUB were say £700 per person affected by the decision, it might be that £600 worth of uncertainty is associated with the relative efficacy parameters for the new intervention versus current care, and £200 worth of uncertainty is associated with uncertain utility values, and just £5 is associated with uncertain cost parameters. (Note that these partitioned parameter uncertainty EVPPI values do not have to add up to the overall £700 because if there are interactions in the model between the parameters the sum of the partitioned values can be either higher or lower than the overall PUB measure).

The EVPPI gives the decision-maker an indication of which parameters cause the most decision uncertainty and by how much this could potentially be reduced if all of the decision uncertainty surrounding that (group of) parameter(s) could be resolved. EVPPI calculations can be really helpful in deciding which outcome measures might be useful for further evidence collection. For example, if, instead of the figures given earlier, a different technology appraisal had an overall PUB of £700, but the utility parameters had £650 worth

of associated uncertainty, whilst the relative efficacy parameters had just £50 worth, then a large randomised controlled trial to give a much more accurate estimate of efficacy is not likely to make much difference to decision-making, whereas a study to resolve some of the very large uncertainty in utility values would be much more likely to be of benefit to decision-makers in resolving whether the new intervention is truly more cost-effective than current care.

The partitioning into groups of the parameters for the EVPPI calculations can also be important in helping to clarify whether studies assessing multiple outcomes (e.g. a trial which collects both efficacy AND utility data) might be synergistic and more valuable than either a trial or utility study alone. It can also help identify how much of the decision uncertainty is associated with parameters which are outside the realm of trial data collection; for example, the level of uncertainty associated with natural history or disease progression parameters, or safety outcomes only assessable with a long-term horizon post marketing study.

In the past, EVPPI calculations were time-consuming because they usually required a further level of simulation within each PSA sample, but now the EVPPI is easily estimated within seconds straight from the standard PSA results of a model using the freely available "SAVI" online web-application tool (http://savi.shef.ac.uk/SAVI/). The SAVI tool automatically generates EVPPI results for all individual parameters, and can also generate results for any groups of parameters defined by the user, again within just a few seconds.

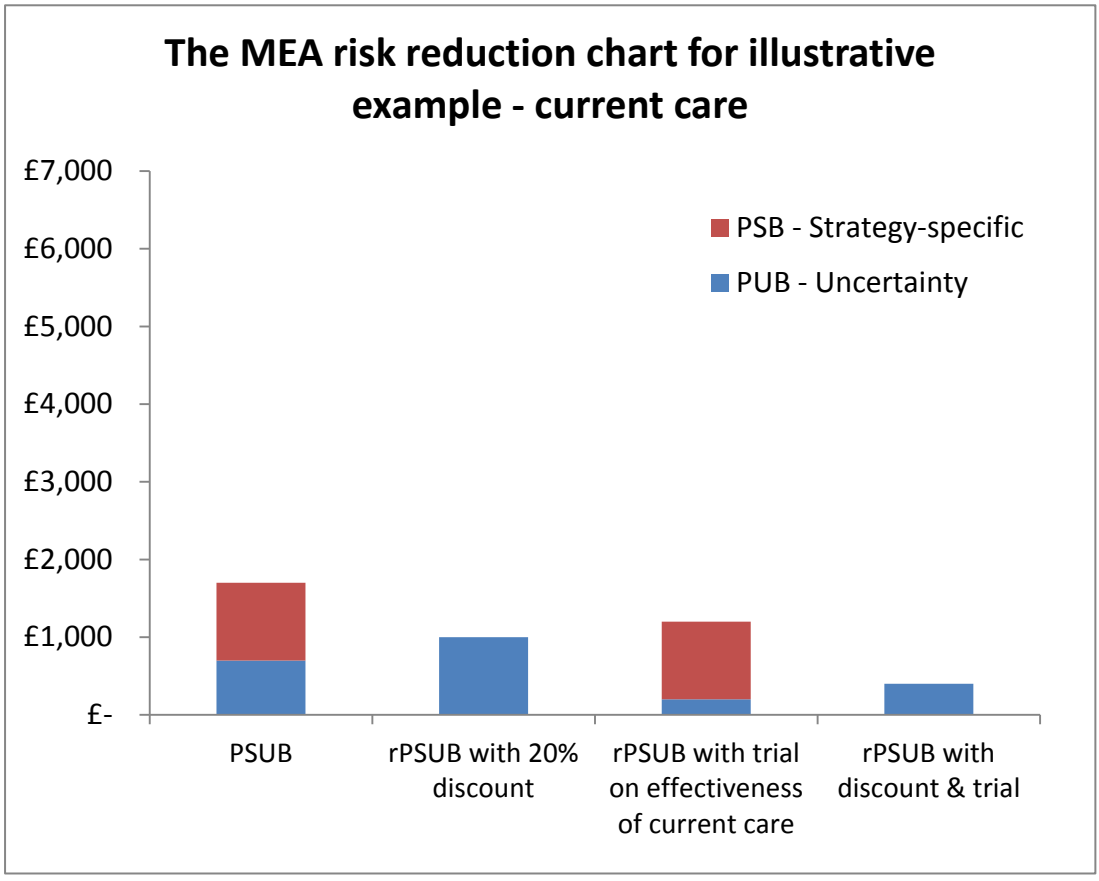
Assessing the value of a possible study or trial design in terms of its potential PUB reduction requires the use of another established concept, the *Expected Value of Sample Information (EVSI).*[18] The EVSI quantifies by how much a specific trial can reduce the PUB. Again, in the past such calculations have been time-consuming but nowadays, the EVSI is easily estimated with SAVI when a simulation of resulting parameter values for the trial or study proposed is available.[19] The simulation of trial results is the only complex component of the calculation processes presented in this report. For each row of the PSA, we need to generate a simulated study result from the proposed study design. Technically, this will involve statistical modelling to specify a likelihood function for the study data conditional on the parameter values in the PSA, and then a simulation process to produce a simulated study result. This results in a simulated study result for each row of the PSA, and then the EVSI is simply calculated in the SAVI tool as the EVPPI of the simulated data. The PUB is then

reduced by the EVSI, yielding the residual PUB. For any interventions within the technology appraisal which are still expected to be cost-ineffective, the P-SUB is composed of the revised PUB and the PSB.

To assess the combination of utilizing both a price reduction and a RwR scheme, the above processes need to be combined. That is, a revised PSA is first undertaken with the price reduction scheme in place, and then the trial simulation and EVSI is conducted. Together this enables the residual P-SUB to be calculated.

To compare the value of different MEA schemes, we propose presenting the residual P-SUBs in the MEA risk reduction chart. An example MEA risk reduction chart showing the P-SUB associated with current care in our worked example, and residual P-SUBs for current care with an example price reduction scheme, an example research scheme and a combination of both is shown in Figure 16.

Figure 16: MEA risk reduction chart for illustrative example – current care



2.5.2. Processes for Calculating the Effects of an MEA on the number of people affected and over time

The question that we pose in this section is: can we trade off the Payer Strategy Burden now (from recommending a technology that is not cost-effective based on current evidence) against data collection later (a RwR agreement that means that we will have a stronger evidence base on which to re-evaluate the decision)? Is this rational?

We can estimate the overall expected strategy-specific burden due to adopting a technology that is not cost-effective (this is the PSB, multiplied by the number of years from the decision to the trial report and decision reassessment). We can also estimate the value of the proposed trial (this is the EVSI of the trial, multiplied by the number of years from the trial report and decision reassessment date up until the time horizon of the new technology). We give a mathematical description of this trade-off in the appendix, where we also include the effect of discounting.

This highlights that the value of research is dependent on the time at which the trial reports as well as the overall time horizon of a technology. The time until research reports is critical because it defines how long we may incur a penalty for having recommended a strategy that was not cost-effective and, together with the time horizon, it defines also the time over which we will benefit from the new knowledge contributed by the trial. The overall time horizon of a technology is influenced by the degree to which there is competition in the system, and the pace of change in the particular sector in which the new technology is situated.

In summary, the steps required to utilise these concepts and methods to make a choice regarding the effects of proposed MEA schemes are:

1. Cost-effectiveness: Assess expected cost-effectiveness and expected net monetary benefits by performing a PSA.
2. Risks associated with decision uncertainty and each strategy: Assess the burdens of decision uncertainty and recommending each strategy (PUB, PSB and P-SUB).
3. Sources of uncertainty: Evaluate uncertainty caused by individual / groups of parameters (EVPPI calculations using SAVI tool).
4. Choice of MEA schemes: Select those MEA schemes that are appropriate and should be assessed further. At this point, it is advisable to consider the issues and guidance questionnaire in Table 3.

5. Assessing MEA schemes: Assess chosen MEA schemes in terms of their reduction of the P-SUB.

3. FEASIBILITY OF MEA RISK ANALYSIS FRAMEWORK DEMONSTRATED IN FOUR TECHNOLOGY APPRAISALS

In this section, we demonstrate the feasibility of assessing risk associated with decision uncertainty and each strategy in most standard technology appraisals using our MEA risk analysis framework. All that is needed for this is the PSA costs, effects and parameter values reported for each simulation. Those results can then be uploaded to SAVI and the PUB and EVPPI analysis is performed at one click. This section also demonstrates how these results alone give an indication of whether and what MEA schemes may be useful.

We selected four past appraisals in which there was some, and in some cases large, decision uncertainty. We provide a short description of the issues at the time of the appraisal, and present the findings for the PUB and P-SUB.

3.1. METASTATIC COLORECTAL CANCER AFTER FIRST-LINE CHEMOTHERAPY

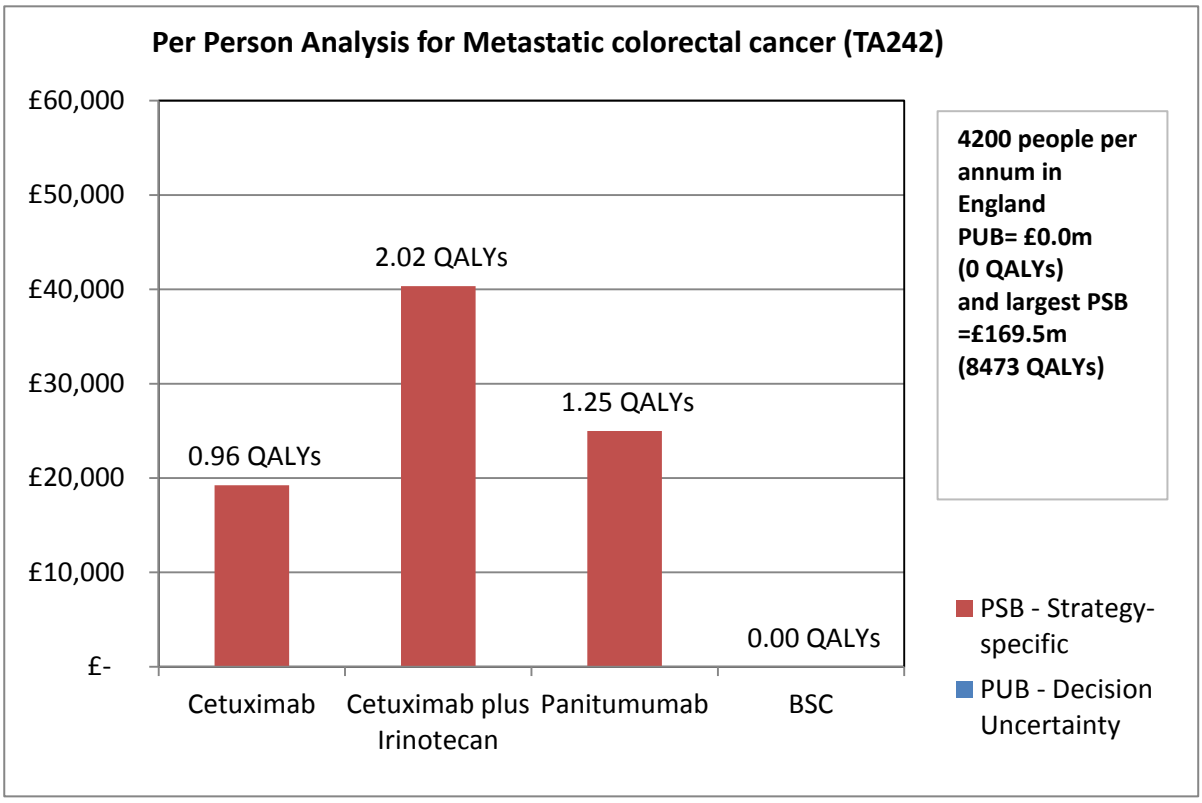
In 2012, NICE appraised cetuximab, a combination of cetuximab and irinotecan, bevacizumab, and panitumumab for the treatment of metastatic colorectal cancer after first-line chemotherapy (TA242).[20] The assessment group had provided a cost-effectiveness model for all the technologies but bevacizumab as they did not think to have sufficient evidence for that. None of the above treatments were recommended.

The committee considered whether end of life criteria were fulfilled by any of the comparators. The first end of life criterion of short life expectancy of less than 24 months was fulfilled for all of them. Bevacizumab did not meet the other criteria of providing an additional three months to patients' lives and of being licensed or indicated for a small patient population. Cetuximab, while fulfilling the criterion of prolonging survival, failed on the small patient population criterion. Panitumumab fulfilled the small patient population criterion and was considered to fulfil the criterion of extending patients' lives, although with a mean incremental survival of between 2.7 and 3.2 months this was uncertain. End of life valuation could therefore be considered for panitumumab only.

The following shows a brief summary of the results from using SAVI on the PSA outputs of the model done by the assessment group. The PUB is zero and the PSB of best supportive care is also zero (Figure 17). This indicates that the best option is best supportive care and

that the decision is not uncertain at all. The PSBs of the interventions are £19,000, £40,000 and £25,000 for cetuximab, cetuximab plus irinotecan and panitumumab, respectively (using an acceptable ICER of £20,000 per QALY for all). The largest PSB results in £170 million or more than 8,400 QALYs.

Figure 17: Risk analysis chart in the metastatic colorectal cancer appraisal



The large PSBs and the fact that there is no decision uncertainty mean that, without the cost of the interventions coming down, it is not possible to recommend these technologies without incurring a major loss. Even if EoL valuation was applied, none of the interventions would be cost-effective, with their ICERs being £90,000, £88,000 and £150,000 per QALY for cetuximab, cetuximab plus irinotecan and panitumumab, respectively.

Cost-effectiveness planes of all interventions against best supportive care (BSC) also show that there is no decision uncertainty. The ICER distributions of all three interventions lie to the top left of the diagonal (Figure 18 to Figure 20). The CEAC shows that up to very high acceptable ICERs, the most likely cost-effective strategy is best supportive care (Figure 21).

Because the interventions are so unlikely to be cost-effective, the partial EVPI analyses return values of zero only. Just to get an idea of the parameters that cause the greatest decision

uncertainty, we changed the maximum acceptable ICER to a high value of £90,000 per QALY. With this, the parameters contributing the most to decision uncertainty are overall and progression-free survival parameters for cetuximab plus irinotecan (53 and 33% of overall PUB respectively) and the utility values associated with all health states (post-progression and progression-free survival) and all comparators (between 44-45% of overall PUB for all utilities). It should be noted that with even higher maximum acceptable ICERs, the survival parameters of panitumumab are likely to contribute more to the PUB.

In conclusion, with the ICER distributions of all the technologies being that far removed from the acceptable ICER, a RwR scheme is not a good strategy. The only possible MEA schemes are thus price reductions or a combination of price reductions and research.

Figure 18: C-E plane for Cetuximab vs BSC in colorectal cancer appraisal (£20k/QALY)

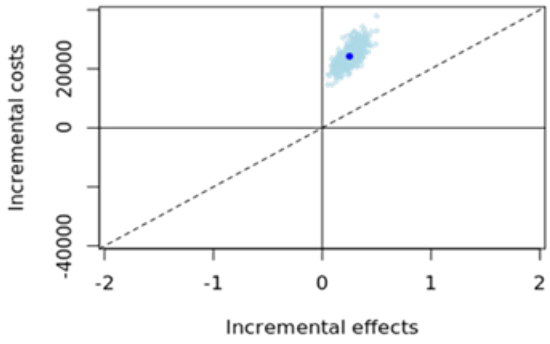


Figure 19: C-E plane for Cetuximab + irinotecan vs BSC in colorectal cancer (£20k/QALY)

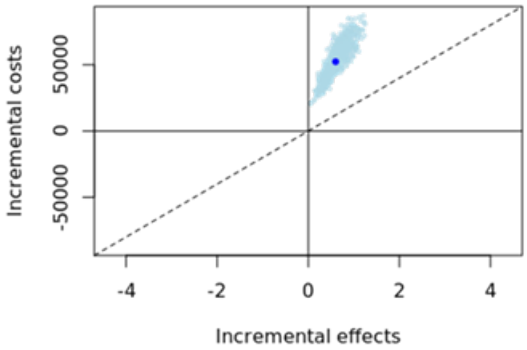


Figure 20: C-E plane for Panitumumab vs BSC in colorectal cancer (£50k/QALY)

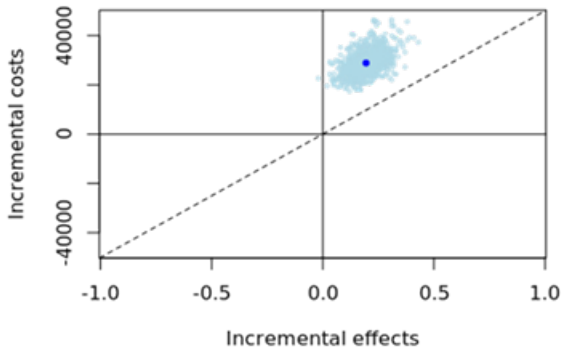
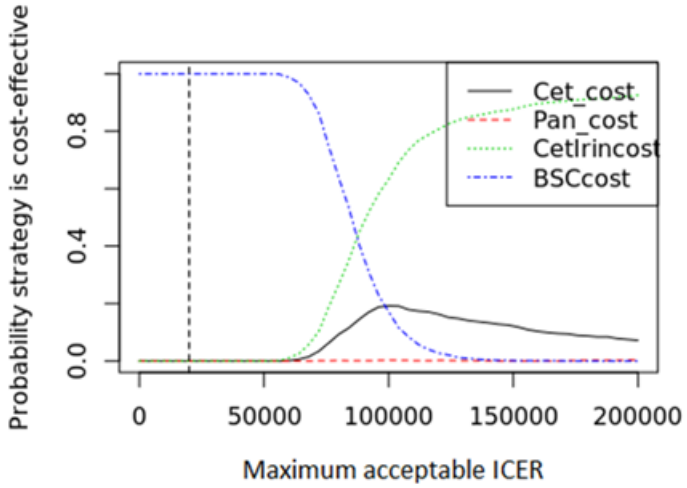


Figure 21: CEAC in colorectal cancer appraisal



3.2. TRABECTEDIN FOR THE TREATMENT OF ADVANCED SOFT TISSUE SARCOMA

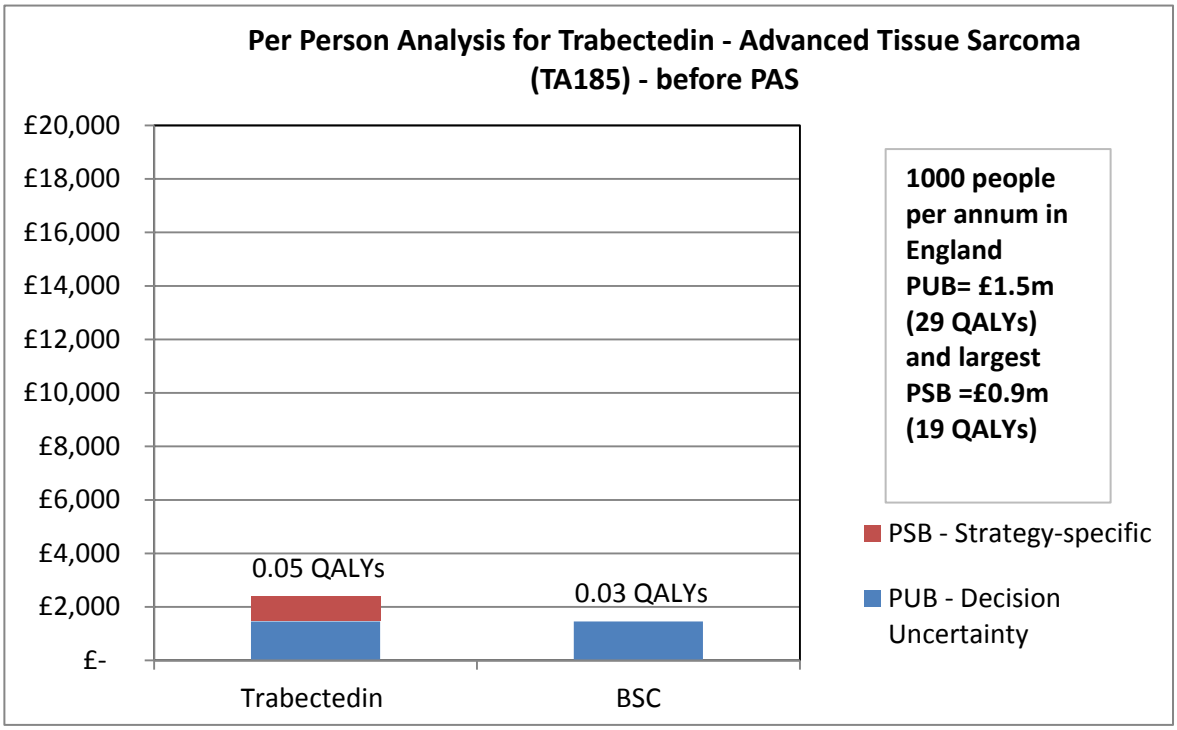
In 2010, NICE appraised trabectedin for the treatment of advanced soft tissue sarcoma (TA185).[21] Trabectedin was recommended if treatment with anthracyclines and ifosfamide has failed or patients are intolerant of or have contraindications for treatment with anthracyclines and ifosfamide. A PAS was also put in place whereby the acquisition cost of trabectedin for treatment needed after the fifth cycle had to be met by the manufacturer, a utilisation cap scheme (scheme E).

The committee concluded that end of life criteria were fulfilled as patients with advanced soft tissue sarcoma have a life expectancy of less than 24 months and trabectedin provides an extension to lives of more than three months. Advanced soft tissue sarcoma affects a very small patient population which meant that all the criteria were fulfilled to place additional weight on the QALYs gained from trabectedin. There is no recommendation of a specific threshold to be employed when EoL criteria hold, but for simplicity, we assume that the adopted maximum acceptable ICER was £50,000 per QALY.

The PUB associated with this decision was £1,457 and the PSB of recommending trabectedin at list price was £1,000 per person (Figure 22). At a population level, this translates into £0.9 million or 19 QALYs worth of PSB, while the PUB would cause losses of £1.5 million or 29 QALYs. This suggests that best supportive care was the best decision option, but with the PUB being larger than the PSB, there was large decision uncertainty. It should be noted that

we have not presented the risk analysis chart on the same scale for all of these appraisals due to the great variability in the P-SUB values.

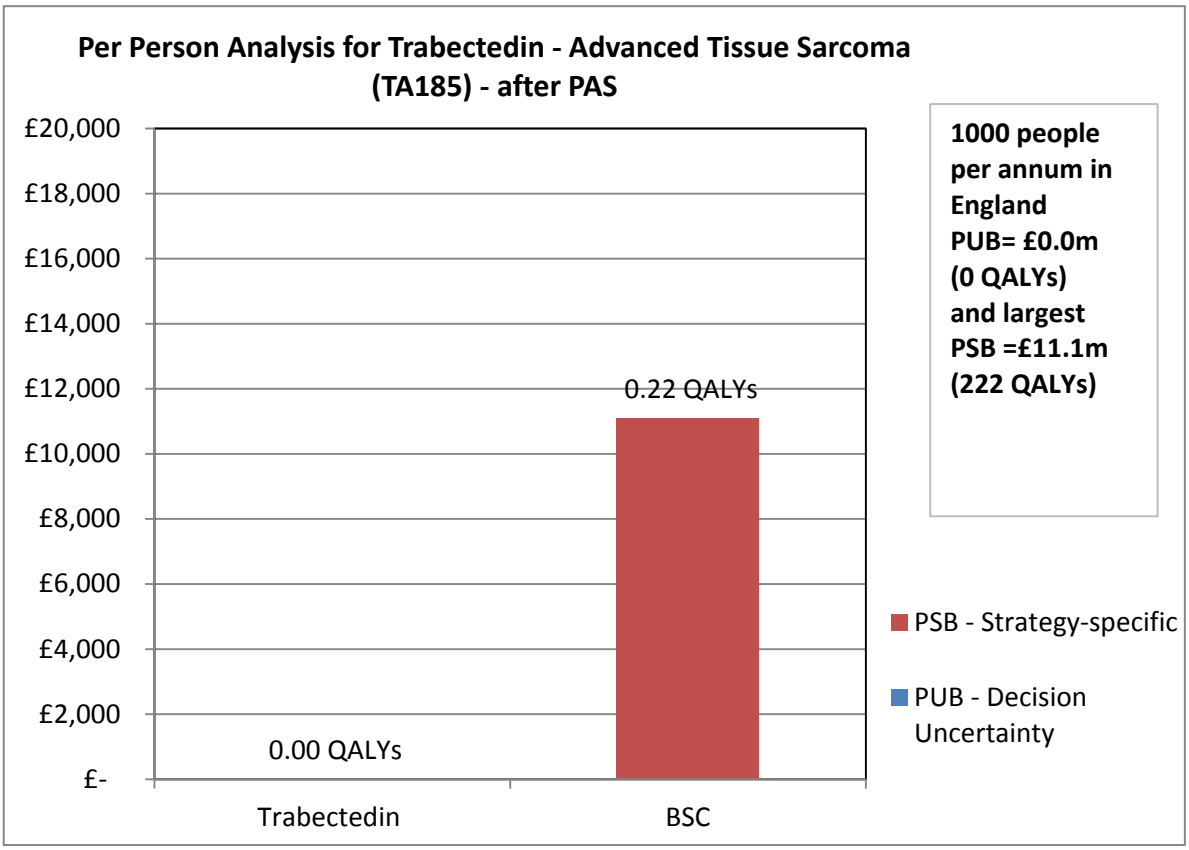
Figure 22: Risk analysis chart in trabectedin for advanced soft tissue sarcoma appraisal at list price



The expected ICER of trabectedin against best supportive care was at £51,772 per QALY when using the list price. If a maximum acceptable ICER of £50,000 per QALY was used, trabectedin would not be cost-effective but it would be close. The cost-effectiveness plane illustrates this and it also highlights the uncertainty in the effectiveness of trabectedin reflected in the horizontal spread to each side of the diagonal (Figure 24). In fact, a good half of the PUB is caused by the survival parameters, which are the main drivers of uncertainty.

At the agreed PAS, the decision reversed and the PUB associated with recommending trabectedin would be £8.5 per person, but the PSB of not recommending trabectedin was much larger at £11,100 per person or £11.1 million at the population level (Figure 23). The cost-effectiveness plane in Figure 25 shows that the ICER distribution has shifted to the right. While there is still large uncertainty in some effectiveness parameters, the decision uncertainty is now small, with only a very small part of the ICER distribution being situated to the top left of the diagonal.

Figure 23: Revised risk analysis chart in trabecetin for advanced soft tissue sarcoma appraisal with scheme E



This appraisal illustrates that when the PSB is relatively small, it may be worthwhile exploring price reduction schemes for their effect on the P-SUB. In this case, the PAS reduced the PUB significantly and eliminated the PSB, creating a situation in which the course of action was clear and trabectedin could be recommended without incurring losses due to decision uncertainty.

Figure 24: C-E plane for trabectedin vs BSC at list price at £50k / QALY

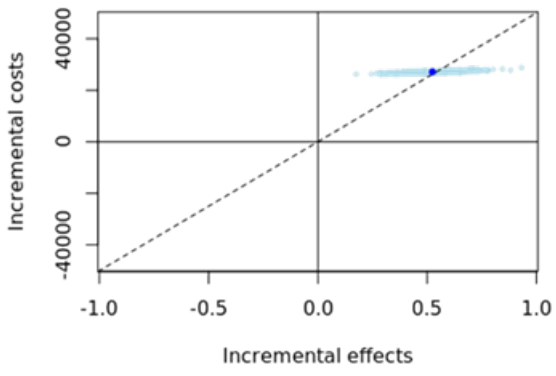
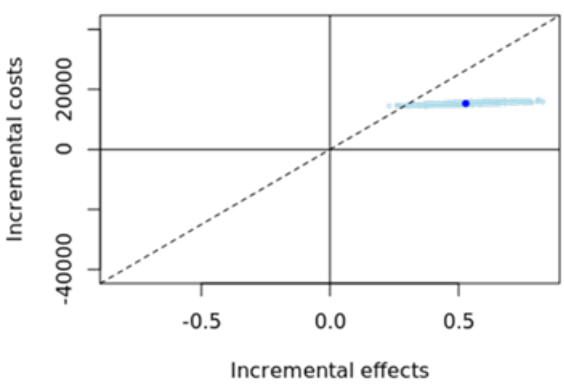


Figure 25: C-E plane for trabectedin vs BSC with PAS at £50k / QALY

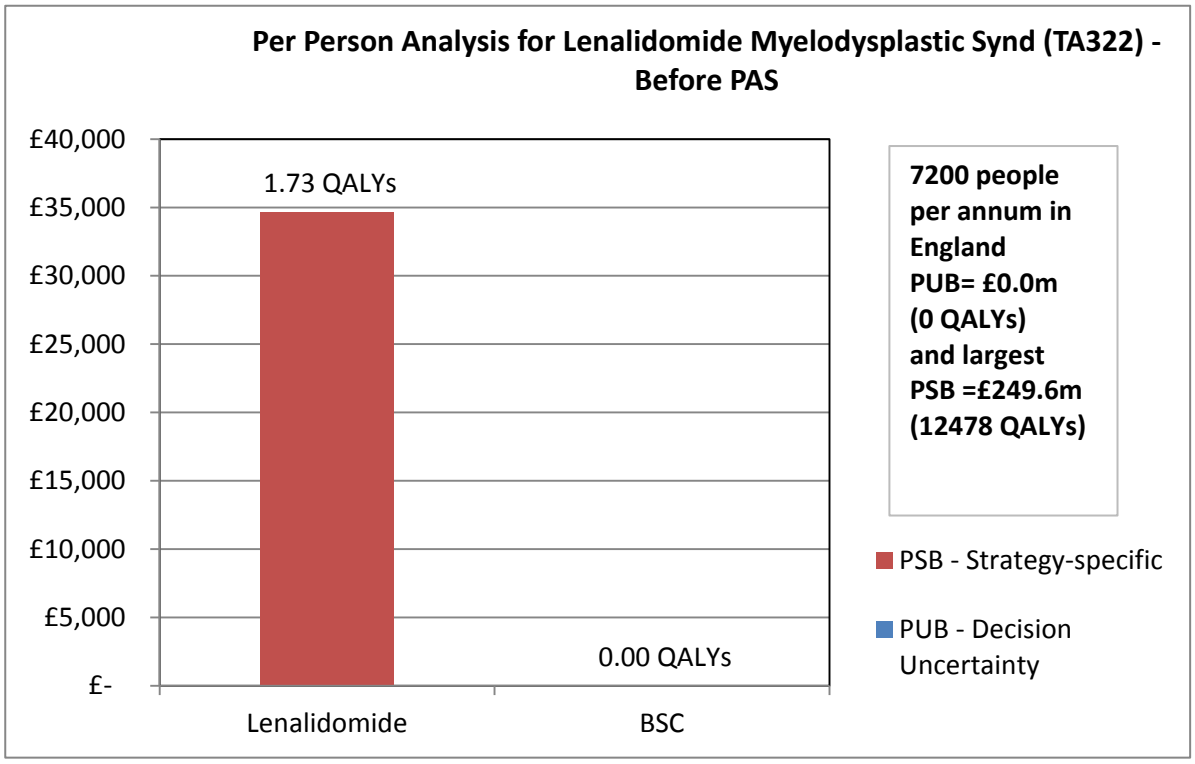


3.3. LENALIDOMIDE FOR TREATING MYELODYSPLASTIC SYNDROMES

Lenalidomide was appraised by NICE in 2014 and recommended as an option for people with transfusion-dependent anaemia, caused by low or intermediate risk myelodysplastic syndromes associated with an isolated deletion 5q cytogenetic abnormality, when other treatments have not worked well enough, conditional on an agreed PAS.[22] The PAS was a utilisation cap scheme (scheme E) by which the manufacturer would not be reimbursed for more than 26 treatment cycles per patient.

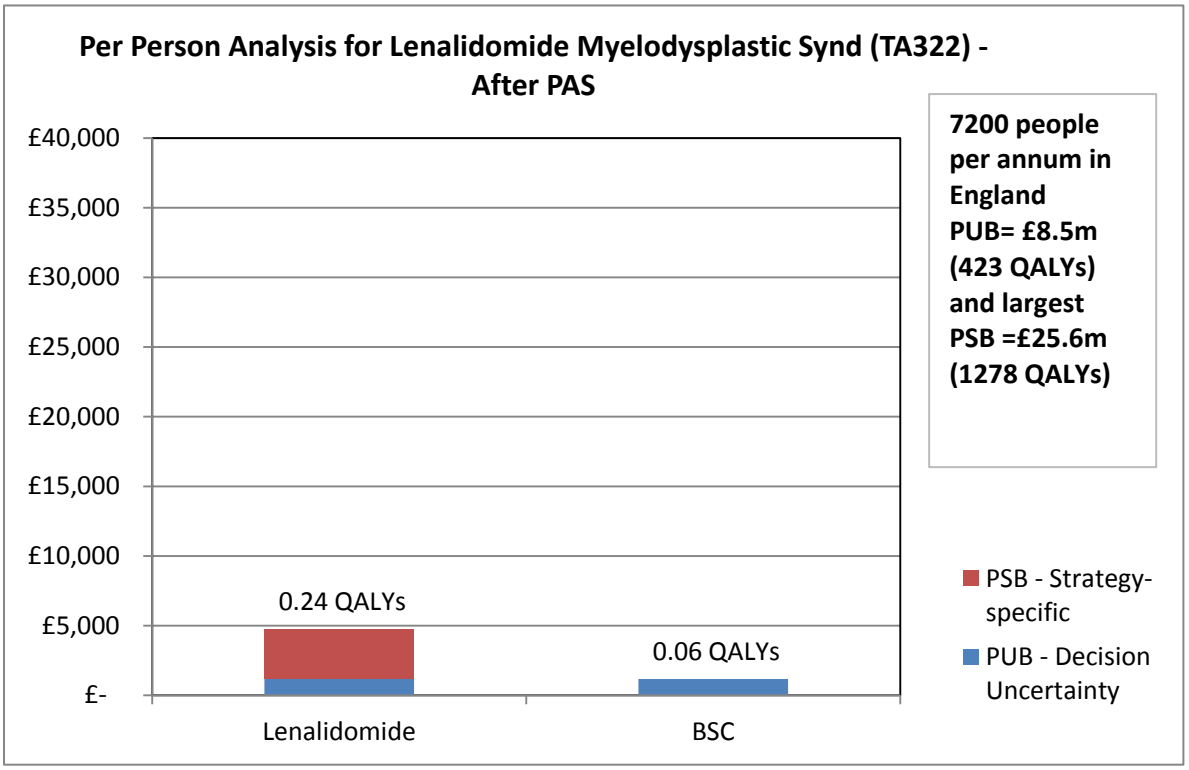
Without the PAS, the analyses were clearly in favour of best supportive care with a PUB of zero and a PSB associated with lenalidomide of £34,660, or £250 million at the population level (Figure 26). The fact that lenalidomide was not cost effective is also reflected in the ICER of £70,129 per QALY. At a maximum acceptable ICER of £20,000 per QALY, lenalidomide should therefore not be cost effective. Figure 28 shows that all of the ICER distribution was located to the left of the diagonal.

Figure 26: Risk analysis chart in lenalidomide for MDS appraisal at list price



The PUB with the PAS was £1,175 per person and £8.5 million at the population level (Figure 27). Best supportive care remained the most cost effective option, with the PSB of recommending lenalidomide still at £4,726 per person, or £25.6 million or 1,280 QALYs at the population level.

Figure 27: Revised risk analysis chart in lenalidomide for MDS appraisal with scheme E



Effectively, the PAS has not made lenalidomide cost-effective, but created decision uncertainty by bringing the ICER closer to the maximum acceptable ICER (the expected ICER with the PAS was £25,135 per QALY (Figure 29)). It has, however, reduced the PSB significantly. It would, nevertheless, be desirable to reduce the price even further and eliminate all of the strategy-specific risk.

Three quarters of the overall PUB were explained by the probability of having acute myeloid leukaemia, the utility associated with that and the response duration associated with lenalidomide. Further research could have been worthwhile to reduce the decision uncertainty.

Figure 28: C-E plane for lenalidomide v s BSC at list price

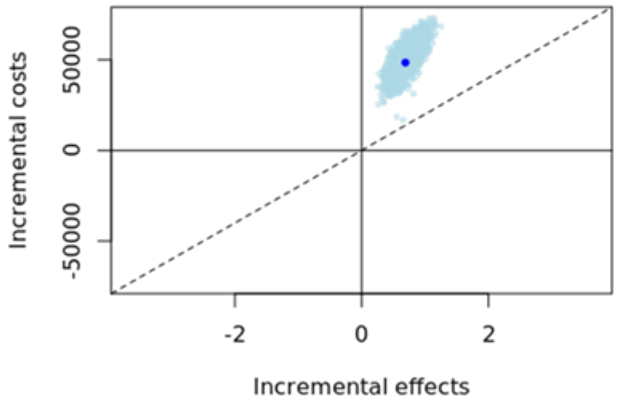
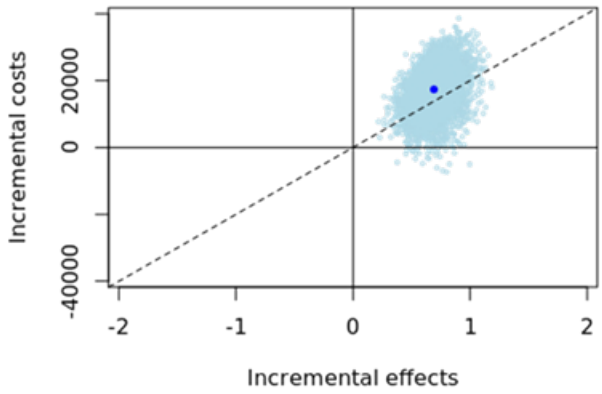


Figure 29: C-E plane for lenalidomide v s BSC with PAS



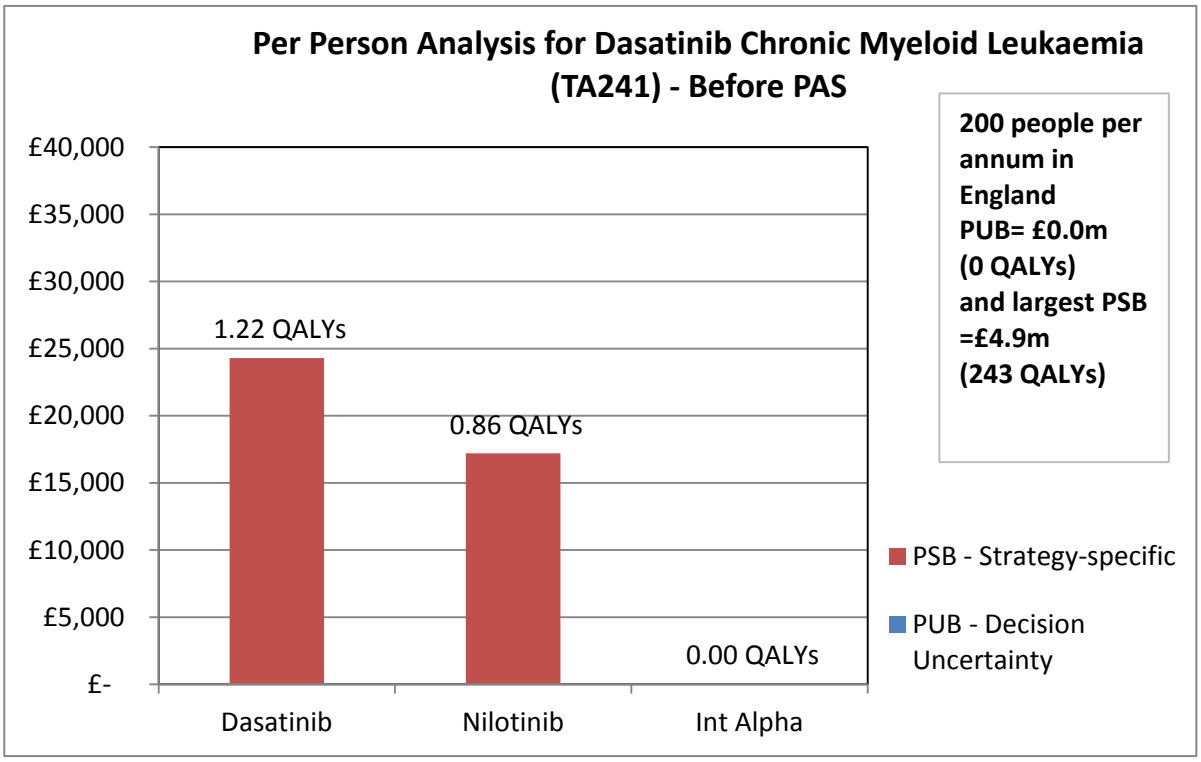
3.4. DASATINIB FOR THE TREATMENT OF IMATINIB RESISTANT CHRONIC MYELOID LEUKAEMIA

In 2012, NICE appraised dasatinib, high-dose imatinib and nilotinib for the treatment of imatinib-resistant chronic myeloid leukaemia (CML) and for the treatment of people with CML who are intolerant to imatinib. An existing comparator was interferon-α. The manufacturer of nilotinib had agreed to a PAS in the shape of a straight discount. NICE recommended nilotinib for both patient groups, but did not recommend dasatinib or high-dose imatinib. For illustration purposes, we focus on the uncertainty associated with dasatinib and nilotinib for treatment of imatinib-resistant CML.

Against interferon-α, both dasatinib and nilotinib had ICERs of approximately £25,000 per QALY. Because of the relative certainty with which interferon-α could be considered cost effective at £20,000 per QALY, the PUB was only £8 per person (Figure 30). The PSB of

recommending dasatinib nevertheless would be £24,300 and the PSB associated with recommending nilotinib £17,200, resulting in a potential strategy-specific burden of £4.9 million or 243 QALYs at the population level.

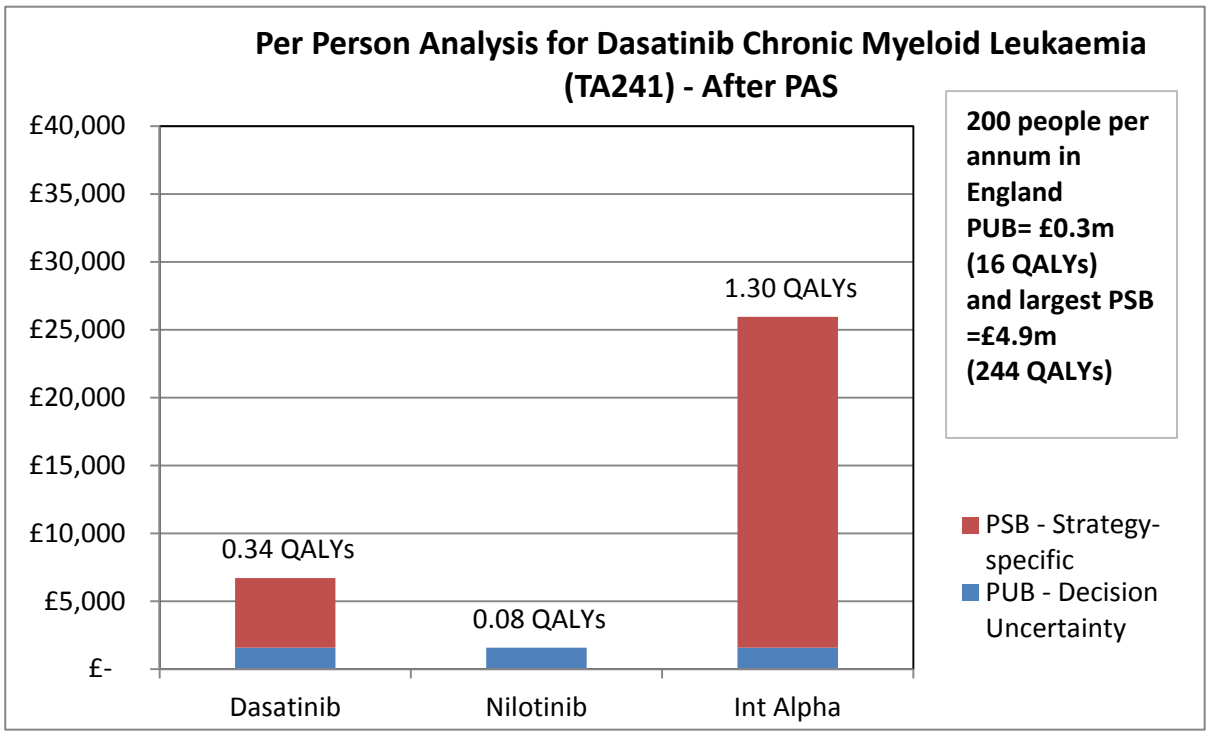
Figure 30: Risk analysis chart in dasatinib and nilotinib for CML appraisal at list price



With the PAS for nilotinib in place, nilotinib was significantly cheaper than dasatinib. At a maximum acceptable ICER of £20,000 per QALY, recommending interferon-α would still have been the optimal decision option, with a probability of being the optimal strategy of 80% (Figure 34).

We therefore considered a maximum acceptable ICER of £30,000 per QALY instead. The PUB would then be large, with £1,500 per person (Figure 31). There would be a strategy-specific burden of £5,000 per patient associated with recommending dasatinib instead of nilotinib, and of £24,000 when interferon-α was recommended instead, driving the potential population PSB up to almost £5 million or 163 QALYs.

Figure 31: Risk analysis chart in dasatinib and nilotinib for CML appraisal at discount



Contributing most to the PUB were the parameters for progression-free survival associated with dasatinib and nilotinib, together making up approximately 75% of the overall PUB. When combined with the parameters for major cytogenetic response (MCyR), that is, response to treatment that can be shown in the marrow, for dasatinib and imatinib, these parameters explain 87% of the PUB. And when combined with the hazard ratio for treatment response (MCyR) in general, this value goes up to 94%.

In conclusion, the recommendation of both dasatinib and nilotinib would ideally be associated with a price reduction scheme for both, as well as a research recommendation on the identified parameters.

Figure 32: C-E plane of nilotinib vs interferon-alpha in CML appraisal at £20k / QALY

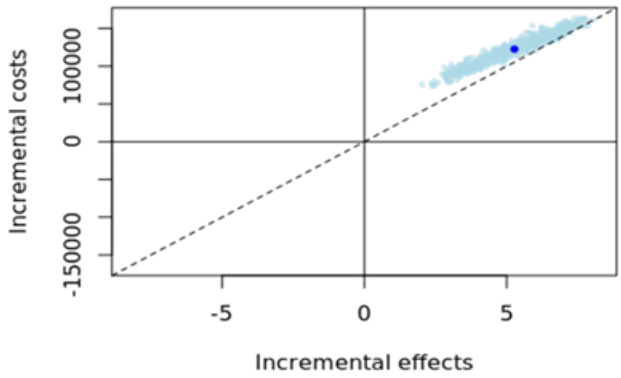


Figure 33: C-E plane for dasatinib against nilotinib in CML appraisal at £30k / QALY

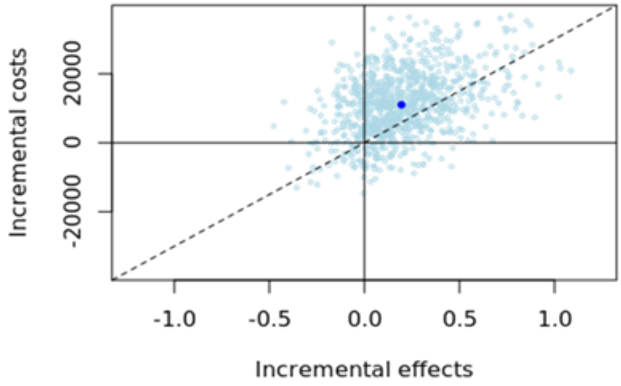
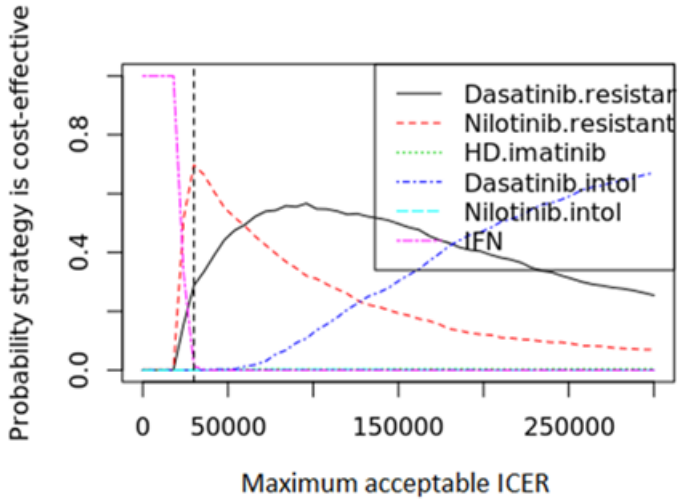


Figure 34: CEAC for CML appraisal (with acceptable ICER of £30k / QALY indicated)



3.5. CONCLUDING REMARKS ON APPLICATION OF THE MEA RISK ANALYSIS FRAMEWORK IN FOUR APPRAISALS

1. This section has demonstrated that the PUB and P-SUB can aid the decision-making process by highlighting whether and what MEA schemes may be appropriate for health technologies.
2. In all of the four appraisals, the SAVI output has provided an idea of whether price reduction schemes or research schemes might be of value.
3. When the PSB is large, only price reductions should be considered.
4. When the PSB is small, price reductions can also be considered before RwR schemes because the price reduction may make a technology cost-effective and reduce the PUB as well.
5. Some price reductions are, however, not sufficient to eliminate the PSB.
6. If there is large decision uncertainty, a small PSB could be acceptable while further research is ongoing but this needs to be carefully examined in further analysis.
7. When the PSA results are available, the PUB and PSB are easily obtained, using SAVI.

4. APPLICATION OF MEA RISK ANALYSIS FRAMEWORK AND IDENTIFICATION OF POTENTIAL MEA SCHEMES IN THREE APPRAISALS

In this section, we use three further past technology appraisals, assess the risk associated with the appraisals and discuss potentially appropriate risk-reducing MEA schemes for each appraisal. For this, we use our MEA risk analysis framework part I (described in Section 2.4) to inform the MEA design guidance questionnaire that was described in Section 2.3.

4.1. TNF-Alpha Inhibitors for Treatment of Ulcerative Colitis

In early 2015, NICE appraised infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (TA326).[23] Infliximab, adalimumab and golimumab have the same marketing authorisation in the UK for the 'treatment of moderately to severely active ulcerative colitis in adult patients who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications for such therapies'.

4.1.1. MEA design guidance questionnaire in the ulcerative colitis appraisal

1. What are the (number and characteristics of) treatment options?

Adalimumab, golimumab and infliximab are monoclonal antibodies that inhibit the pro-inflammatory cytokine, TNF-alpha, called TNF-alpha inhibitors. Alternative treatment options for ulcerative colitis are colectomy and conventional therapy that includes corticosteroids and mercaptopurine or azathioprine.

All the TNF-alpha inhibitors entail induction therapy for the first few weeks a patient is on treatment, in which the dose may differ from that in maintenance therapy. Because of that, cost in induction therapy differs from cost of maintenance therapy. For golimumab and infliximab, the appropriate dose also depends on body weight.

For both induction and maintenance therapy, adalimumab is the cheapest strategy, followed by golimumab, with infliximab being the most costly of the TNF-alpha inhibitors. The cost of adalimumab induction therapy is £2,113 and the cost for maintenance therapy after four

weeks is £704. For infliximab, the cost of induction therapy is at £5,035; the cost of four weeks of infliximab maintenance therapy is £839. Golimumab has a cost of induction therapy of £2,289 and four week maintenance cost of £763. A patient access scheme was agreed with the manufacturer of golimumab that made the higher dose (for patients exceeding 80kg of body weight) available at the same cost as the lower dose.

It should also be noted that different subgroups have been considered in the appraisal: adult and children populations. As this is for illustration purposes only, we focus on the adult population. As different MEAs may apply to both populations, this process would need to be followed for both populations in order to take a decision for both.

2. What is the base-case cost-effectiveness?

The ulcerative colitis model was a state-transition Markov cohort model simulating eight states. When colectomy was an option, adalimumab, golimumab, infliximab and conventional therapy were all dominated by colectomy (Table 5). Colectomy was associated with the largest expected net benefits by far, followed by conventional treatment, adalimumab, golimumab and infliximab at both maximum acceptable ICERs of £20,000 and £30,000 per QALY.

In some patients, colectomy is not an option. In the committee meeting, there was also some doubt as to whether colectomy was indeed a comparator or rather a treatment for the most severe cases. When colectomy was excluded as an option, infliximab was dominated by adalimumab (although the difference in QALYs was small), and golimumab was extendedly dominated by adalimumab and conventional therapy (Table 5).

Table 5: Expected cost-effectiveness and net benefits in ulcerative colitis appraisal (10,000 PSA runs)

	Conventional treatment	**Colectomy**	**Infliximab**	**Adalimumab**	**Golimumab**
Expected Costs (£)	73,624	56,295	96,598	91,225	90,091
Expected QALYs	10.48	14.72	10.82	10.83	10.65
ICER against colectomy	Dominated	-	Dominated	Dominated	Dominated
ICER against conv. treatment	-	NA	67,133	50,353	99,109
Expected net benefits (at £20k/QALY)	136,024	238,103	119,894	125,413	122,880
Expected net benefits (at £30k/QALY)	240,848	385,302	228,140	233,733	229,365

3. What is the nature and scale of risk in this appraisal?

3.a What is the nature and scale of risk captured by the PSA?

According to the assessment group that developed the model, most of the uncertainty came from synthesising data and especially extrapolating short-term trial data to a life-time horizon. There were nine relevant randomised controlled trials (RCT) of which all but one compared either adalimumab, golimumab or infliximab to placebo. The other one was UC-SUCCESS and compared infliximab to azathioprine and to infliximab together with azathioprine. The primary end points in all RCTs were clinical response or remission. Because there were no head-to-head studies, a meta-analysis was performed.

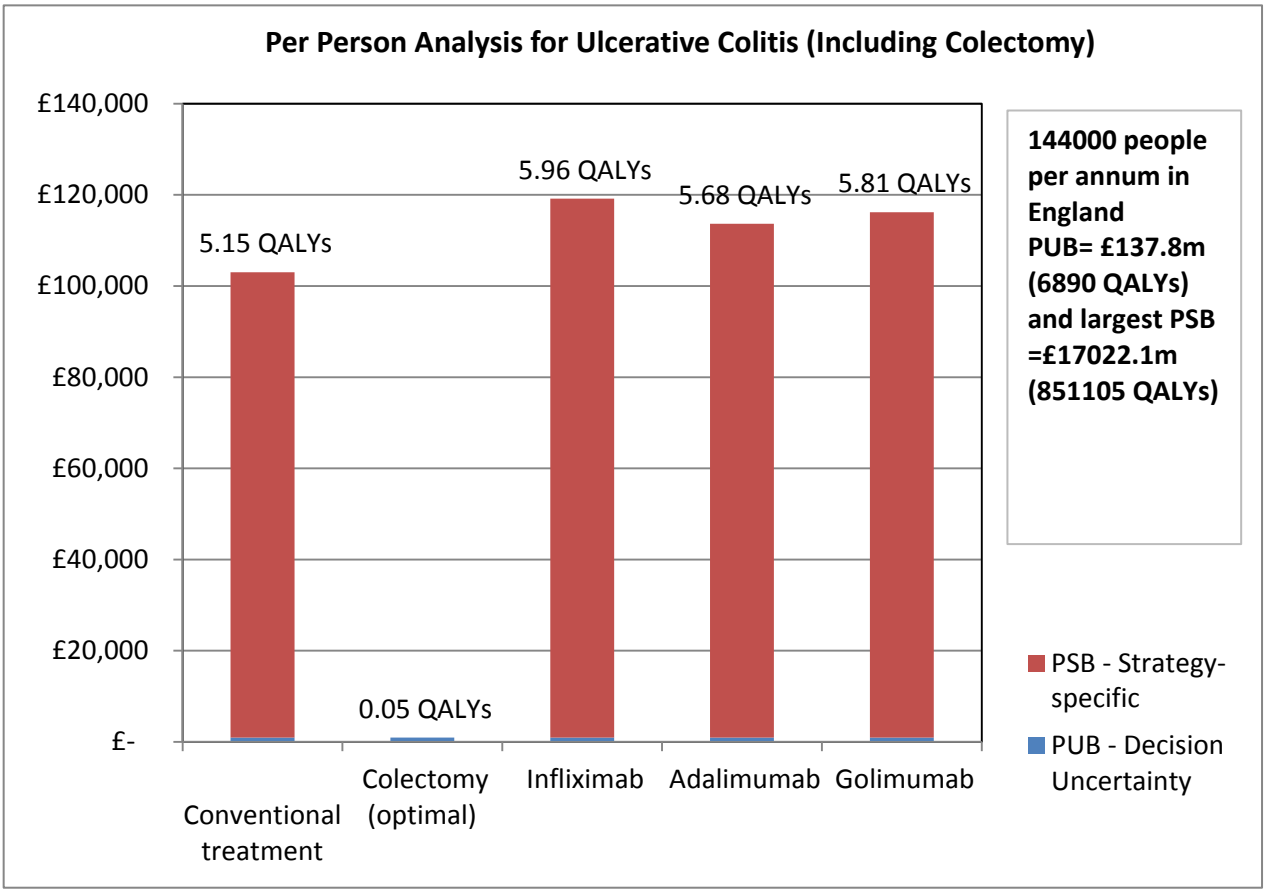
Two trials were excluded from the analysis because the definition of remission differed and most patients had not had azathioprine before. One more trial was excluded from the base-case because patients were only from Japan (this was included in a sensitivity analysis). For the base-case, only the data relating to patients who had not had TNF-alpha inhibitors before was used.

Several sensitivity analyses were conducted, in all of which golimumab remained extendedly dominated and ICERs for adalimumab and infliximab varied slightly but remained in the £50,000 per QALY area. The TNF-alpha inhibitors were dominated in all sensitivity analyses except in one which used an alternative set of utility values. In this scenario, golimumab and

conventional therapy were dominated and the ICER for adalimumab compared with colectomy was £80,315 per QALY.

There was some cost associated with the decision uncertainty, and at a per person level the PUB had a value of £957 (Figure 35). Accrued over the large patient population that is affected per annum, this translates into a PUB of £138 million (Figure 35). When colectomy was an option, the PSB was much larger than the PUB, however, at more than £100,000 per person for all TNF-alpha inhibitors. This translates into a PSB of £17 billion or 851,105 QALYS at the population level. The PSB being so much larger than the PUB suggests that while there is some value in resolving decision uncertainty, but without a reduction in price, recommending the TNF-alpha inhibitors will result in a large cost to the NHS.

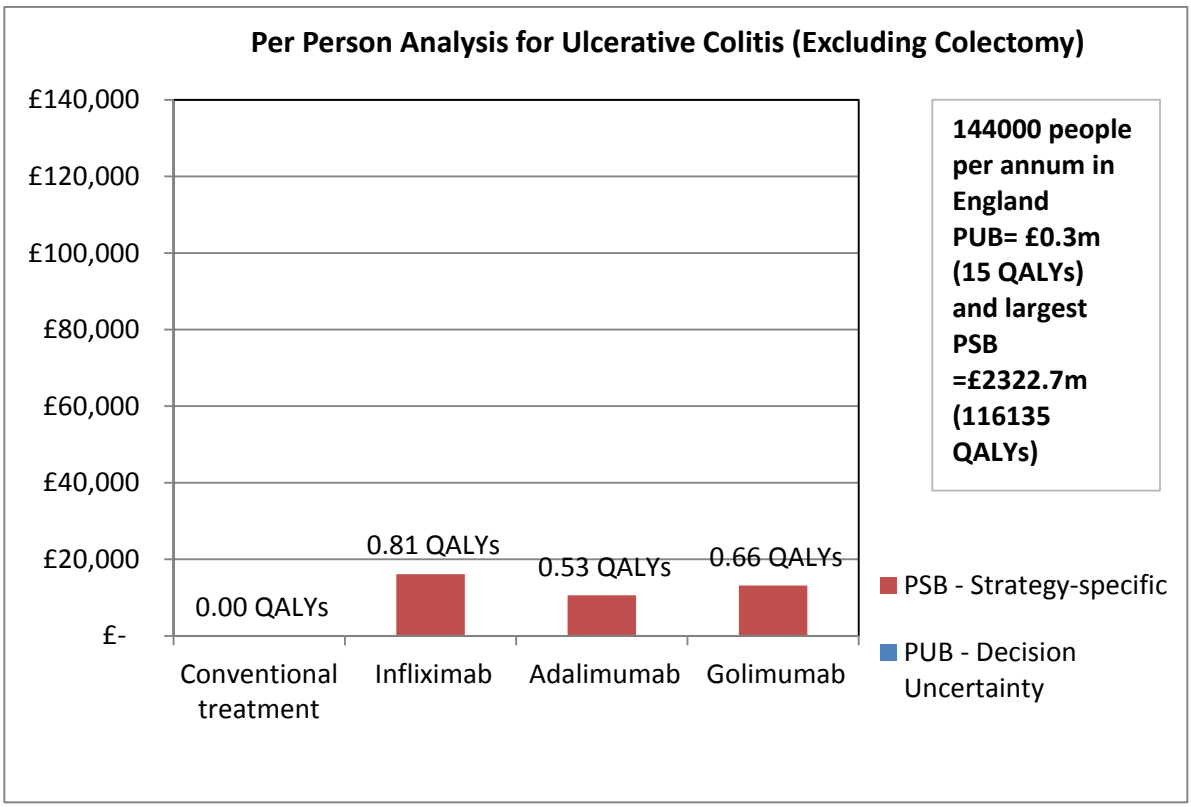
Figure 35: Risk analysis chart in the ulcerative colitis appraisal with colectomy included



When colectomy was excluded from the analysis, conventional treatment was the optimal strategy and the PUB at only £2 per person (Figure 36). Accrued over the annual population,

this translates into £0.3 million or 15 QALYs. The PSB associated with the TNF-alpha inhibitors was then between £10,000 and £16,000 (Figure 36), which translates into £2.3 billion of 116,000 QALYs for infliximab. The most appropriate MEA scheme thus appears to be a price reduction. If such a reduction in price were implemented, the PUB would be re-assessed and further research might also become an important consideration.

Figure 36: Risk analysis chart in ulcerative colitis appraisal when colectomy is excluded



To illustrate this in a different way, the cost-effectiveness plane of adalimumab against colectomy shows how far the expected ICER of adalimumab was removed from the maximum acceptable ICER diagonal (Figure 38). The wide spread into the north-west quadrant and most of the ICER distribution lying above the diagonal explain the large PSB compared with the PUB of all the TNF-alpha inhibitors.

When colectomy was not an option, the cost-effectiveness plane of adalimumab against conventional therapy at £20,000 per QALY shows that the largest part of the ICER distribution lies above the diagonal and there is therefore a very limited chance of it being cost-effective (Figure 37).

Figure 37: C-E plane of adalimumab vs conventional therapy in ulcerative colitis appraisal

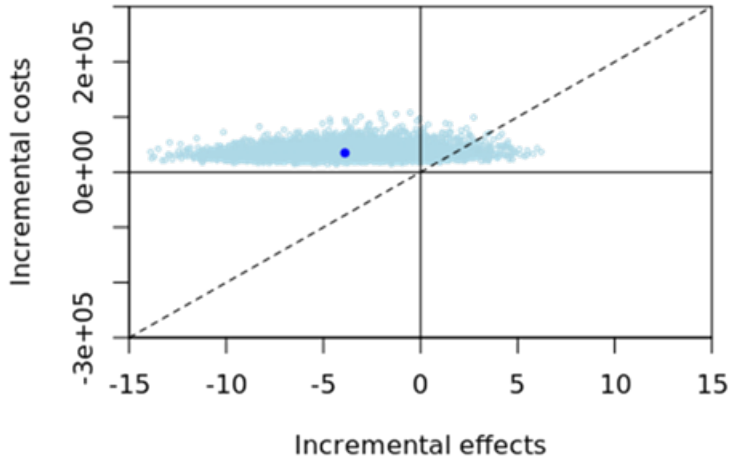
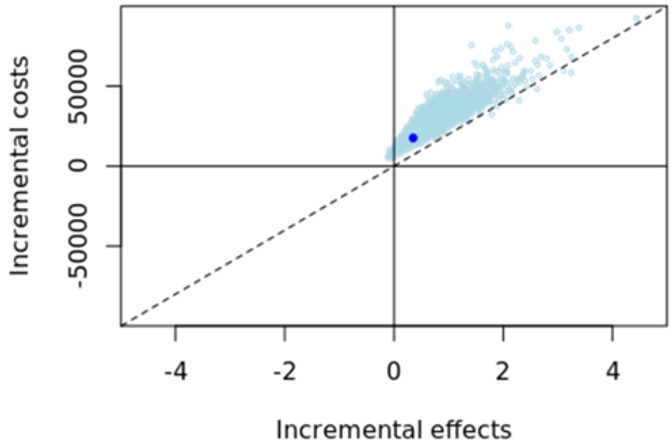


Figure 38: C-E plane of adalimumab vs colectomy in ulcerative colitis appraisal



3.b What is the nature of uncertainty not captured by the PSA?

As mentioned above, a few sensitivity analyses were conducted that are not reflected in the results of the PSA: 1. The Japanese study was included, 2. data relating to patients who had TNF-alpha inhibitors before was included, 3. both were included. Furthermore, an alternative set of utility values sourced from a different study was also explored in a sensitivity analysis.

The Committee recommended all the TNF-alpha inhibitors, because although the ICERs were higher than would be considered cost-effective it was agreed that they were largely over-estimated. This was due to uncertainties in the model, namely that the utility decrement associated with the post-surgery state was under-estimated, the rate and cost associated with surgery was under-estimated and that costs and utility decrements associated with adverse effects of corticosteroids were excluded. In conclusion, this is an example of the PSA not

capturing all uncertainties. The results of the MEA risk analysis framework would therefore have to be taken with caution.

3.c What is the temporal nature of uncertainty, e.g. is there more uncertainty beyond the trial period or is it resolvable with open-label follow up?

The uncertainties surrounding the long-term benefits of TNF-alpha inhibitors beyond the trial duration as well as the optimal duration of treatment with TNF-alpha inhibitors were deemed to be an issue. Furthermore, the uncertainty surrounding long-term relapse and response rates at re-treatment if treatment was stopped was thought to affect an alternative strategy that entailed stopping treatment with a TNF-alpha inhibitor after a certain time (see Question 5.).

4. What is the uncertainty by groups of parameters?

No single parameters contributed the most to the overall PUB – but one group of them did. All utility values together, that is the utility for the state of remission, the dis-utilities associated with a loss of remission and a loss of response, the post-surgery utility and the utility decrement of chronic pouchitis together accounted for approximately 100% of the overall PUB.

5. What alternative treatment strategies might be available?

It was thought that a stopping rule similar in design to the one used in the appraisal for TNF-alpha inhibitors for treating Crohn's disease could be an alternative licensed treatment strategy. This would entail stopping treatment at 12 months or at treatment failure (which would require surgery) and only continue after if there is clear evidence of ongoing active disease. For this, patients should have their disease re-assessed at least every 12 months and, on relapse, should be given the option of starting treatment again. This treatment strategy could potentially improve cost-effectiveness of TNF-alpha inhibitors but it was not modelled as there was no evidence on patients' health outcomes when treatment was stopped, patients' rates of relapse or the effectiveness of re-treatment being the same as the effectiveness on treatment the first time.

6. What measures of patient-based outcomes are available and measurable?

The severity of ulcerative colitis can be assessed using the Mayo scoring system for assessment of ulcerative colitis activity with scores ranging from 0-12 and classifying the states of no disease (0-1), mild disease (2-5) and moderate to severe disease (6-12).

7. Is price a substantial part of overall costs associated with treatment?

Price was a substantial part of overall costs.

8. Are there any precedent PASs in place?

There was a PAS for golimumab which made the higher dose available at the same price as the lower dose. It was included in the base-case and did not make golimumab cost-effective.

9. Could price agreements be national or local?

It was stated in the guidance that costs of treatments may vary because there were negotiated procurement discounts but this is stated in all guidance. Other than that there was no further information on price agreements.

4.1.2. Possible MEAs in the ulcerative colitis appraisal

Starting with effective price reduction schemes, we briefly discuss the suitability of the available MEA schemes in Table 6 and Table 7.

Table 6: Verdicts on the usefulness of different price reduction schemes in the ulcerative colitis appraisal

Price MEA scheme	**Possible use**	**Verdict**
A. Money back guarantee	For patients who have no response, the manufacturer could provide a refund.	Needs including in the model and testing for cost-effectiveness.
B. Conditional treatment continuation	A combination of this with scheme E has been done. The patient utilitisation cap meant that patients would not continue treatment for longer than one year but furthermore treatment continuation was conditional on treatment success: if treatment had failed before the end of the year, TNF-alpha inhibitors would have been withdrawn sooner.	This MEA scheme lacked the evidence to support it and needed more research done on parameters such as health outcomes in patients whose treatment is stopped, their rates of relapse and whether the effectiveness of re-treatment would be the same as the effectiveness of treatment for the first time.
C. Price linked to outcome	Since there are defined outcome measures in the shape of the Mayo score, the price of treatment after a certain time point (for instance, after induction therapy) could be linked with the health states mild, moderate or remission.	Needs including in the model and testing for cost-effectiveness.
D. Discounted treatment initiation	This could be realised in the shape of the manufacturer providing treatment for free or at a reduced price for the duration of a fixed period, such as a year, such that it would become cost-effective for those patients who only have the first year of treatment.	Needs including in the model and testing for cost-effectiveness.
E. Utilisation cap	See scheme B.	See scheme B.

F. Fixed cost per patient	This could ensure that no over-spend would be made for patients who are treated for long times. Since ulcerative colitis is a life-long condition, this may not be realistic.	Probably not applicable.
G. Discount	With the cost of TNF-alpha inhibitors being very high leading to them not being cost-effective, a discount could be a good option.	Needs including in the model and testing for cost-effectiveness.
H. Expenditure cap	There is no clear coherent rationale for this.	Not applicable.
I. Price volume agreement	There is no clear coherent rationale for this.	Not applicable.

Table 7: Verdicts on the usefulness of different research schemes in the ulcerative colitis appraisal

Research MEA scheme	**Possible use**	**Verdict**
4. RwR with reimbursement only	There are some model parameters that would benefit from research. The main uncertain parameters were the utility values used in the model. For instance, a post-surgical utility study would be highly useful to clarify whether colectomy is as cost-effective as it looked in the model results. Other possible study designs relate to the possibility of the alternative treatment strategy of stopping treatment after 12 months in responders. One study design could entail giving different TNF-alpha inhibitors to patients in different trial arms for one year and following these patients up long-term. This	Uncertainty associated with the utility values could be reduced. To take account of the stopping rule in the cost-effectiveness analyses a study on the effectiveness on this treatment regimen would be essential to produce evidence on many of the associated parameters. However, given the large PSB, research cannot stand alone but has to be accompanied by a price reduction.

	study could also be extended to include an additional arm following up on patients after surgery. With both of these studies it would be possible to elicit utility values using an EQ-5D questionnaire. An observational study could also be possible, in which patients treated with different TNF-alpha inhibitors could be followed up long-term. Relapse rates and re-treatment response rates as well as EQ-5D data on utilities could be obtained from such a study.	
5. RwR with reimbursement and refund agreement	This could limit the cost to the payer through implementing a rule whereby a refund will be provided for patients who do not reach remission after a certain time period of treatment.	See 4, but more desirable from the payer's perspective.
6. RwR with a conditional flexible pricing agreement	A possible pricing agreement could involve interim analyses at certain time points at which the health outcomes (for instance, number of patients in remission) are assessed. The price may then be adapted for the ICER to fall in an agreed acceptable range.	Could be used but similarly to 4 and 5 should not stand alone without a price reduction.

4.1.3. Conclusion on ulcerative colitis appraisal

In conclusion,

1. The chosen MEA scheme B and E of conditional treatment continuation and a utilisation cap lacked the evidence to support it and may not reduce the P-SUB.
2. Some MEA schemes can be viewed as alternative treatment strategies on which research needs to be conducted.

3. This case study also highlighted how much larger the risk can be for the payer when a strategy is recommended that is not cost-effective: recommending adalimumab for instance would be associated with a P-SUB that is a hundred times larger than not recommending it and instead maintaining colectomy as the preferred treatment option.
4. This appraisal is an example of large uncertainty not captured in the PSA which means that the MEA risk analysis framework results have to be interpreted with caution.

4.2. ADJUVANT IMATINIB FOR GIST

In 2010, NICE appraised a treatment regimen of three years of imatinib for the adjuvant treatment of gastrointestinal stromal tumours (GIST) after surgery (TA196). At that time, imatinib was not recommended by NICE despite ICERs that were close to a maximum acceptable ICER of £20,000 per QALY because the evidence base of its effectiveness was considered too immature to enable conclusions to be drawn about key aspects of imatinib's clinical effectiveness. People at high risk of recurrence were however eligible to receive adjuvant imatinib for up to three years via the Cancer Drugs Fund.

The appraisal was reviewed in 2014 in TA326 after the manufacturer submitted new evidence.[24] This new evidence resulted in the ICER falling below £20,000 per QALY. Imatinib was then recommended as an option as adjuvant treatment for up to three years for adults who are at risk of relapse after surgery for KIT-positive gastrointestinal stromal tumours.

We therefore investigate whether closer examination of the uncertainties at the time of the 2010 appraisal could have resulted in a different decision where adjuvant imatinib would have been recommended with a MEA scheme in place.

4.2.1. MEA design guidance questionnaire in the GIST appraisal

1. What are the (number and characteristics of) treatment options?

Imatinib is a selective kinase inhibitor which binds to activated c-KIT receptors and blocks the cell signalling pathway, preventing uncontrolled cell proliferation. For the treatment of GIST, imatinib was considered to be prescribed for one year or three years after surgery in

TA326 but only the three year treatment strategy was considered in TA196. In both cases, the comparator was no adjuvant treatment after surgery. Imatinib is available in doses of 100 mg (60-tab pack) and 400 mg (30-tab pack) at net prices per pack of £862.19 and £1,724.39 respectively. At a dose of 400 mg per day, drug costs for a course of treatment would be approximately £20,700 for one year and £62,100 for three years.

2. What is the base-case cost-effectiveness?

The manufacturer's model used a Markov state transition approach in which patients could remain recurrence-free, have a recurrent GIST (first or second recurrence), have progressive disease or die. The analysis included patients at high risk of recurrence. The 2010 base-case ICER for adjuvant imatinib treatment for 3 years against no treatment was £20,264 per QALY, with the expected costs and effects of imatinib both exceeding those of no treatment (Table 8). It should be mentioned that the 2014 base-case ICER for one year of adjuvant imatinib treatment vs no adjuvant treatment was £3,509 per QALY, and the ICER for three years adjuvant imatinib vs one year adjuvant imatinib was £16,006 per QALY.

Table 8: Cost-effectiveness data in GIST appraisal (2010) (1,000 PSA runs)

	Imatinib 3 yrs	**No treatment**
Expected Costs (£)	87,155	46,662
Expected QALYs	6.33	4.33
ICER Imat vs. No tx	20,264	-
Expected net benefits (at £20k/QALY)	39,434	39,960

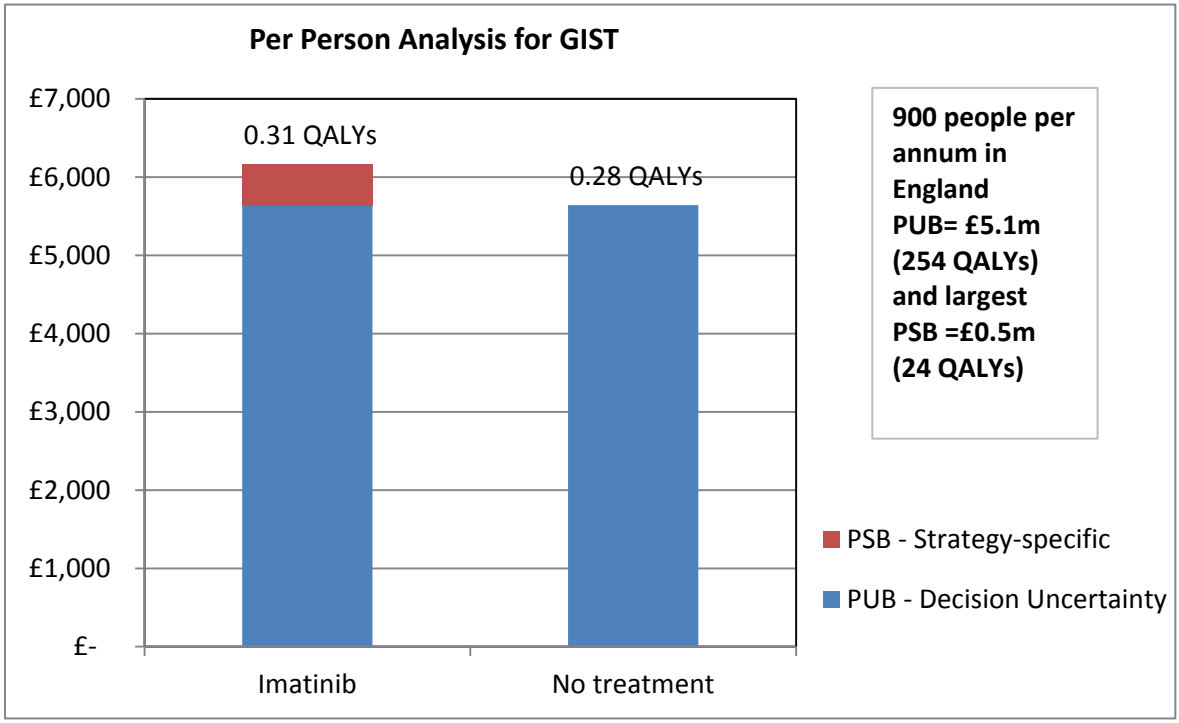
3. What is the nature and scale of risk in this appraisal?

3.a What is the nature and scale of risk captured by the PSA?

In the GIST appraisal (2010), there was large decision uncertainty compared with the strategy-specific risk associated with recommending a course of three years imatinib post-surgery. Over the affected population, the PUB could be as large as £5.1 million or 254 QALYs, whereas the PSB associated with not recommending imatinib was at £0,5 million or 24 QALYs (Figure 39). This suggests that a research recommendation may have been an appropriate course of action, but a price reduction may have been even more effective. This is

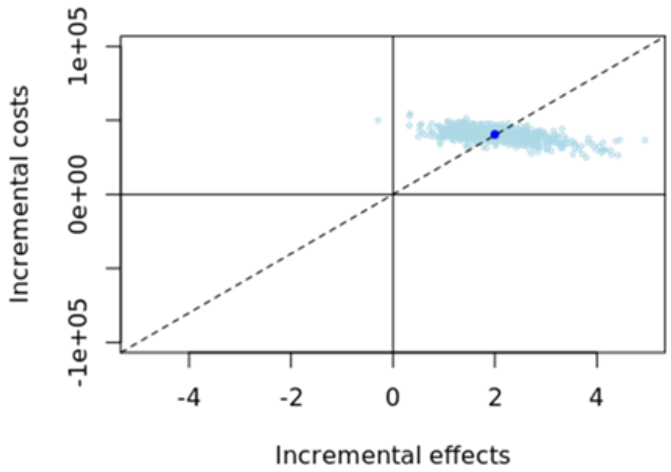
because reducing the price could not only have reduced the PSB but also the PUB, by bringing the ICER down and further away from the threshold.

Figure 39: Risk analysis chart in the GIST appraisal (2010)



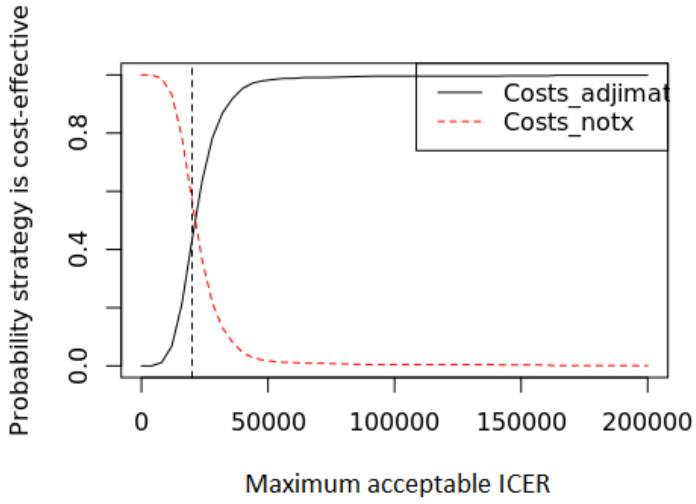
The cost-effectiveness plane of imatinib against no treatment at £20,000 per QALY shows that the ICER distribution is spread out across the diagonal with the mean cost-effectiveness estimate to its slight left (Figure 40). The larger part of the ICER distribution lies in the north-west quadrant.

Figure 40: C-E plane for imatinib against no treatment in GIST appraisal (2010)



The probability of imatinib being the most cost-effective strategy at a maximum acceptable ICER of £20,000 was at 43.4% (Figure 41).

Figure 41: CEAC of imatinib and no treatment in GIST appraisal (2010)



3.b What is the nature of uncertainty not captured by the PSA?

We are not aware of other uncertainties that may have been considered at the time.

3.c What is the temporal nature of uncertainty, e.g. is there more uncertainty beyond the trial period or is it resolvable with open-label follow up?

To reflect changes in recurrence rates over time, different hazard ratios were used for each year after surgery for up to five years. Because trial follow-up was not as long, the later hazard ratios were much more uncertain than the first year hazard ratio.

4. What is the uncertainty by groups of parameters?

The PUBs for either decision option are relatively large reflecting considerable uncertainty. The parameters that are most responsible for this large uncertainty are the rate of GIST recurrence five years after surgery as well as the hazard ratios for the different years after surgery (Table 9). This suggests that uncertainty could be resolved by a study with longer follow-up.

Table 9: Parameters causing decision uncertainty in the GIST appraisal (2010)

Parameters causing decision uncertainty	**Per Person EVPPI (£) at optimal decision (no treatment)**
Recurrence of GIST 5 years after surgery	4,099
Recurrence of GIST with imatinib 3 years after surgery	989

Grouped parameter analysis suggests that a large part of the existing decision uncertainty (83%) could be resolved by completely removing uncertainty surrounding the longer term GIST recurrence rates and imatinib hazard ratios (Table 10). We also tested whether information on only three years of recurrence rates and hazard ratios would reduce its value compared to five years and obtained the paradoxical result that information on three years would be better than on five years, although the difference was small. This result may be explained by the standard errors being large which is caused by only having run 1,000 PSA simulations and highlights the importance of using a sufficiently large number of iterations.

Table 10: Grouped parameters causing decision uncertainty in GIST appraisal (2010)

Parameters causing decision uncertainty	**EVPPI (£) at optimal decision (no treatment)**
GIST recurrence rates for up to 5 years after surgery	4,490
GIST recurrence rates for up to 3 years after surgery	4,510

5. What alternative treatment strategies might be available?

One year of imatinib treatment.

6. What measures of patient-based outcomes are available and measurable?

Recurrence-free and overall survival can be measured.

7. Is price a substantial part of overall costs associated with treatment?

Price is a substantial part of overall costs.

8. Are there any precedent PASs in place?

None identified.

9. Could price agreements be national or local?

It was stated in the guidance that costs of treatments may vary because there are negotiated procurement discounts but this is stated in all guidance. Other than that there was no further information on price agreements.

4.2.2. Possible MEAs in the GIST appraisal

Below are descriptions and verdicts on possible MEA schemes in the GIST appraisal.

Table 11: Verdicts on the usefulness of different price reduction schemes in the GIST (2010) appraisal

Price MEA scheme	**Possible use**	**Verdict**
A. Money back guarantee	For patients who experience recurrence of the GIST, the manufacturer could provide a refund.	Needs including in the model and testing for cost-effectiveness.
B. Conditional treatment continuation	Not applicable as patients stop treatment upon recurrence.	Not applicable.
C. Price linked to outcome	A price linked to outcome scheme by which the price will be paid over a longer period than the treatment is given could be suitable in this case and reduce risk. This could work	Needs including in the model and testing for cost-effectiveness but appears to be desirable given the present uncertainties on long-term effectiveness.

	along the lines of a leasing scheme that was proposed by Edlin *et al.*.[12] Instead of paying for the treatment when it occurs in the first three years after surgery, an alternative time horizon of, for example, five years could be chosen and the three year treatment price paid over that time horizon. If a patient experiences recurrence at any point in time, treatment as well as payments could be stopped.	
D. Discounted treatment initiation	This could have been realised in the shape of the manufacturer providing treatment for free or at a reduced price for the duration of a fixed period, which in this case need not be long to make imatinib cost-effective. It would not address the uncertainty, however.	Needs including in the model and testing for cost-effectiveness but should probably not stand alone.
E. Utilisation cap	A type of a utilisation cap is already in place in limiting the treatment period to three years. Another utilisation cap could have been explored at that time: limiting the treatment period to one year.	Evidence on effectiveness of one year treatment would have been needed at the time, therefore this does not seem feasible.
F. Fixed cost per patient	A fixed cost per patient is not applicable as treatment is already given for a fixed period and at fixed doses.	Not applicable.
G. Discount	With the cost of imatinib being slightly too high for it to be cost-effective, a discount could help both in bringing down the PSB and the PUB.	Needs including in the model and testing for cost-effectiveness.
H. Expenditure cap	There is no clear coherent rationale	Not applicable.

	for this.	
I. Price volume agreement	There is no clear coherent rationale for this.	Not applicable.

Table 12: Verdicts on the usefulness of different research schemes in the GIST (2010) appraisal

Research MEA scheme	Possible use	Verdict
4. RwR with reimbursement only	There are some model parameters that would benefit from research. A possible study design that could help reduce most of the decision uncertainty is a randomised study with long-term follow up on patients with and without imatinib treatment after surgery. An observational study could also be possible, in which patients treated with imatinib are followed up long-term, however, this may not address the uncertainty on long-term recurrence rates when no treatment is given.	A RwR scheme could be an option. Considerations are the length of the study, which will make re-visiting the decision possible at a much later stage only, delaying the benefits to be had from further research, and the overall time horizon that seems appropriate for imatinib for the adjuvant treatment of GIST. Modelling this trial, calculating its EVSI and the life-time value of it compared with the strategy-specific burden is therefore necessary. This scheme could be more worthwhile when paired with a discount or money back guarantee.
5. RwR with reimbursement and refund agreement	This could limit the cost to the payer through implementing a rule whereby a refund will be provided for patients experience recurrence.	See 4, but this scheme is more desirable from the payer perspective.
6. RwR with a conditional flexible pricing agreement	A possible pricing agreement could be a good option as an optimal price can be re-assessed at interim analyses, for instance, at each year of the study. Of course, each year, only a small part of the uncertainty would be reduced.	This would need assessing formally and comparing with another RwR scheme or price linked to outcome scheme.

Based on an examination of the uncertainty of the 2010 appraisal, we would have investigated different schemes for their expected value, with the most likely candidates being a price linked to outcome scheme, a discount and RwR schemes possibly combined with a discount. Further analyses, as proposed in the MEA risk analysis framework, could have indicated whether one or more of these schemes could and should have been adopted.

At the time, however, imatinib was not recommended and four years later re-assessed when more evidence had become available. Up to three years adjuvant treatment with imatinib was then recommended, as the committee concluded that the true value of the ICERs was between £3,610 and £12,100 per QALY gained for 1 year adjuvant imatinib compared with no adjuvant treatment, and between £16,700 and £30,000 per QALY gained for 3 year adjuvant imatinib compared with 1 year adjuvant imatinib.

The uncertainty in the base-case was greater with a PUB of £9,000 which was largely caused by changes in the model, the consideration of an additional treatment arm (one year imatinib) and a consideration of different parametric survival models. The committee noted that the main limitations of the model were associated with extrapolating effectiveness beyond the follow-up period, as none of the included trials had a follow-up period of more than three years (only one trial with such a follow-up period). There was also no trial that compared no adjuvant treatment with three years imatinib and an indirect comparison proved difficult as the assumption of proportional hazards did not hold.

4.2.3. Conclusion on adjuvant imatinib for GIST appraisal

In conclusion,

1. The small population PSB of 24 QALYs may have warranted price reduction and a RwR scheme.
2. An analysis of MEA schemes, such as the one proposed in our framework, could have resulted in a better study design that would have resulted in a reduction of decision uncertainty. This could then possibly have prevented a situation such as the one in 2014, in which decision uncertainty was larger than before.

4.3. DRY POWDER INHALERS FOR USE IN CYSTIC FIBROSIS

In 2013, NICE appraised colistimethate sodium and tobramycin dry powders for inhalation for treating pseudomonas lung infection in cystic fibrosis (TA276).[25] Treatment options for the pulmonary component of cystic fibrosis had so far included inhaled antibiotics effective against pseudomonas aeruginosa, the most frequent cause of lung infection in people with cystic fibrosis, such as nebulised colistimethate sodium or tobramycin. The potential drawback of nebulised treatment administration compared to the dry powder inhalers is reduced compliance due to the additional treatment burden associated with them, which may result in reduced effectiveness. For instance, using the nebuliser takes more time (up to an hour for patients in good health) than using a dry powder inhaler and involves preparing and cleaning of the nebuliser equipment.

At the time of appraisal, nebulised colistimethate and tobramycin were in use. Patients would tend to receive the less costly colistimethate as the first-line treatment, and receive tobramycin if colistimethate did not achieve response, or was associated with unacceptable adverse events, an excessive number of acute exacerbations or a loss of lung function. To assess the cost effectiveness of the dry powder inhalers (DPI), ideally the colistimethate DPI would have been compared with nebulised colistimethate sodium and the tobramycin DPI with nebulised tobramycin but, due to lack of evidence, both DPIs were compared with nebulised tobramycin.

The technology assessment group's model resulted in colistimethate DPI being dominated by nebulised tobramycin when the list price was used, with it being less effective and more costly. The tobramycin DPI was more effective than nebulised tobramycin but also more costly, with an ICER of £123,563 per QALY at list price. Both manufacturers offered a commercial-in-confidence discount which resulted in a cost-saving for each QALY lost for colistimethate DPI; and tobramycin DPI being more effective and less costly, thus dominating nebulised tobramycin. With those PASs in place, NICE recommended colistimethate DPI for use in patients who would benefit from colistimethate sodium but cannot tolerate it in its nebulised form and recommended tobramycin DPI when colistimethate sodium was contraindicated, was not tolerated or did not produce adequate clinical response. With many model parameters causing decision uncertainty, we explore whether other MEA schemes could have posed alternatives or complements to the employed PASs.

4.3.1. MEA design guidance questionnaire in the cystic fibrosis appraisal

1. What are the (number and characteristics of) treatment options?

Colistimethate sodium dry powder inhaler (DPI) is a formulation of colistimethate sodium supplied as hard capsules for use with an inhaler. It works by disrupting the structure of the bacterial cell membrane, leading to bacterial death. It is indicated for the management of chronic pulmonary infections caused by P. aeruginosa in patients with cystic fibrosis aged six years and older.

The recommended dosage for colistimethate sodium DPI is 1 capsule to be inhaled twice daily using the 'Turbospin' inhaler device which is a breath-activated, reusable dry powder inhaler. The price for a 28-day pack including one Turbospin inhaler is £968. The list price cost for 56 days of treatment is therefore £1,936. The manufacturer of colistimethate sodium DPI has agreed a patient access scheme with the Department of Health which makes colistimethate sodium DPI available with a discount applied to all invoices.

Tobramycin DPI is a formulation of tobramycin supplied as hard capsules for use with an inhaler. It primarily disrupts protein synthesis leading to altered cell membrane permeability, progressive disruption of the cell envelope and eventual cell death. Tobramycin DPI is indicated for the suppressive treatment of chronic pulmonary infection caused by P. aeruginosa in adults and children aged six years and older with cystic fibrosis.

The recommended dosage for tobramycin DPI is 112 mg tobramycin (4×28-mg capsules), administered twice daily for 28 days using the Podhaler device in alternating cycles of 28 days on treatment followed by 28 days off treatment. The price for a pack of 56×28-mg capsules and 1 Podhaler device is £447.50. The list price cost for 56 days of treatment is therefore £1,790. The manufacturer of tobramycin DPI had agreed a patient access scheme with the Department of Health which makes tobramycin DPI available with a discount applied to all invoices.

The above technologies will be referred to as colistimethate and tobramycin compared with nebuliser treatment in the following.

2. What is the base-case cost-effectiveness?

There were two models used in this appraisal. One compared colistimethate DPI with tobramycin nebuliser and the other compared tobramycin DPI with tobramycin nebuliser. The expected cost-effectiveness and net benefits at list price and agreed discounts for both colistimethate and tobramycin DPIs compared with the tobramycin nebuliser are shown in Table 13.

Table 13: Cost-effectiveness data in DPI for cystic fibrosis appraisal (5000 PSA runs)

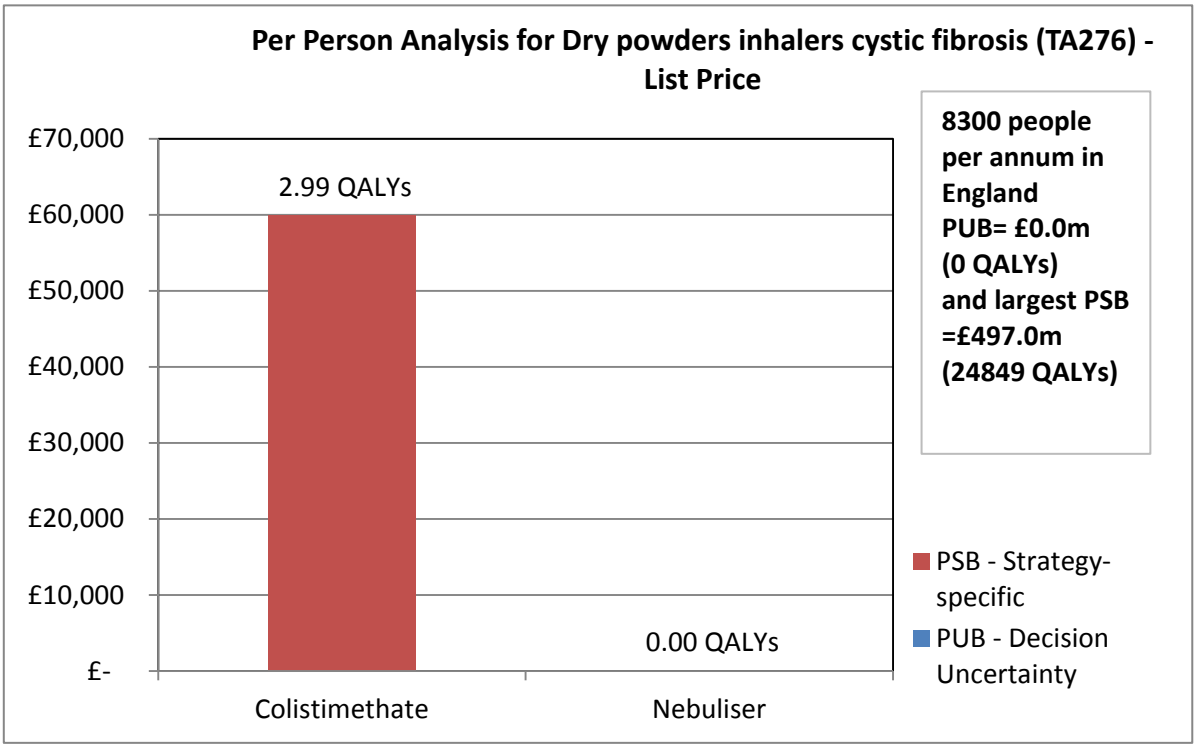
	List price				**Discount**			
	Colistimethate	Nebuliser	Tobramycin	Nebuliser	Colistimethate	Nebuliser	Tobramycin	Nebuliser
Expected Costs (£)	167,566	110,228	136,965	94,512	100,000	110,519	75,246	94,512
Expected QALYs	9.5	9.6	8.7	8.4	9.5	9.6	8.7	8.4
ICER vs nebulizer (£)	Dominated	-	123,571	-	52,672 (per QALY lost)	-	Dominating	-
Expected net benefits (at £20k/QALY)	21,507	81,384	37,539	73,121	85,955	81,659	99,258	73,120

3. What is the nature and scale of risk in this appraisal?

3.a What is the nature and scale of risk captured by the PSA?

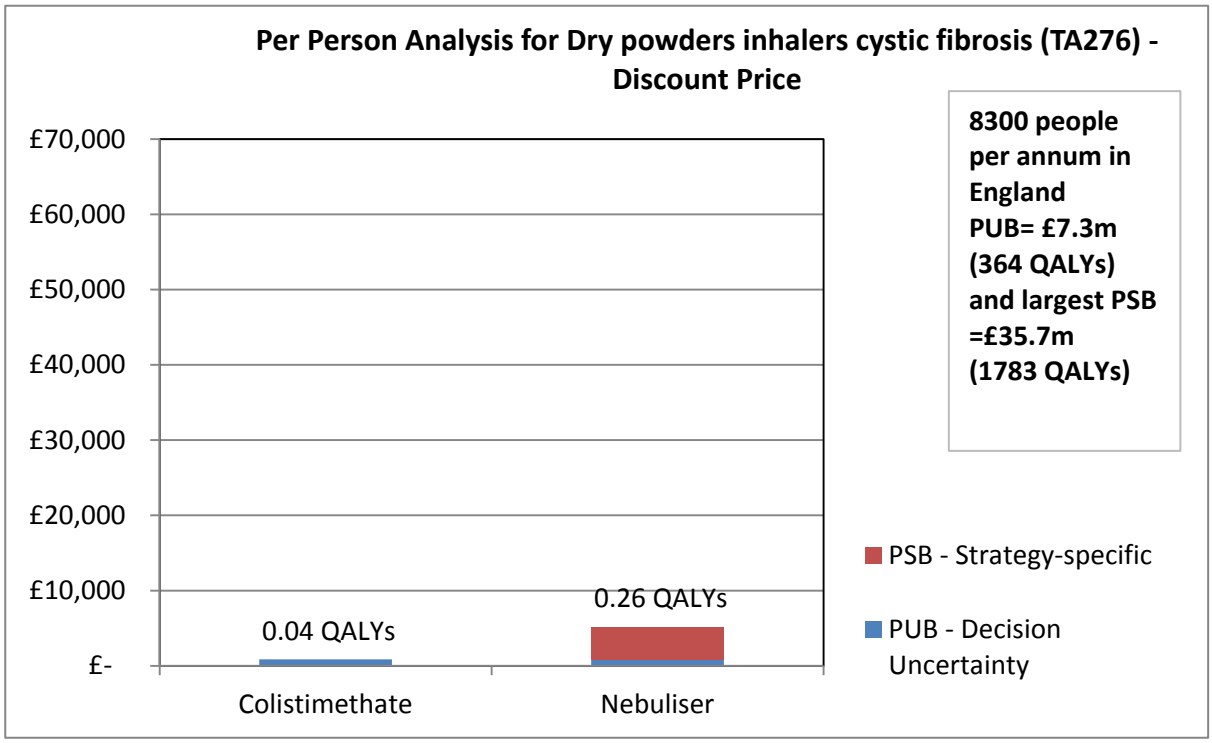
In the colistimethate comparison without the discount, the PUB associated with the decision is zero (Figure 42). This is because it is very unlikely for the colistimethate DPI to be cost-effective compared with the tobramycin nebuliser, which is also reflected in the large PSB of almost £60,000 associated with colistimethate, which at the population level translates into almost £500 million or 25,000 QALYs lost when recommending colistimethate DPI without the discount.

Figure 42: Risk analysis chart in colistimethate DPI for cystic fibrosis at list price



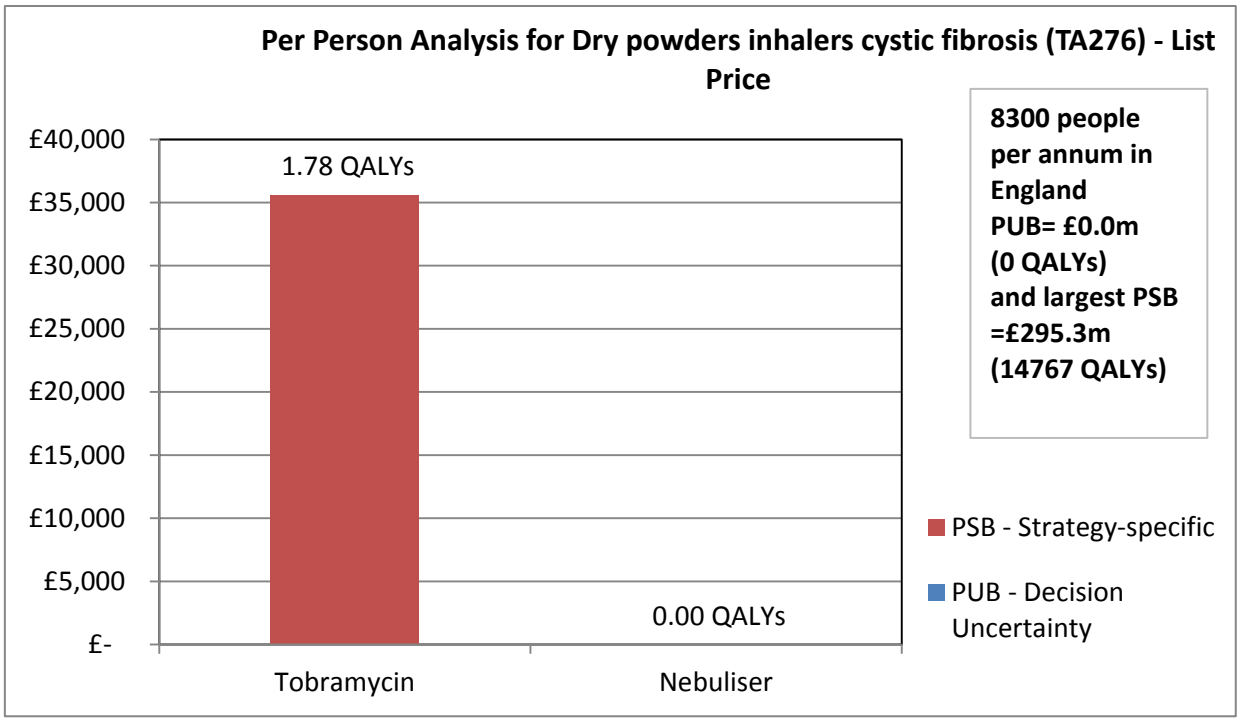
With the discount, the colistimethate DPI is still less effective than the nebuliser, but it is also much less costly. The PUB is now £876 per person and the PSB of not recommending the colistimethate DPI is £4,300 per person or £35.7 million at the population level (Figure 43).

Figure 43: Revised risk analysis chart in colistimethate DPI for cystic fibrosis at discount



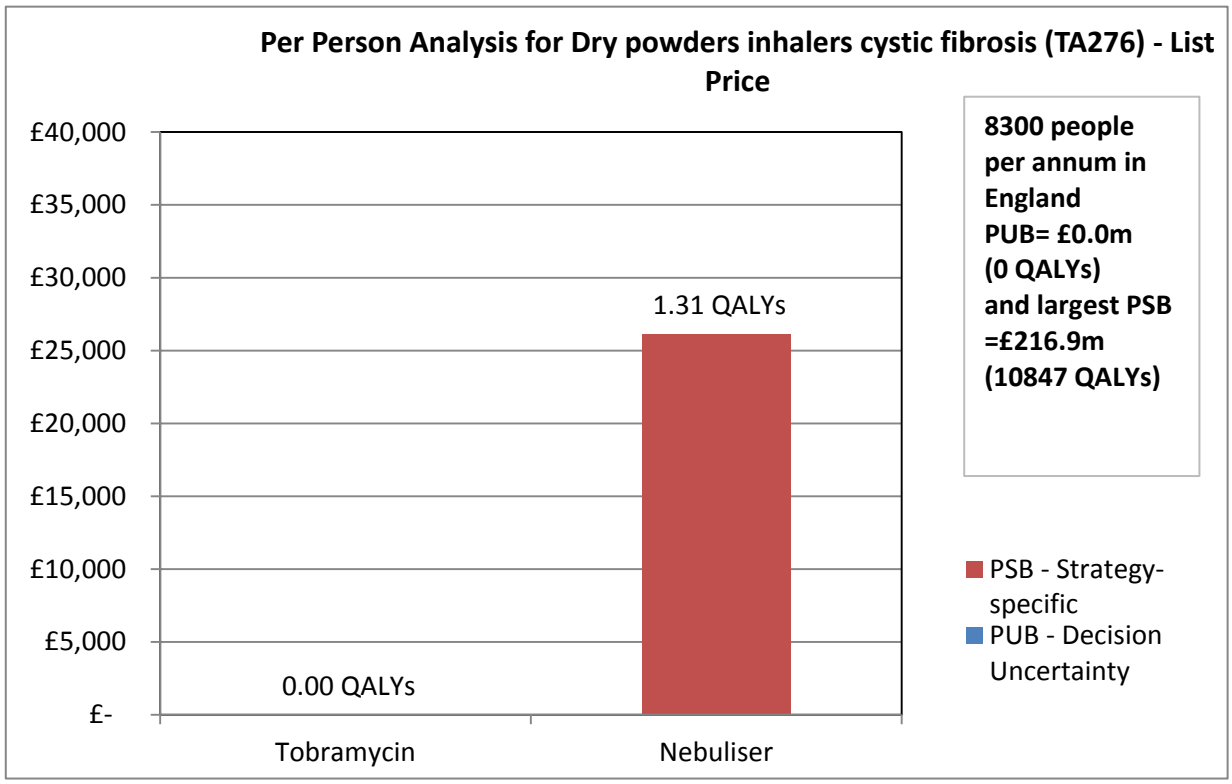
For the tobramycin DPI, the PUB is again zero (Figure 44), suggesting that it is far from the threshold. The PSB of recommending it nevertheless is large at almost £300 million or 15,000 QALYs at the population level.

Figure 44: Risk analysis chart in tobramycin DPI for cystic fibrosis at list price



At discount, the tobramycin DPI is dominating the nebuliser comparator, that is, it is more effective and less costly. This result is not associated with a PUB, but there is a PSB associated with not recommending the DPI of £26,000, or £217 million at the population level (Figure 45).

Figure 45: Revised risk analysis chart in tobramycin DPI for cystic fibrosis at discount



There is no P-SUB associated with recommending tobramycin DPI which is due to the fact that at its discounted price, it is clearly dominating nebulised tobramycin (Figure 47). At list price, Figure 47 shows that the mean cost-effectiveness estimate of tobramycin DPI would have been far removed from the acceptable ICER diagonal, explaining the large P-SUB (Table 13) associated with recommending it. The cost-effectiveness planes in Figure 46 and Figure 47 show that without an effective price reduction scheme, both DPIs cannot be cost-effective compared with nebulised tobramycin.

Figure 46: C-E plane colistimethate vs nebulizer with (bottom) and without (top) discount

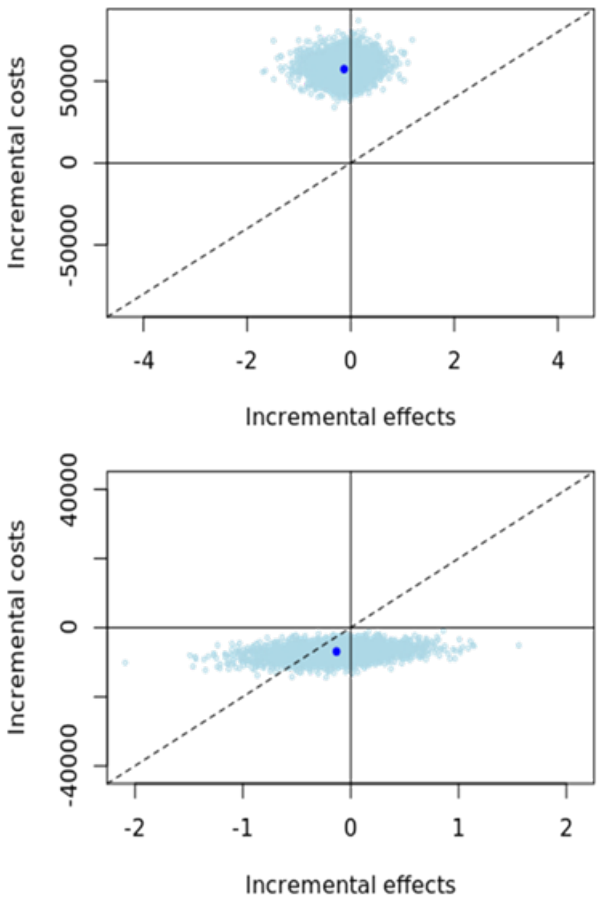
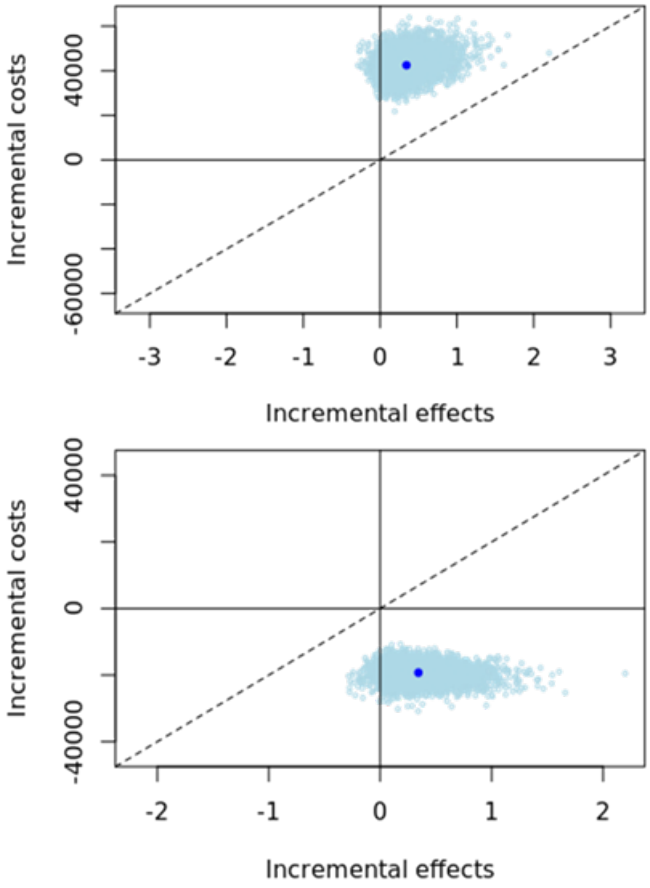


Figure 47: C-E plane tobramycin vs nebulizer with (bottom) and without (top) discount



With the PAS in place, the probability of colistimethate being the most cost-effective strategy compared with nebulised tobramycin at a maximum acceptable ICER of £20,000 was at 79% (Figure 48). The same probability for the tobramycin DPI was at 100% (Figure 49).

Figure 48: CEAC colistimethate DPI vs nebuliser

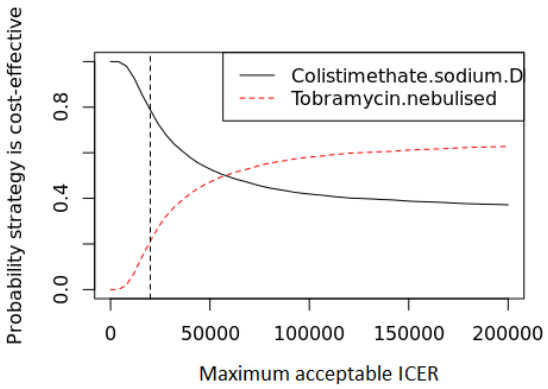
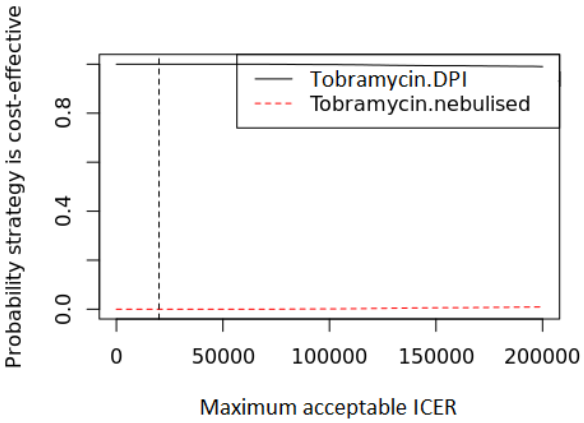


Figure 49: CEAC tobramycin DPI vs nebuliser



3.b What is the nature of uncertainty not captured by the PSA?

Other uncertainties included the existence of other treatment strategies in practice. For instance, some patients would be given nebulized colistimethate in the off-treatment cycle of tobramycin treatment. This has not been modelled as there is no evidence base for this.

Another source of uncertainty was that colistimethate DPI was compared with nebulized tobramycin while it should have been compared with nebulized colistimethate to reflect the real-world decision. However, there was no trial evidence comparing those two strategies.

The two trials included for use in the model were non-inferiority studies of colistimethate and tobramycin DPI against nebulized tobramycin (COLO/DPI/02/06 and EAGER, respectively). There was hence only evidence to suggest that either dry powder inhaler was not worse than nebulizer treatment but no evidence to show that either was more effective or equal.

One of the postulated benefits of DPIs versus nebulized treatments is their potential effect on patient compliance, which then would result in a QALY gain. The COLO/DPI/02/06 trial however, showed, to the contrary, improved adherence with nebulized treatment and there was no clear compliance data from the EAGER trial.

3.c What is the temporal nature of uncertainty, e.g. is there more uncertainty beyond the trial period or is it resolvable with open-label follow up?

Further uncertainty stemmed from the fact that the model used assessments at 24 weeks of treatment in the trial to extrapolate treatment effects over a life-time.

4. What is the uncertainty by groups of parameters?

With the decisions being very clear in three out of the four comparisons (Table 13) resulting in PUBs of zero for all decisions but one, we performed EVPPI analysis for the only decision option associated with an expected cost of uncertainty, or PUB, of greater than zero, that is recommending colistimethate DPI at discount. The individual parameters most responsible for decision uncertainty in the discounted colistimethate DPI model were different transition probabilities (Table 14). As we have observed in the previous appraisals, there were synergies in the way individual parameters contribute to decision uncertainty. In the discounted colistimethate model, 88% of the PUB associated with recommending colistimethate DPI was explained by all transition probabilities for patients moving between the health states and the EQ-5D utility values. When the utility values were removed, the value of the transition probabilities' contribution to decision uncertainty reduced to 73% of the overall PUB.

Table 14: Uncertain parameters in the colistimethate DPI at discount model

Parameters causing decision uncertainty	**EVPPI (£) at optimal decision (colistimethate DPI)**
Transition probability FEV<40% to FEV<40% (colistimethate DPI)	79.7
Transition probability FEV<40% to FEV40-69% (colistimethate DPI)	42.6
Transition probability FEV70-99% to FEV70-99% (tobramycin nebuliser)	39.8

5. What alternative treatment strategies might be available?

Apart from switching between tobramycin and colistimethate treatment which was discounted on the grounds that there was no evidence for it, none were identified.

6. What measures of patient-based outcomes are available and measurable?

Pulmonary function tests such as spirometry can be undertaken to measure lung function and yield the FEV1/FVC ratio (FEV% score). Because patients are heterogeneous with respect to their lung function, their age and other parameters which all influence the FEV% score, it is difficult to use it as an outcome measure based on which treatment would be recommended or not. It could potentially be used against each patient's baseline to determine an outcome based price.

7. Is price a substantial part of overall costs associated with treatment?

Price is a substantial part of overall costs.

8. Are there any precedent PASs in place?

Both, colistimethate and tobramycin are sold with straight discounts which are commercial in confidence.

9. Could price agreements be national or local?

It was stated in the guidance that costs of treatments may vary because there are negotiated procurement discounts but this is stated in all guidance. Other than that there was no further information on price agreements.

4.3.2. Possible MEAs in the DPI for use in cystic fibrosis appraisal

The different MEA schemes are discussed in Table 15 and Table 16.

Table 15: Verdicts on the usefulness of different price reduction schemes in the DPI in cystic fibrosis appraisal

Price MEA scheme	Possible use	Verdict
A. Money back guarantee	For patients who do not experience an improvement of agreed nature, the manufacturer could provide a refund. There is cystic fibrosis registry data available that could facilitate such a scheme. However, health states for cystic fibrosis can only be used for evaluation contingent on the patient's baseline. A suitable point in time for such an assessment would need to be defined, but there may not be evidence for this.	Needs including in the model and testing for cost-effectiveness.
B. Conditional treatment continuation	A type of this was used in the case of colistimethate as patients who do not respond to the treatment will be switched to tobramycin. This should not be considered a MEA scheme as the comparators are confounded here: colistimethate is simply recommended as first-line treatment, with tobramycin as second-line but the real comparator of colistimethate DPI is nebulised colistimethate and this was not modelled.	Not applicable.
C. Price linked to outcome	A price linked to outcome scheme in which price would be linked to patients' lung function relative to their baseline could be used in this case.	Needs including in the model and testing for cost-effectiveness.
D. Discounted treatment initiation	This could be an option for colistimethate by having the manufacturer providing treatment for free or at a reduced price for the	Needs including in the model and testing for cost-effectiveness.

	duration of a fixed period until response for each patient is established or not.	
E. Utilisation cap	Does not seem applicable as cystic fibrosis is a life-long condition.	Not applicable.
F. Fixed cost per patient	Does not seem applicable as cystic fibrosis is a life-long condition with individual exacerbations.	Not applicable.
G. Discount	With only marginal QALY gains and losses, cost-effectiveness results are very much dependent on the cost of the intervention. A discount, as was offered by the manufacturers, therefore seems to be a good option.	Results have shown this to make a major difference.
H. Expenditure cap	There is no clear coherent rationale for this.	Not applicable.
I. Price volume agreement	There is no clear coherent rationale for this.	Not applicable.

Table 16: Verdicts on the usefulness of different research schemes in DPI in cystic fibrosis appraisal

Research MEA scheme	**Possible use**	**Verdict**
4. RwR with reimbursement only	The above analysis had shown that both DPIs should not be recommended without an effective price reduction scheme in place. With the employed PASs, there is little doubt that tobramycin DPI is a cost-effective option. But for colistimethate DPI, there is residual uncertainty (PUB of £876 with discount compared to almost £60,000 at list price) as well as some uncertainty with respect to the chosen comparator. The model parameters that would benefit the most from research would mainly be the transition probabilities and utility values. A possible study design that could help reduce	An RwR scheme could be explored but to assess its desirability, a new model would be required that compares colistimethate DPI with nebulised colistimethate. Any proposed RwR scheme should only be used in combination with a price reduction scheme.

	most of the decision uncertainty is a randomised study with long-term follow up on patients using nebulised colistimethate against colistimethate DPI, recording utility values at each points of assessment.	
5. RwR with reimbursement and refund agreement	This could limit the cost to the payer through implementing a rule whereby a refund will be provided for patients in whom treatment with the DPI proves inferior.	See 4, but this scheme is more desirable from the payer perspective.
6. RwR with a conditional flexible pricing agreement	Is not applicable as the cost has been brought down through the PAS.	Not applicable.

4.3.3. Conclusion on DPI for cystic fibrosis appraisal

1. The above analysis showed that, if the appropriate comparators are not modelled, it is difficult to assess MEA schemes for their value. It is therefore advisable that the appropriate comparisons are modelled, even when there is no evidence to support this, and to allow this uncertainty to be reflected in the chosen probability distributions.

4.4. CONCLUDING REMARKS ON IDENTIFYING SUITABLE MEA SCHEMES USING THE MEA RISK ANALYSIS FRAMEWORK PART I IN THREE APPRAISALS

1. In this section, we illustrated the use of the MEA risk analysis framework, including the MEA design guidance questionnaire and the taxonomy, to assess the risk due to decision uncertainty, the strategy-specific risk and consider different MEA options that could reduce that risk.
2. We have shown that, based on the knowledge of the PUB and the PSB as well as a detailed analysis of the issues present at the time of appraisal, many MEA schemes can be eliminated from the options available to the decision-maker in each setting.
3. Further quantitative analysis of the effect of the remaining MEAs needs to be undertaken to make an informed and transparent decision.
4. Our analysis showed that sometimes, MEA schemes such as conditional treatment continuation lack the evidence to support their clinical benefit.

5. The importance of modelling the appropriate comparators, even with little evidence, was shown in two appraisals.
6. This chapter highlighted that large PSBs can only be overcome by significant reductions in price – but small PSBs could be addressed with price reductions or accepted in order for further research to be conducted within an RwR decision.
7. When an MEA analysis is not undertaken, this may result in study designs that do not reduce decision uncertainty and subsequent appraisals in which there is greater decision uncertainty than before.

5. APPLICATION OF THE FULL MEA RISK ANALYSIS FRAMEWORK: PAZOPANIB FOR THE TREATMENT OF ADVANCED RENAL CELL CARCINOMA

5.1. BACKGROUND ON THE PAZOPANIB APPRAISAL

We chose the pazopanib for treatment of advanced renal cell carcinoma as our main case study because it reflected well the situation that decision-makers may find themselves in with the recent regulatory changes. In 2011, the National Institute for Health and Care Excellence (NICE) appraised pazopanib for the first-line treatment of advanced renal cell carcinoma (TA215).[26] Pazopanib at the time had a conditional marketing authorisation for 'the first-line treatment of advanced renal cell carcinoma and for patients who have received prior cytokine therapy for advanced disease'. Conditional marketing authorisations are constrained in their time and conditional on emerging further evidence confirming the effectiveness or safety profile of a new technology. Pazopanib obtained marketing authorisation conditional on the then ongoing head-to-head non-inferiority trial of pazopanib versus sunitinib (COMPARZ trial).

At the time of appraisal, two technologies were in use in the UK NHS for the treatment of advanced renal cell carcinoma, sunitinib and interferon-α. Sunitinib had been appraised in 2009 (TA169) and found to be cost-effective against interferon-α based on end of life valuation. End of life criteria were fulfilled as evidence suggested that sunitinib provided a gain in survival of more than three months in a small patient population that has an expected life expectancy of less than 24 months.[4] End of life valuation allows a QALY weighting that is different from the typically employed maximum acceptable ICER range of £20,000 to £30,000 per QALY.[4]

Evidence for pazopanib was suggestive of a similar survival gain to sunitinib and an improved safety profile. The caveat was that the magnitude of the survival gain associated with pazopanib was highly uncertain. There had only been one study, the VEG105192 trial, assessing pazopanib and best supportive care against placebo and best supportive care with the primary outcome of progression-free survival and the secondary outcome of overall survival. Due to patients in the placebo arm receiving pazopanib at disease progression, the

estimates of overall survival were confounded by cross-over and highly uncertain, despite using different methods of adjusting for cross-over.

Hazard ratios for overall survival were obtained from a network meta-analysis that included seven studies. All hazard ratios used in the model were obtained by using a Cox proportional hazard model against interferon-α and an underlying Weibull survival model. The hazard ratio of pazopanib against sunitinib was not calculated, thereby assuming that their hazard rates were uncorrelated.

In the final decision, the survival gain produced by pazopanib was assumed to be comparable to sunitinib, implying that EoL criteria would be fulfilled for both pazopanib and sunitinib. The Committee agreed that the comparison of pazopanib versus interferon-a would meet the End Of life criteria and an additional QALY weighting could be applied. The extent of this weight was not reported in the FAD and therefore, for simplicity, a MA-ICER of 50k was used for our analyses. Because pazopanib was not likely to demonstrate an additional survival gain of three months over sunitinib, in the comparison of pazopanib against sunitinib, acceptable ICERs of £20,000 to £30,000 per QALY were used.

At those acceptable ICERs, pazopanib was shown to be cost-effective compared with interferon-α at the price suggested by the manufacturer. It was, however, not cost-effective compared with sunitinib. But the manufacturer proposed a patient access scheme of a straight discount, at which pazopanib became less costly and more effective than sunitinib.

Two features of this appraisal hence stood out: 1. The application of end of life (EoL) valuation to pazopanib and sunitinib, but not the other comparators and 2. The lack of evidence for the overall survival gain of pazopanib against the other comparators. The former meant that the choice of the appropriate comparator and maximum acceptable ICER for pazopanib was not straightforward. The latter, the lack of evidence, meant that not only would any decision be made under considerable uncertainty, leading to a large payer uncertainty burden, but also, the decision of whether end of life criteria applied or not would be highly uncertain.

The objectives of this study hence became to use the MEA risk analysis framework to 1. analyse the risk present at the time of appraisal, 2. Assess the value of different MEA

schemes and 3. incorporate end of life criteria and end of life valuation in the MEA risk analysis framework.

5.2. METHODS FOR APPLYING THE MEA RISK ANALYSIS FRAMEWORK IN THE PAZOPANIB APPRAISAL

The results section presents two different analyses: the first one investigates the cost-effectiveness data and decision uncertainty present in the pazopanib appraisal using EoL valuation. This is closer to reflecting the true decision-making at the time. In the second analysis, a maximum acceptable ICER of £20,000 per QALY is used, neglecting EoL valuation. The use of the two different analyses serves illustration purposes: we will demonstrate the potential of different types of MEA schemes in reducing the PUB when there is no additional strategy-specific risk in the first analysis; and the value of different MEA schemes in terms of the P-SUB reduction when there is a large strategy-specific risk in the second analysis.

Because of the large uncertainty surrounding the end of life criterion of a greater than three months incremental survival gain of pazopanib over comparators, we developed a framework for incorporating incremental survival gain into the probabilistic net benefit analysis. In our EoL framework, the maximum acceptable ICER could vary in each iteration of the PSA according to the incremental survival gain provided by pazopanib and sunitinib in that iteration. More specifically, the acceptable ICER that was applied in the pazopanib and sunitinib net benefit calculations could adopt the values of either £50,000 per QALY for fulfilled EoL criteria in each iteration of the PSA or £20,000 per QALY when the EoL criterion was not fulfilled in that iteration. More detail on these steps can be found in the technical appendix.

For both analyses, we followed the quantitative MEA risk analysis framework developed in Section 2 and present results in five sections.

1. To obtain cost-effectiveness data, the PSA was performed with a number of 50,000 simulations for both price scenarios. The large number of simulations was chosen because of the large number of model parameters. For each simulation, parameter values and values for

resulting costs and effects for each of the four technologies (pazopanib, sunitinib, interferon-α and best supportive care) were stored for both price scenarios.

2. The PUB and PSB were calculated using an extension to the SAVI tool.

3. SAVI also provided results for parameter expected value of perfect information (EVPPI) analysis, identifying those parameters which contributed most to the overall decision uncertainty. Group EVPPI analysis gave an indication of uncertainty that could be addressed through further research.

4. We then considered the choice of appropriate MEA schemes by answering the MEA design guidance questionnaire, for both EoL and non-EoL analyses. Based on these findings, we gave a potential description of and verdict on each MEA scheme and selected the most appropriate ones.

5. We then assessed the chosen MEA schemes by modelling the selected price reduction and research schemes. To model price reduction schemes, we simply changed the price to its discounted version and then performed the PSA and other calculations again. For a more complex price reduction scheme, such as the money back guarantee scheme, it was necessary to put in the price as a function of survival gain in each row of the PSA (more detailed methods on the implementation of the money back guarantee scheme are in the appendix).

For research schemes, the value of a proposed research study can be evaluated by modelling the trial and performing EVSI analysis. Modelling the trial requires a statistical model being specified, that describes the data to be collected, e.g. a normal, or Weibull distribution. The sample size and follow-up period also need to be defined. A dataset from the proposed study is then simulated for each iteration of the PSA, conditional on the parameter values of that iteration. Running an EVPPI analysis on the generated trial data in SAVI then provides the EVSI of that trial.

As mentioned above, at the time of appraisal pazopanib held a marketing authorisation conditional on the results from the then ongoing COMPARZ trial. While this is therefore not a classic RwR scheme in which the payer reimburses for the drug used in research as well as in the general patient population, we are treating it as a RwR scheme here, to consider its

value. With the main uncertainty being contributed by the uncertainty surrounding the hazard ratios of pazopanib, the COMPARZ trial was a multi-centre head-to-head non-inferiority trial of pazopanib and sunitinib with 1,100 patients recruited and randomised into two trial arms with the primary outcome of progression-free and secondary outcome of overall survival. The trial had no defined follow-up time as patients were recruited over a long period of time and the trial stopped only five months after the last patient was recruited. We therefore simulated four different specifications of sample sizes and follow up durations.

The above-mentioned money back guarantee scheme was coupled with the observation of registry data to gather evidence on the real-world survival data of pazopanib. This required modelling another research study, of the size of the annual patient population affected by advanced renal cell carcinoma, with different follow-up durations of one, two or three years, this time with only one trial arm of patients receiving pazopanib. The EVSI of that real-world evidence was then obtained by performing EVPPI analysis on the generated data using SAVI.

5.3. RESULTS OF THE MEA RISK ANALYSIS FRAMEWORK APPLIED IN THE PAZOPANIB APPRAISAL

5.3.1. The MEA risk analysis framework applied to pazopanib appraisal with end of life valuation

The NICE Committee agreed the End of Life criteria were met. The evidence to support the 3-month survival extension for pazopanib or sunitinib compared with interferon-α was, however, highly uncertain. Our newly developed EoL framework can help address this.

1. Cost-effectiveness data for pazopanib under end of life valuation

Under EoL valuation, pazopanib was the technology with the highest expected net benefit (Table 17). Of course, these results are also influenced by our rule that interferon-α and best supportive care only ever be valued at a maximum acceptable ICER of £20,000 per QALY, which was based on these comparators not meeting all EoL criteria.

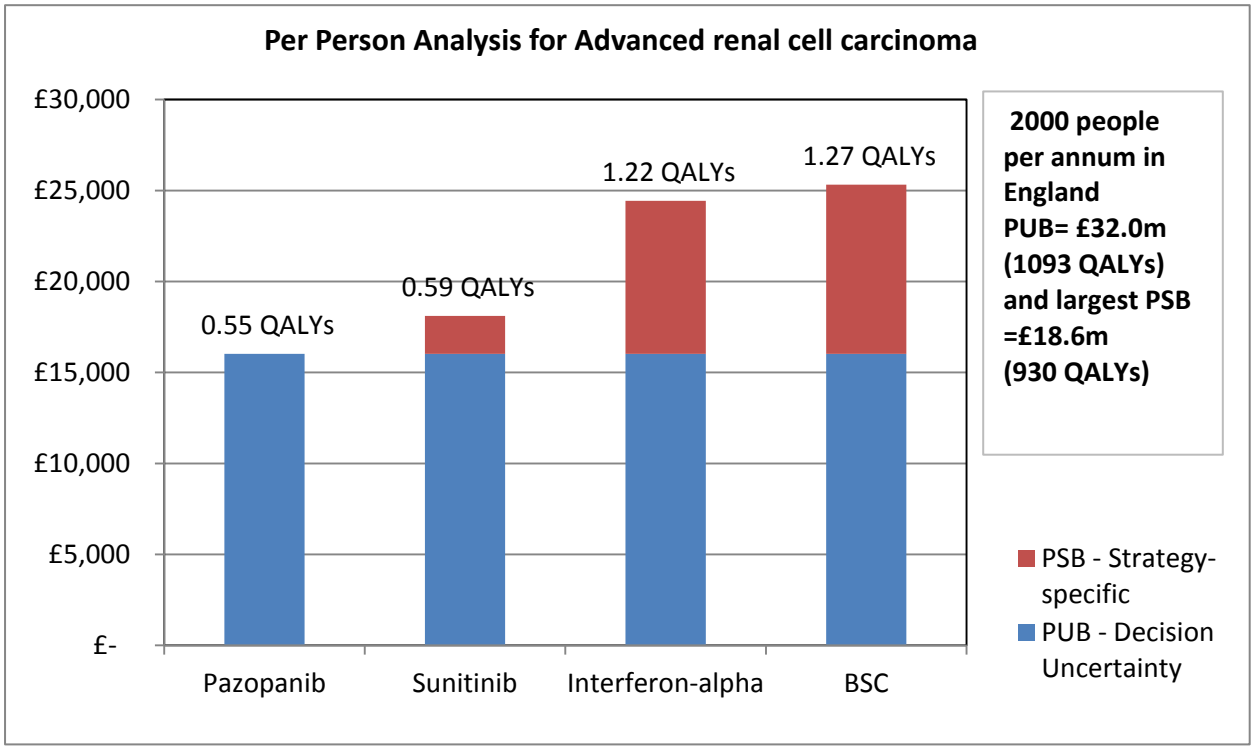
Table 17: Cost-effectiveness data in pazopanib appraisal (50,000 PSA runs)

	Pazopanib	**Sunitinib**	**Interferon-α**	**Best supportive care**
Expected Costs (£)	40,148	36,366	8,383	4,103
Expected QALYs	2.02	1.90	1.25	0.99
ICER against interferon-α	41,100	42,767	-	NA
ICER against sunitinib	31,901	-	NA	NA
Expected net benefits (EoL valuation)	25,007	22,925	16,591	15,708
Expected net benefits (£20k/QALY)	284	1,695	16,591	15,708

2. The risk associated with each strategy and decision uncertainty under EoL valuation

The PUB was large at approximately £16,000 per person (Figure 50). Pazopanib, being the cost-effective strategy, did not exhibit a PSB, but the strategy-specific burdens associated with the other interventions are shown in Figure 50. In England, there are approximately 2,000 people that are affected by advanced renal cell carcinoma on an annual basis. The annual PUB is therefore at £32 million, or almost 1,100 QALYs. If best supportive care was the recommended strategy, the P-SUB, that is the cost that the payer incurs with this decision option, caused by uncertainty and the strategy-specific risk, could increase to a total of more than £50 million annually.

Figure 50: Risk analysis chart in the pazopanib appraisal with EoL valuation



3. Uncertainty caused by parameters in pazopanib appraisal under EoL valuation

The most uncertainty is caused by the parameters associated with the overall survival and progression-free survival hazard ratios of pazopanib against interferon-α and sunitinib against interferon-α (Table 18).

Table 18: Parameters causing the most decision uncertainty in pazopanib appraisal (EoL valuation)

Parameters causing decision uncertainty	**EVPPI (£)**
Overall survival hazard ratio of pazopanib vs interferon-α	14,394
Overall survival hazard ratio of sunitinib vs interferon-α	2,925
Progression-free survival hazard ratio of pazopanib vs interferon-α	1,310
Progression-free survival hazard ratio of sunitinib vs interferon-α	412

4. Choice of MEA schemes under EoL valuation

Even when a new technology is considered cost-effective, it may still be worthwhile considering the use of a MEA scheme. Pazopanib, despite yielding the largest expected net benefit, still remains a largely uncertain decision option. Two MEA schemes were put in place at the time of appraisal: a straight discount and re-visiting the decision based on new evidence becoming available from the COMPARZ trial.

To select other possibly valuable MEA schemes that could have been put in place to aid a recommendation of pazopanib, answers to the MEA design guidance questionnaire are in Table 19, and a verdict on each of the MEA schemes identified in the taxonomy in Table 20 and Table 21.

Table 19: MEA design guidance questionnaire for pazopanib appraisal

1. What are the (number and characteristics of) treatment options?	Best supportive care (BSC), interferon-α, sunitinib, pazopanib
2. What is the base-case cost-effectiveness?	Pazopanib was the most effective but also the most costly treatment strategy. It was not cost-effective at a maximum acceptable ICER of £20,000 per QALY. It was the strategy with the largest expected net benefit when EoL valuation was used.
3. What is the nature and scale of risk in this appraisal?	
3.a What is the nature and scale of risk captured by the PSA?	The PSA reflects that there is large decision uncertainty.
3.b What is the nature of uncertainty not captured by the PSA?	If a £20,000 / QALY threshold is used, the PSA does not capture EoL valuation. If a larger threshold is used, the uncertainty surrounding the decision of whether EoL criteria apply or not is not captured. Our approach with a variable threshold addresses this.
3.c What is the temporal nature of uncertainty, e.g. is there more uncertainty beyond the trial period or is it resolvable with open-label follow up?	None identified
4. What is the uncertainty caused by individual / groups of parameters?	There is large uncertainty associated with the hazard ratios for overall and progression-free survival of pazopanib and sunitinib compared with interferon-α which appear to cause the largest part of overall decision uncertainty.
5. What alternative treatment strategies might be available?	None identified
6. What measures of patient-based outcomes are available and measurable?	Overall survival, progression-free survival
7. Is price a substantial part of overall costs associated with treatment?	Yes
8. Are there any precedent PASs in place?	The manufacturer offered a discount of 12.5% to the daily cost of pazopanib.
9. Could price agreements be national or local?	Probably national.

Table 20: Verdicts on the usefulness of different price reduction schemes in the pazopanib appraisal

Price MEA scheme	Possible use	Verdict
A. Money back guarantee	For patients with shorter than an agreed, or a comparator's overall or progression-free survival, the manufacturer could provide a refund.	Needs including in the model and testing for cost-effectiveness.
B. Conditional treatment continuation	Discontinuation of pazopanib in case of disease progression is part of the treatment pathway and therefore not an option for a MEA.	Not applicable.
C. Price linked to outcome	The price of treatment could be linked with the length of overall survival in some function.	Needs including in the model and testing for cost-effectiveness but as we do not know what functional form the price link with outcomes could take, this seems less applicable than a money back guarantee (scheme A).
D. Discounted treatment initiation	This could be realised in the shape of the manufacturer providing treatment for free or at a reduced price for the duration of a fixed period, but since survival can be relatively short in advanced renal cell carcinoma this may not be a good option.	Probably not a good option.
E. Utilisation cap	With the aim of treatment being to prolong patients' lives, the longer they are on it the better. A utilisation cap would penalise the manufacturer for prolonged survival and does not seem appropriate.	Not appropriate.
F. Fixed cost per patient	This could ensure that no over-spend would be made for patients who are treated for a long time. Similarly to scheme E, this is inappropriate.	Not appropriate.
G. Discount	A discount had already been	Different values of discounts need

	proposed and although it did not address the large uncertainty, it is a good option to reduce costs. Different values of the discount could therefore be tested.	testing.
H. Expenditure cap	There is no clear coherent rationale for this.	Not applicable.
I. Price volume agreement	There is no clear coherent rationale for this.	Not applicable.

Table 21: Verdicts on the usefulness of different research schemes in the pazopanib appraisal

Research MEA scheme	Possible use	Verdict
4. RwR with reimbursement only	The main parameters benefitting from research clearly were the hazard ratios associated with progression-free, and more importantly, overall survival. The most promising RwR scheme is therefore seen in a trial similar to the COMPARZ trial.	The trial needs to be modelled to assess the expected value of research. Assumes that based on the new evidence, the appropriate decision will be taken, even if it means reversal of the previous decision.
5. RwR with reimbursement and refund agreement	This could limit the cost to the payer through implementing a rule whereby a refund will be provided for patients enrolled in the study who do not reach remission after a certain time period of treatment. This scheme is more desirable from the payer perspective.	The trial needs to be modelled and the cost-saving achieved through the refund rule needs to be assessed. Assumes that based on the new evidence, the appropriate decision will be taken, even if it means reversal of the previous decision.
6. RwR with a conditional flexible pricing agreement	A possible pricing agreement could involve interim analysis at certain time points at which the health outcomes (for instance, hazard ratios of overall survival) are assessed. The price may then be adapted for the ICER to fall in an agreed acceptable range. The data collection is more short-term, therefore there is less scope for this scheme.	Not applicable.

Summarising the verdicts in Table 20 and Table 21, the MEA schemes that were actually used at the time of appraisal appeared to have been appropriate. With the uncertainty being greatest for a group of parameters that can be resolved in one trial, a RwR scheme with a trial such as the COMPARZ trial seemed to be a good strategy. The choice for the discount scheme was also justified as the cost of pazopanib was the greatest among the different treatment strategies.

The most promising schemes alternative to those schemes that were actually in place were scheme A., a money back guarantee where registry data are used to identify patient-level survival data, and scheme G., a discount of a different value to the one proposed by the manufacturer. Scheme A. could furthermore be combined with a rule of re-visiting the decision after having collected a few years' worth of registry data, basing the decision on the collected evidence.

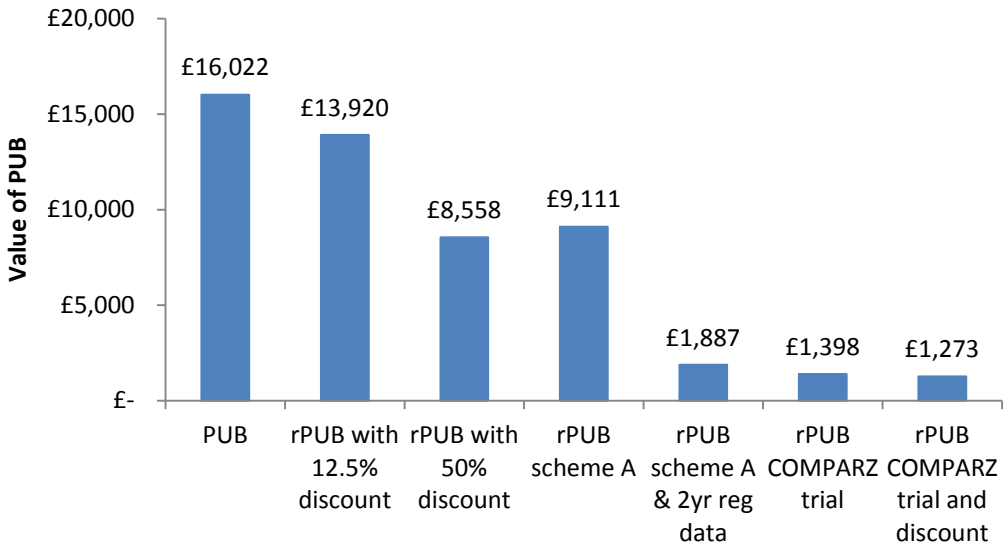
5. Assessment of selected MEA schemes under EoL valuation

The application of the full MEA risk analysis framework resulted in the COMPARZ trial paired with a discount providing the greatest reduction in the PUB (Figure 51). Only the 12.5% discount proposed by the manufacturer reduced the PUB by more than 10% (Figure 51) but the PUB still remained large at £14,000 per person. Of course, larger discounts would enable the PUB to come down further, as shown with the example value of a 50% discount in Figure 51. The money back guarantee scheme (Scheme A) reduced the PUB more than the discount proposed by the manufacturer, but less than the 50% discount.

Research schemes enabled further reduction of the PUB. When scheme A was combined with collecting data and re-visiting the decision at a future point in time (chosen here to be at two years after the current appraisal), the PUB reduces further to £1,900. A trial similar to the COMPARZ trial can help reduce the largest part of the PUB, achieving a residual PUB of only £1,400 (Figure 51). Together with the discount, the PUB can be reduced by another £100, making the combination of the trial and the discount the most effective MEA scheme when EoL valuation is in place. Of course, this only holds true if a re-evaluation is scheduled for the time the research evidence becomes available.

The recommendation of both the COMPARZ trial and the discount meant a saving to the payer of £14,800 per person, the difference between the initial PUB and the residual PUB. Over the population, this results in £29.5 million per year of savings.

Figure 51: MEA risk reduction chart in pazopanib appraisal under EoL valuation



5.3.2. The MEA risk analysis framework applied to a hypothetical appraisal based on pazopanib but without end of life valuation

1. *Cost-effectiveness data for pazopanib without end of life valuation*

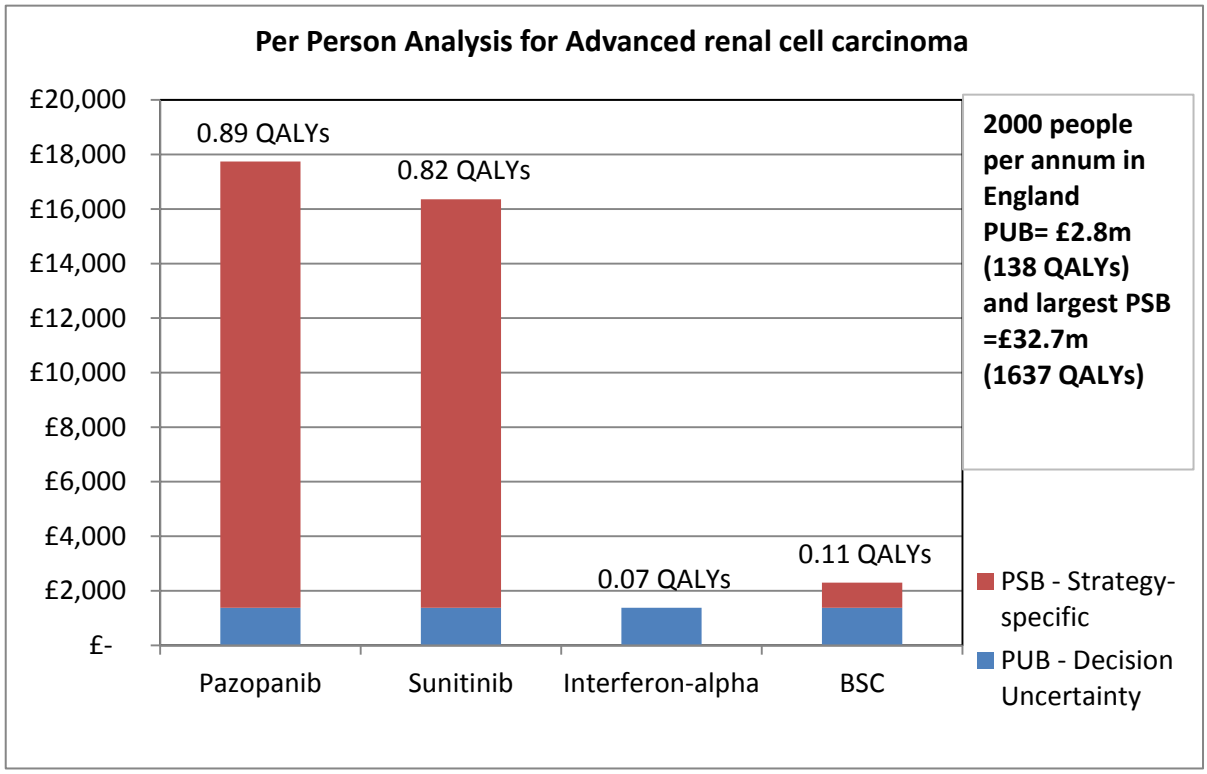
When the EoL framework was not applied and a maximum acceptable ICER of £20,000 per QALY was used, pazopanib would have been expected to be the strategy with the lowest expected net benefit (Table 17). Of course, this is a hypothetical situation and does not reflect the committee decision making at the time. Pazopanib would not be cost-effective compared with interferon-α or sunitinib at an acceptable ICER of £20,000 (or £30,000) per QALY. Best supportive care was less effective than all the other strategies. Interferon-α would be the strategy with the largest expected net benefit at a maximum acceptable ICER of £20,000 per QALY (Table 17). This made interferon-α, rather than pazopanib, the best decision option.

2. The risk associated with each strategy and decision uncertainty without EoL valuation

The PUB associated with this hypothetical example in which we ignored EoL valuation was large at approximately £1,400 (Figure 52), but it was by far not as large as under EoL valuation (£16,000) (Figure 50). With interferon-α now being the best choice, there is no strategy-specific burden associated with it (Figure 52). Pazopanib had the smallest expected

net benefit which is reflected in the greatest PSB among the different interventions (Figure 52).

Figure 52: Risk analysis chart in the hypothetical pazopanib appraisal without EoL valuation



The importance of considering the strategy-specific burdens is highlighted by the size of the P-SUB associated with recommending pazopanib instead of the optimal option. Over the estimated population affected by advanced renal cell carcinoma (at approximately 2,000 patients annually), the annual opportunity loss due to uncertainty was estimated at £2.8 million or 138 QALYs lost until the decision changes, if interferon-α was adopted (Figure 52). These are large numbers, especially considering that these are annual and would be multiplied by the number of years this decision remains relevant. But the P-SUB of recommending pazopanib is twelve times the size of the PUB. The P-SUB reflects the risk the payer, here the UK NHS, takes when pazopanib is recommended and this risk is quantified at more than £35 million or 1,768 QALYS lost per year.

To illustrate the strategy-specific risk in a different way, the cost-effectiveness plane is shown in Figure 53. The combination of mean incremental QALYs and costs associated with pazopanib against interferon-α is indicated by the blue dot in the middle which lies to the top

left of the ICER diagonal. If pazopanib was chosen over interferon-α, the distance between that blue dot and the ICER diagonal equals the PSB associated with pazopanib. Pazopanib would therefore have been expected to be much less cost-effective than interferon-α (Figure 53). However, the wide spread of the ICER distribution represents the large uncertainty in this decision as the true ICER could lie anywhere within that distribution, also to the right of the ICER diagonal, which would then make pazopanib cost-effective.

The cost-effectiveness acceptability curve (CEAC) shows that pazopanib was the most likely cost-effective strategy at acceptable ICERs larger than approximately £35,000 per QALY (Figure 54). Below that, best supportive care was the strategy most likely to be cost-effective at lower ICERs up to approximately £15,000 per QALY, and interferon-α between £15,000 and £35,000 per QALY gained.

Figure 53: C-E plane of pazopanib vs interferon at £20k / QALY

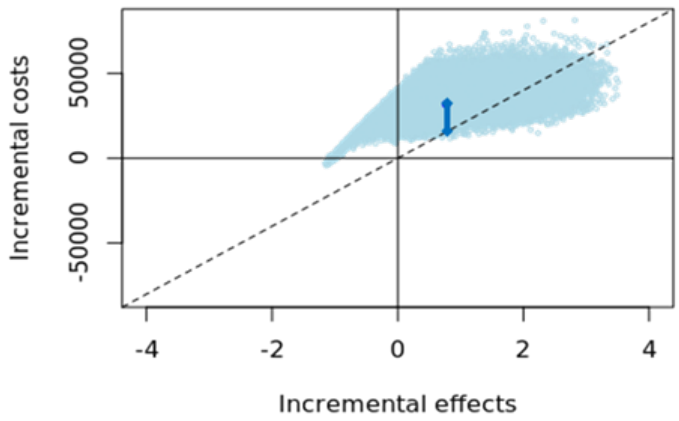
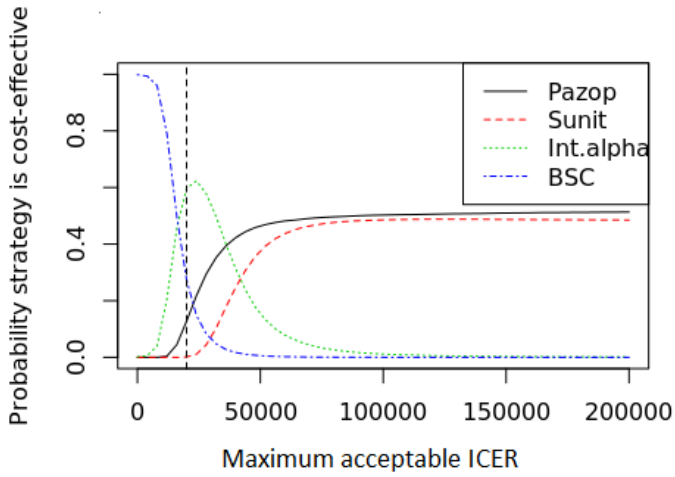


Figure 54: CEAC of pazopanib and comparators with indication of the acceptable ICER of £20k / QALY



3. Uncertainty caused by parameters in pazopanib appraisal without EoL valuation

The parameters contributing the most to overall decision uncertainty without EoL valuation differed from the ones identified under EoL valuation. This was because at different maximum acceptable ICERs the decision was between different strategies, as shown in the above CEAC. And for different decision options, different sets of parameters may matter more.

At an acceptable ICER of £20,000 per QALY, the parameters causing the most uncertainty included the overall survival hazard ratio of pazopanib against interferon-α and the overall survival hazard ratio for best supportive care against interferon-α as well as cost parameters related to certain health states (Table 22).

Table 22: Parameters causing the most decision uncertainty in pazopanib appraisal (£20,000 / QALY)

Parameters causing decision uncertainty	**EVPPI (£)**
Overall survival hazard ratio of pazopanib vs interferon-α	664
Overall survival hazard ratio of best supportive care vs interferon-α	221
Other monthly cost related to post-progression with interferon-α	101
Other monthly cost related to post-progression with best supportive care	26

4. Choice of MEA schemes in hypothetical pazopanib appraisal without EoL valuation

We assumed that the potentially useful MEA schemes were the same as under the EoL valuation. This assumption was based on reductions in price being even more relevant with a large PSB. And while the parameters contributing the most to decision uncertainty had changed, the one with the greatest effect remained the overall survival hazard ratio of pazopanib against interferon-α, which we could learn about in a trial similar to the COMPARZ trial. Of course, a trial of pazopanib against best supportive care might reduce more of the decision uncertainty, but the overall survival hazard ratio of best supportive care against interferon-α only accounts for 15% of the overall decision uncertainty and research on this parameter would therefore have a relatively small effect. We therefore tested the value of different trial designs, different discounts, a money back guarantee scheme, and monitoring registry data.

5. Assessment of selected MEA schemes without EoL valuation

With the 12.5% discount to daily cost proposed by the manufacturer, pazopanib would look slightly less costly than sunitinib. However, this was not enough to yield the highest expected net benefit at a maximum acceptable ICER of £20,000 per QALY (Table 23).

Table 23: Cost-effectiveness data in hypothetical pazopanib appraisal at £20k / QALY with discount (50,000 PSA runs)

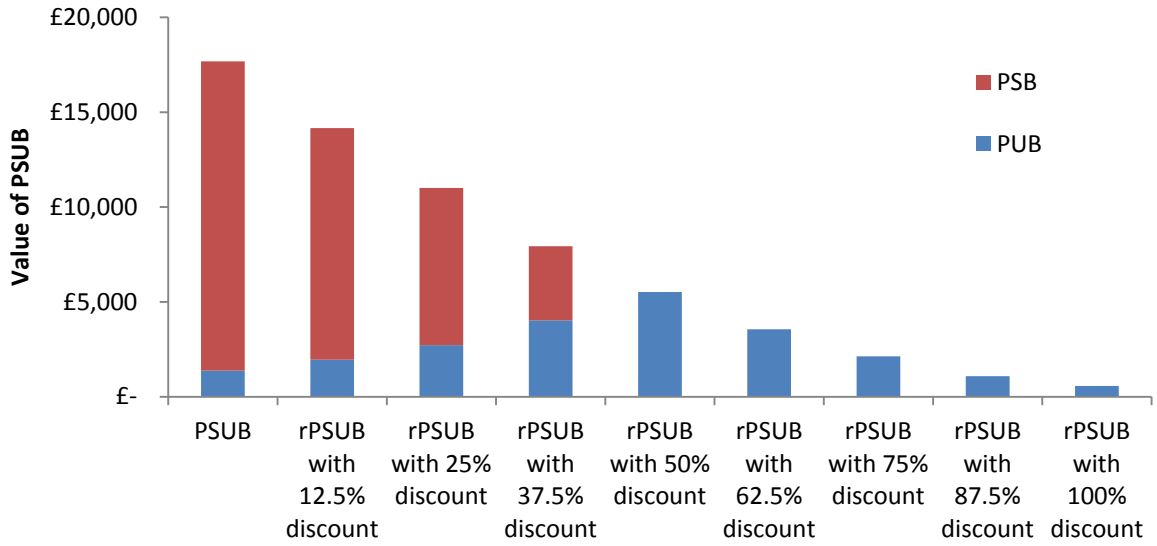
	Pazopanib	**Sunitinib**	**Interferon-α**	**Best supportive care**
Expected Costs (£)	36,144	36,414	8,376	4,098
Expected QALYs	2.03	1.90	1.25	0.99
ICER against interferon-α	35,686	42,788	-	NA
ICER against sunitinib	NA	-	NA	NA
Expected net benefit (£20k/QALY)	4,395	1,669	16,600	15,710

With the discount proposed by the manufacturer, the PUB at an acceptable ICER of £20,000 per QALY increased because it was now less certain that interferon-α was the most cost-effective strategy (Figure 55). Likewise, the P-SUB associated with recommending pazopanib was reduced from approximately £17,000 at list price to £14,000 with the discount (as shown in the second stacked bar in Figure 55). Because the P-SUB is composed of the PUB and the PSB, the fact that the P-SUB decreased while the PUB went up must be caused

by the PSB decreasing more than the PUB increased. This meant that while the discount increased uncertainty, it reduced the loss that was incurred in terms of net benefit forgone by choosing a strategy that was not the most cost-effective.

We found that a larger discount to the daily cost of pazopanib could potentially have a greater impact on the reduction of the P-SUB. The decision to recommend pazopanib at a maximum acceptable ICER of £20,000 per QALY was associated with a positive strategy-specific burden at no discount but this was resolved between a discount of 37.5% and 50% (Figure 55). The P-SUB of recommending pazopanib became increasingly smaller with increasing discounts (Figure 55) until pazopanib became the most cost-effective strategy and the P-SUB continued to decrease. The PUB increased up to that threshold discount. This effect followed from decision uncertainty for interferon-α increasing with an increasing discount, as it became less and less likely to be cost-effective.

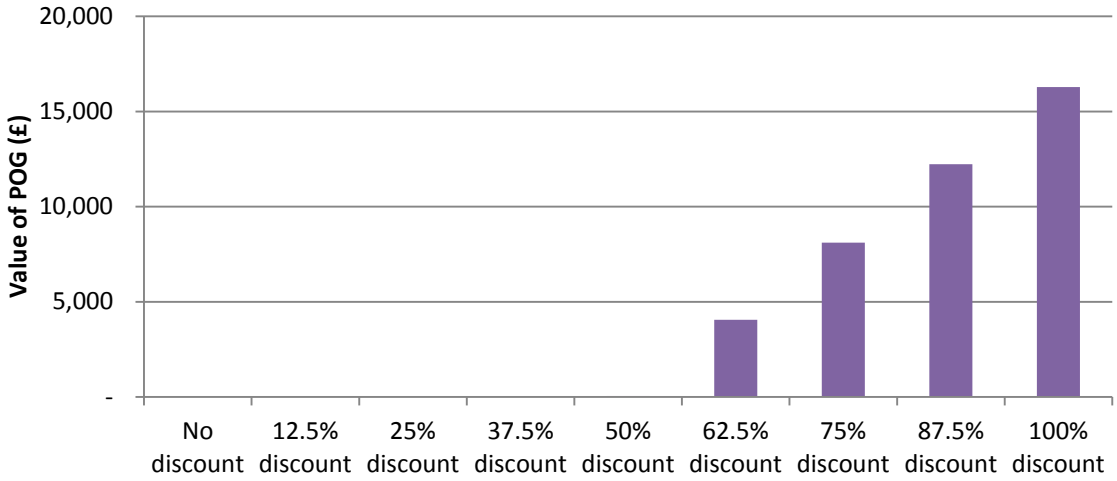
Figure 55: MEA risk reduction chart for different values of discounts in pazopanib appraisal without EoL valuation



With pazopanib becoming cost-effective at a discount of 50%, there was a payer optimality gain (POG) associated with this discount. This POG was the difference between the expected net benefit of the new intervention and the expected net benefit of the intervention that was previously cost-effective. At the maximum acceptable ICER of £20,000 per QALY, the previously cost-effective strategy was interferon-α. The POGs associated with the different

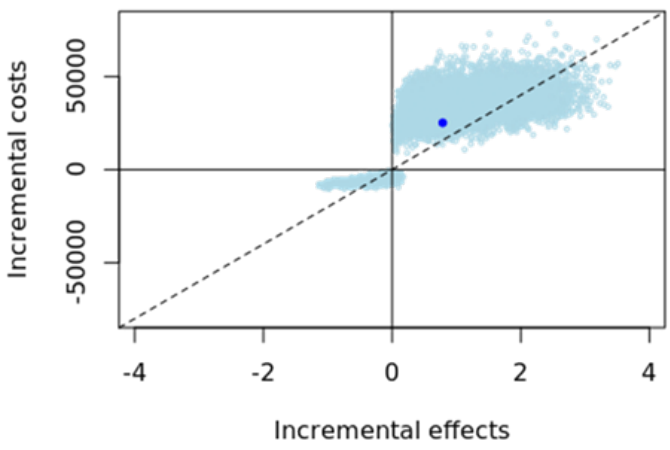
levels of discount are shown in Figure 56. The interpretation of that optimality gain is that the payer will incur savings with the new intervention that can be used for other purposes. The health displaced by funding an expensive intervention is therefore reduced.

Figure 56: The POG at different discounts in the pazopanib appraisal without EoL valuation



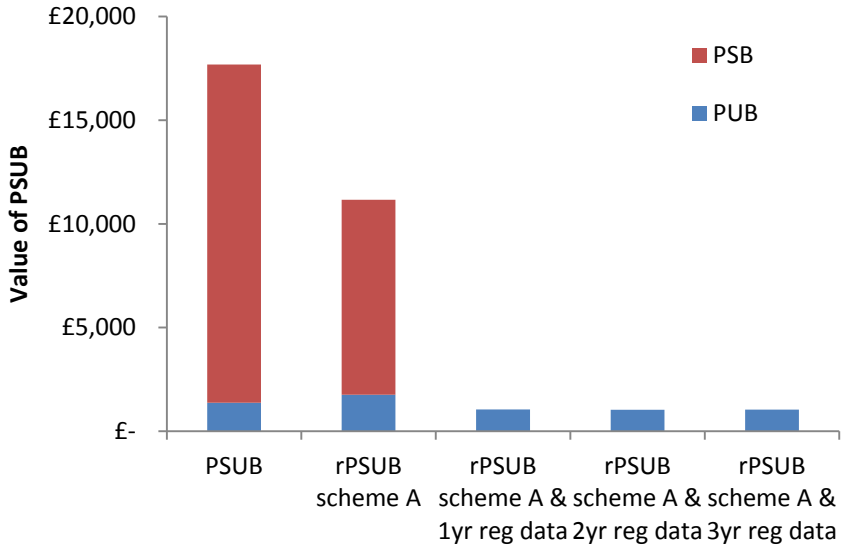
With scheme A in place, the expected ICER for pazopanib versus interferon-α dropped to £31,900 per QALY. Two centres of concentration in the cost-effectiveness plane now reflect the different prices associated with different survival gains (Figure 57), with the additional "bubble" reflecting the lowered price when survival gains were low. With scheme A, pazopanib was also more effective and less costly than sunitinib.

Figure 57: C-E plane for money back guarantee scheme for pazopanib vs interferon-α at £20,000 / QALY



With scheme A, the P-SUB of pazopanib dropped to £11,162 (Figure 58). This was mainly composed of the PSB (£9,400 versus the PUB of £1,762), reflecting that, while the PSB had decreased compared to the original composition of the P-SUB without any MEA in place, pazopanib was still not the most cost-effective strategy. Monitoring registry data for a year and re-visiting the decision based on that reduced the P-SUB further, eliminating the strategy-specific burden; assuming that the then cost-effective technology would be recommended at the future appraisal (Figure 58). Longer time spans of monitoring registry data only marginally reduced the P-SUB beyond that. If the decision was revisited at these time points, the PSB could be reduced to zero.

Figure 58: Risk analysis chart with scheme A and monitoring registry data in pazopanib appraisal at £20k / QALY



We also tested RwR schemes and to explore the effect of different sample sizes and follow up durations on the reduction of the PUB, we modelled four trials similar in design to the COMPARZ trial.

- Large and long trial: sample size of n = 550 in both trial arms and a follow-up period of 24 months.
- Large and shorter trial: n = 550 in both trial arms and a follow-up period of 6 months (this is the closest design to the COMPARZ trial).
- Medium sized trial: n = 300 in both trial arms and a follow-up period of 12 months.
- Small and shorter trial: n = 30 in both trial arms and a follow-up period of 6 months.

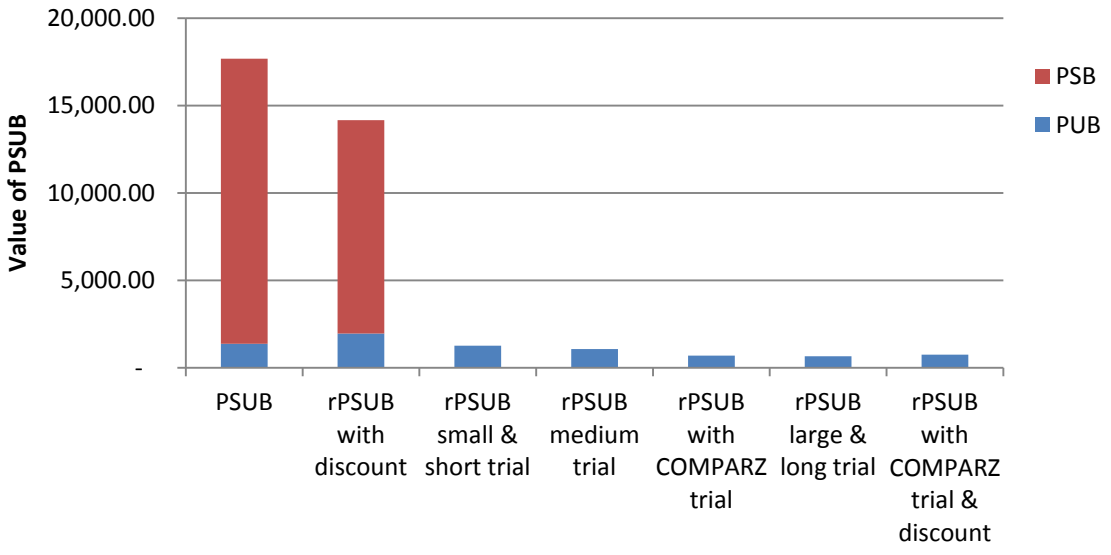
The value of these trial designs, traditionally known as the expected value of sample information (EVSI), is shown in Table 24. The hazard ratios for overall and progression-free survival of pazopanib against sunitinib make up for the greater part of the uncertainty in the model (66%) (Table 24). A large part of the uncertainty contributed by those parameters could be reduced by a trial similar to the COMPARZ trial. Indeed, the different trial designs would reduce the PUB significantly, by 53% for the trial that is closest to the COMPARZ trial, as shown in the last column. Smaller sample sizes reduced the value of the trial.

Table 24: Value of different trial designs in pazopanib hypothetical example (£ 20,000 / QALY)

Trial design	Sample size	Trial duration	Value of trial (£)	Indexed to PUB (%)
Uncertainty contributed by hazard ratios Paz vs Sun			**906**	**66**
Large and long trial	1100	24 months	723	53
Large and shorter trial	1100	6 months	679	49
Medium sized trial	600	12 months	306	22
Small and shorter trial	60	6 months	116	8

Under the assumption that the trial would report at the time of the appraisal, gathering new evidence reduced the P-SUB more than just the discount, with the condition that the decision was going to be revisited and the PSB eliminated by recommending the cost-effective strategy based on the new evidence (Figure 59).

Figure 59: MEA risk reduction chart in pazopanib appraisal without EoL valuation



Of course, new evidence does not usually report immediately, or else one might just take the new evidence into account straight away. A more likely scenario was that trial results were going to be reported two to three years after the appraisal, longer for some longer-term studies, and the value of further research would therefore only be accrued starting at the time at which the decision would be revisited.

It is therefore necessary to take the timing of the trial into account when recommending research. The annual risk measured by the P-SUB (that is discounted for future years up to the decision relevance horizon of 5 years) is now presented for each year in Figure 60. To enable comparison between the impact of a research scheme on the risk to the payer over future years and no research scheme, we present the annual risk for pazopanib with the discount (Part A of Figure 60), for pazopanib with the discount and the trial (Part B of Figure 60) and for interferon-α with the discount employed for pazopanib. Part A shows that the original per patient P-SUB associated with pazopanib simply decreases in future years, due to the discounting. In Part B, we observe that in the first two years up to the point of re-appraisal based on the new research evidence, the payer incurs the full P-SUB (that is discounted for future years) (Figure 60, Part B). Provided that the cost-effective strategy (based on the new evidence and price) is recommended at the time of re-appraisal, the residual PUB is the only risk to the payer in years 3, 4 and 5 (Figure 60, Part B). Part C of Figure 60 shows that the risk to the payer associated with interferon-α and a discount for pazopanib consists only of the PUB (discounted in future years).

Figure 60, Part B shows that performing the research reduces the risk compared to not doing any evidence collection (Figure 60, Part A). It is also obvious from Figure 60 that at the employed decision relevance horizon of five years, the sum of the stacked bars is smallest in Part C, despite the PUB not being reduced. This means that recommending pazopanib with the discount and the trial would be associated with great losses to the payer. These losses are made up of the difference between the lifetime PSB (accrued over the time horizon and the affected patient population) and the lifetime EVSI, which amounts to £42 million (Table 25, for re-appraisal after two years).

Figure 60: Annual risk over the pazopanib appraisal decision relevance horizon

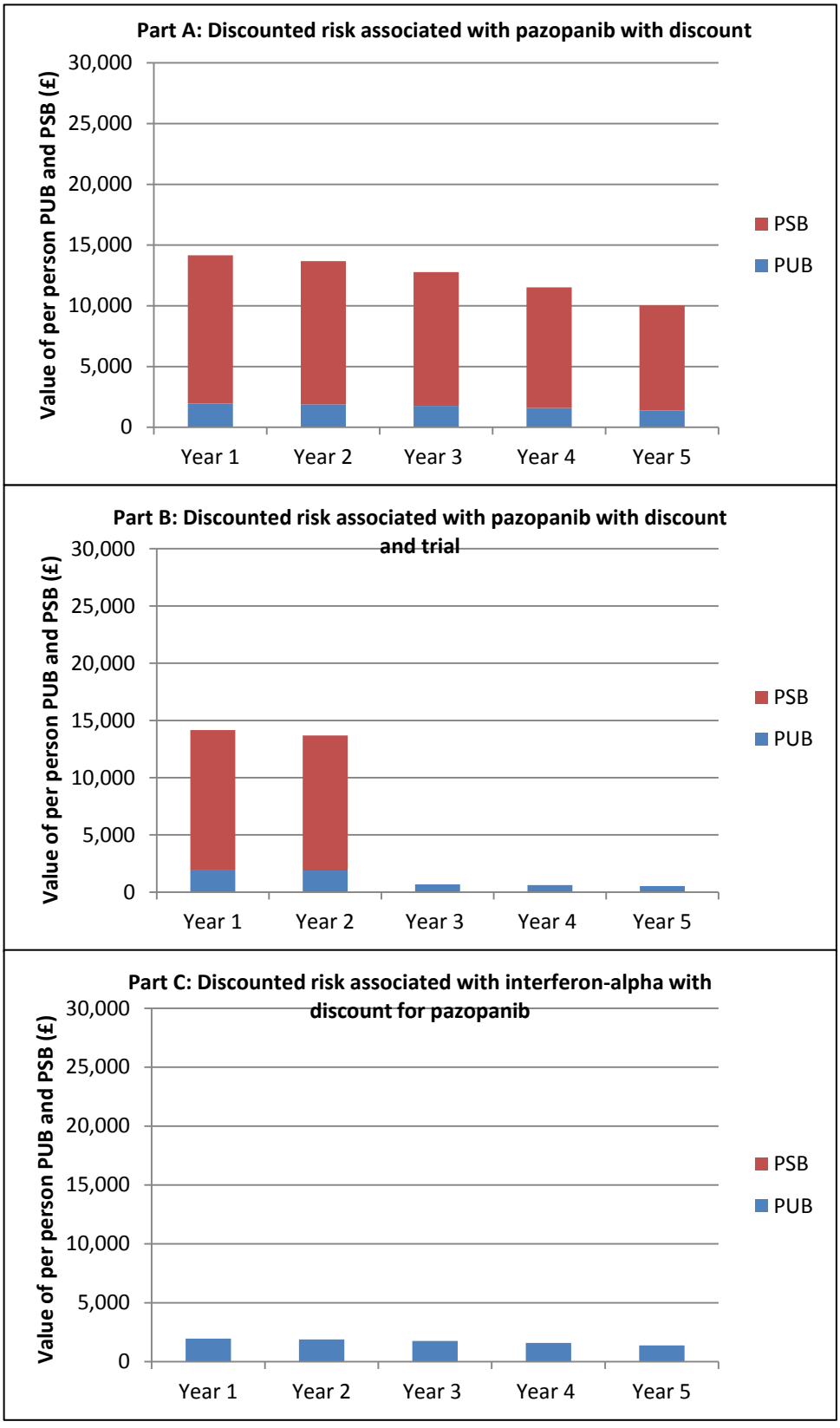


Table 25: Net EVSI for decision relevance horizon and affected population for pazopanib with discount and trial at different schedules for re-appraisal

Values accrued over 2,000 affected patients per year with decision relevance horizon of 5 years	Lifetime EVSI	Lifetime PSB	Lifetime net EVSI
Reappraisal after 1 year	8,190,948	24,410,860	- 16,219,912
Reappraisal after 2 years	5,856,823	47,996,232	- 42,139,409
Reappraisal after 3 years	3,677,891	70,013,429	- 66,335,538
Reappraisal after 4 years	1,712,620	89,871,679	- 88,159,059
Reappraisal after 5 years	0	107,176,997	- 107,176,997

These results indicate that when the PSB is very large compared to the PUB, a RwR decision without the addition of a price reduction does not result in complete risk reduction. The price reduction in this case would have to be substantial enough to bring the lifetime PSB down to a level at which it could be offset by the life-time value of the trial.

RwR schemes can be funded by the manufacturer, the payer or a third party stakeholder. If the COMPARZ trial had not already been ongoing at the time, NICE could have recommended publicly funded research. It would, however, be necessary to evaluate the expected net benefit of sampling (ENBS) as the cost of research has to be subtracted from its life-time value. Of course, in this case, recommending pazopanib with research was already associated with a risk, leading to a lifetime P-SUB of £60 million. If (a crude estimate of) trial costs of £2 million were added to this, the payer would incur a financial loss of £62 million over a five year time horizon when recommending pazopanib with the discount and research.

5.4. Concluding remarks on the application of the MEA risk Analysis Framework in the pazopanib appraisal

In conclusion, the application of the MEA risk analysis framework to the pazopanib appraisal and a hypothetical appraisal based on pazopanib but whithout EoL valuation has shown that:

1. When the PUB is large and results in a significant loss to the payer when calculated over the patient population, MEA schemes should be considered even when there is no strategy-specific burden.
2. In the pazopanib appraisal, the recommendation of research and the discount reduced the P-SUB by £14,800 per person; that is by £29.5 million annually over the patient population.
3. Price reduction schemes primarily reduce the PSB but they can reduce the PUB when the decision option for which the price is reduced is already expected to be cost-effective – they will, however, increase the PUB when it is not expected to be cost-effective.
4. If a RwR scheme is assessed, it is vital to consider the timing at which the trial reports and the overall decision relevance horizon.
5. The life-time value of the research study measured by the LEVSI has to be large enough to offset the life-time strategy burden (LPSB) incurred by recommending a strategy that was not expected to be cost-effective for the time until research reports.
6. If recommendation of a cost-ineffective technology with research is considered, the research scheme has to be combined with price reductions in the period until research reports in order to reduce losses to the payer.
7. Assessing the payer uncertainty burden and the payer strategy burden and presenting them in a risk analysis chart can give a simple overview of the risk associated with different decision options in a technology assessment.
8. Assessing MEA schemes requires more technical analysis. Price reduction schemes can easily be modelled but require the PSA to be performed again. Research schemes require modelling of the research study which can take a few days to conceptualise.

6. DISCUSSION OF FINDINGS AND ISSUES FOR IMPLEMENTATION

6.1. ISSUES FOR IMPLEMENTATION OF THE MEA RISK ANALYSIS FRAMEWORK

The methods that have been described in this report are extensions to Value of Information Methods that have been advocated for several years in the context of HTA. Those methods have struggled to achieve impact and some of those same barriers to implementation are worth addressing in the context of this new proposed framework.

In part it could be argued that the value of these methods has been limited in the past because there has been much less requirement for NICE to consider conditional approval. Guidance from Technology Appraisals has typically been based on evidence that was more certain, despite uncertainties present. Initiatives to speed up the process of licensing for pharmaceuticals will increasingly affect this.

Traditionally it has proved difficult to introduce new technically challenging methods to the Appraisals Committees. This is unsurprising because of the broad cross section of skills represented on those committees. However, in the current situation, it is likely that as the committees are increasingly faced with the appraisal of technologies that are potential candidates for managed entry agreements, they will increasingly recognise the need for the exact types of information that are readily presented in the MEA Risk Analysis Framework. Fundamentally, the types of information that are relevant to the assessment of MEAs, and are already set out in the Methods Guide, are quantitative in nature. Our contention is that, whilst the calculations and methods that underpin PUB and P-SUB are complex, the interpretation of results is straightforward and those results represent exactly the calculations that AC members will otherwise be left to perform in their heads. For the sake of accuracy and transparency the routine provision of these calculations should be an aid to decision makers, not an additional unwelcome complexity. There is an opportunity to provide appropriate information to decision makers, and requiring only minimal instruction on interpretation, before the gap between their needs and what is provided routinely in NICE TAs grows too great.

There are different elements of the framework that may be relevant in different situations and at different points in the process of a TA. First there is the quantification and understanding

of the nature of risk associated with the current evidence base and strategies. Second, there is the formal modelling of alternative candidate MEAs. The former is extremely simple to calculate and can be done in seconds using the SAVI tool. It would not be a substantial burden for those submitting or critiquing evidence (manufacturers or ERGs in the STA process, AGs in the MTA process) to provide estimates of PUB and P-SUB alongside their existing analyses, provided there is an appropriate probabilistic sensitivity analysis. An alternative that may be feasible is for these calculations to be performed in committee once there is agreement on the preferred set of parameter values for the base case analysis. Provided this is one of the scenarios that have already been performed, and the PSA results have been stored, the additional analysis can be undertaken "live".

The second stage of analysis is more involved and is unlikely to be an analysis that would be seen as routine. Rather, once the committee is convinced that the exploration of MEAs is relevant then such analyses could be undertaken. Where a manufacturer intends to propose an MEA from the outset then it may be feasible to present these analyses to the AC at the first opportunity but this would be complicated by not knowing the committee's preferred set of analyses. In the analyses we undertook, we found that modelling the value of a trial as part of an MEA was several days of analyst time.

Of course, all these analyses are contingent on the representation of uncertainty in the model parameters and propagated through the cost effectiveness results using PSA, truly reflecting the AC thinking. Often there are situations where structural uncertainty contributes the largest element of uncertainty in the view of the committee. It is important to establish the frequency with which this occurs, or where the representation is deficient for some other reason, and establish the most suitable analytical approach in those circumstances (e.g. model averaging, discrepancy approach etc.).

6.2. SUMMARY OF FINDINGS AND CONCLUSION

We conclude this report with a set of numbered findings and conclusions which summarise what has been achieved, proposed further work and recommendations for practice.

F1. We developed the MEA risk analysis framework that is designed to help make a transparent and informed choice on the appropriate use of MEA schemes.

F2. The MEA risk analysis framework consists of a taxonomy of the different MEA schemes and the MEA design guidance questionnaire that aid in the choice of potentially appropriate MEA schemes. Quantitative risk analysis using the P-SUB concepts enable to choose the MEA schemes that work best in reducing the risk associated with a decision in a health technology assessment setting.

F3. Risk associated with any strategy in a health technology assessment can be evaluated using the payer strategy burden (PSB) and the payer uncertainty burden (PUB). Together, these concepts form the payer strategy and uncertainty burden (P-SUB).

F4. The PSB quantifies the risk associated with each strategy that the payer, in this case the NHS commissioners, incurs.

F5. The PUB is the risk the payer incurs due to decision uncertainty, or the possibility of recommending the "wrong" strategy.

F6. We demonstrated the feasibility of the MEA risk analysis framework to be used as a potential routine add-on to analyses currently undertaken in the technology appraisal process.

F7. The MEA risk analysis framework is not dependent on the provision of QALYs as an effectiveness measure but can accommodate any outcome measure with which a maximum acceptable spending can be associated.

F8. Recommending MEA schemes can help reduce or eliminate the P-SUB that the payer would incur otherwise.

F9. There are two different dimensions of MEA schemes, effective reductions in price and recommendations with research. There are many different types of price reductions and research schemes that are summarised in the updated taxonomy.

F10. Effective price reduction schemes can be a good option when a technology would otherwise be cost-ineffective. In fact, when the PSB is large compared with the PUB, research will not be sufficient to make a strategy cost-effective, and reductions in price

would have to be considered. According to our findings, recommending technologies that are not cost-effective based on available evidence and price (as represented in the PSA) is associated with a risk of making the wrong decision and consequently a cost to the payer that could be greater than that of decision uncertainty alone (the PSB was one hundred times the PUB in the ulcerative colitis appraisal, for instance).

F11. Research schemes are especially appropriate when there is large decision uncertainty that can be resolved through further research. Price reductions can also reduce the PUB, when the strategy for which the price reduction will be put in place is expected to be cost-effective. When a technology is not expected to be cost-effective and price is reduced, decision uncertainty may increase but if that technology consequently becomes cost-effective, the PUB may reduce again and the PSB reduces to zero, thus reducing the overall cost to the payer. Price reductions and research recommendations can be complementary in reducing the P-SUB.

F12. When considering RwR schemes in an appraisal in which the intervention is not the most cost-effective strategy, it is important to consider the timing of research. This is because the PSB associated with that strategy will be incurred for as long as the new research evidence has not become available. One of our key findings is that it is possible to offset the PSB by conducting research on the key parameters causing decision uncertainty and revisiting the decision at a defined future time point if the future value of the trial outweighs the accrued PSB up to that point. It is important to note that, for this to be a feasible course of action, the decision-maker has to be prepared to reverse their decision if new evidence shows that the new intervention is not cost-effective.

F13. We also demonstrated that it can be desirable to conduct further research when a technology appeared to be cost-effective but had the caveat of large decision uncertainty. For instance, our analysis showed that instead of not recommending imatinib for the adjuvant treatment of GIST in 2010 and appraising it at a later stage, recommending it with a RwR scheme in place may have been the better option as there would have been the opportunity to direct the objectives of the research. We consequently think that MEA schemes should be considered routinely.

F14. When RwR decisions are considered, it is important to consider practical aspects associated with them, such as the reversibility of the decision and ethical issues that may prevent research from being conducted once the technology is widely available. We have not provided a detailed review of these guiding principles for the use of such schemes as that can be found elsewhere.

F15. Our work is foreseen to be of particular relevance in the currently changing pharmaceutical environment in which a greater number of submissions with an evidence base that is smaller or earlier in its evolution are expected.

F16. This work has shown that it is possible to systematically and routinely evaluate MEA schemes in technology appraisals. Such evaluations do not require more than the outputs from probabilistic sensitivity analyses (costs, effects and parameter values for each iteration of the PSA) which can be uploaded to online tools such as the Sheffield Accelerated Value of Information (SAVI) to identify the key drivers of uncertainty. Full EVPPI analysis using SAVI was shown to take only a couple of minutes; more when the number of parameters was large (e.g. for more than two hundred parameters, it could take thirty minutes). To assess RwR schemes, the planned trial would need to be modelled, which requires more time and computational effort.

F17. While the MEA risk analysis framework does not require much additional information from manufacturers and their submissions, the decision-maker may need to make arrangements to incorporate the framework in the appraisal process. This may entail making the discussion of MEA schemes a routine part of committee meetings. It may further be worthwhile to re-consider the decision-making body's ability to propose all types of MEA schemes, whether they be RwR or effective price reduction schemes.

F18. We have demonstrated the applicability of our framework to the real world by conducting eight case studies. Only one of those entailed the full evaluation of different MEA schemes and it may be desirable to repeat the process on other examples in order to explore a greater range of MEA schemes for their value. This was, however, out of the scope of this project.

F19. This work has only considered examples in which phase III trial evidence was available. When this is not the case, there may be other unknowns that can make assessments of this type more difficult. Incomplete decision models that either neglect large parts of the present uncertainty in their PSA or do not model the appropriate comparators, make it difficult to assess decision uncertainty correctly and evaluate MEA schemes (as seen in the DPI for cystic fibrosis appraisal). The assumption in our framework is that all uncertainty is accounted for in the PSA. Results derived by using our framework might therefore be misleading if applied to models in which the majority of parameters were assumed to be certain.

F20. Scope for further research thus lies in applying the framework to further real world appraisals to gain experience, explore the practical considerations associated with the different schemes and provide the opportunity to review our framework at the time of re-appraisal.

F21. Potential further research could assess the feasibility of modelling all the different MEA schemes and optimising the MEA recommendation such that the best value for the NHS is obtained.

F22. If well designed and used appropriately, MEA schemes should help the payer to optimise recommendations regarding new and existing technologies in a predictable, transparent and rational manner. The MEA risk analysis framework may thus be a step in the direction of more efficient decision-making.

F23. We conclude that assessments of the payer's risk of decisions taken under uncertainty and assessments of possible reductions in this risk through MEA schemes, as proposed in our MEA risk analysis framework, are both feasible and desirable.

7. REFERENCES

1. European Medicines Agency: Committee for Medicinal Products for Human Use. Guideline on the Scientific Application and the Practical Arrangements Necessary to Implement Commission Regulation (EC) No 507/2006 on the Conditional Marketing Authorisation for Medicinal Products for Human Use Falling Within the Scope of Regulation (EC) No 726/2004 (EMEA/509951/2006; London: CHMP, 5 December 2006). *London: European Medicines Agency* 2006.

2. European Medicines Agency. Adaptive Pathways. 2014. URL: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000601.jsp (Accessed January 2015).

3. Walker S, Sculpher M, Claxton K, Palmer S. Coverage with evidence development, only in research, risk sharing, or patient access scheme? A framework for coverage decisions. *Value in Health* 2012;15:570-9.

4. NICE. Guide to the methods of technology appraisal 2013. 2013. URL: http://www.nice.org.uk/article/pmg9/chapter/1-Introduction (Accessed May, 2015).

5. Klemp M, Frønsdal KB, Facey K. What principles should govern the use of managed entry agreements? *International journal of technology assessment in health care* 2011;27:77-83.

6. Garrison LP, Towse A, Briggs A, de Pouvourville G, Grueger J, Mohr PE, *et al.* Performance-based risk-sharing arrangements—good practices for design, implementation, and evaluation: report of the ISPOR good practices for performance-based risk-sharing arrangements task force. *Value in Health* 2013;16:703-19.

7. Carlson JJ, Gries KS, Yeung K, Sullivan SD, Garrison Jr LP. Current status and trends in performance-based risk-sharing arrangements between healthcare payers and medical product manufacturers. *Applied health economics and health policy* 2014;12:231-8.

8. Ferrario A, Kanavos P. Dealing with uncertainty and high prices of new medicines: A comparative analysis of the use of managed entry agreements in Belgium, England, the Netherlands and Sweden. *Social Science & Medicine* 2015;124:39-47.

9. Bruegger U. A review of Coverage with Evidence Development (CED) in different countries: What works and what doesn't. Manchester: NICE; 2014.

10. Walker S, Sculpher M, Claxton K, Palmer S. Coverage with evidence development, only in research, risk sharing or patient access scheme? A framework for coverage decisions. York: Centre for Health Economics, University of York; 2012.

11. Stafinski T, McCabe C, Menon D. Funding the unfundable: mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems (vol 28, pg 113, 2010). *Pharmacoeconomics* 2010;28:608-.

12. Edlin R, Hall P, Wallner K, McCabe C. Sharing Risk between Payer and Provider by Leasing Health Technologies: An Affordable and Effective Reimbursement Strategy for Innovative Technologies? *Value in Health* 2014;17:438-44.

13. Edlin R, Hall P, Wallner K, McCabe C. Leasing health technologies: an affordable and effective reimbursement strategy for innovative technologies? ; 2013. URL: medhealth.leeds.ac.uk/download/236/auhe_wp13_02 (Accessed January, 2015).

14. Franken MG, Gaultney JG, Blommestein HM, Huijgens PC, Sonneveld P, Redekop WK, *et al.* Policymaker, please consider your needs carefully: does outcomes research in relapsed or refractory multiple myeloma reduce policymaker uncertainty regarding value for money of bortezomib? *Value in Health* 2014;17:245-53.

15. Claxton K, Palmer S, Longworth L, Bojke L, Griffin S, McKenna C, *et al.* Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. 2012.

16. Drummond M. When do performance-based risk-sharing arrangements make sense? *Eur J Health Econ* 2015:569-71.

17. NICE. List of technologies with approved Patient Access Schemes. 2014. URL: http://www.nice.org.uk/about/what-we-do/patient-access-schemes-liaison-unit/list-of-technologies-with-approved-patient-access-schemes (Accessed January 2015).

18. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation: Oxford university press; 2006.

19. Strong M, Oakley J, Brennan A, Breeze P. Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample: A Fast Nonparametric Regression-Based Method. *Journal* 2015.

20. NICE. Cetuximab, bevacizumab and panitumumab for the treatment of metastatic colorectal cancer after first-line chemotherapy: Cetuximab (monotherapy or combination chemotherapy), bevacizumab (in combination with non-oxaliplatin chemotherapy) and panitumumab (monotherapy) for the treatment of metastatic colorectal cancer after first-line chemotherapy (review of technology appraisal 150 and part review of technology appraisal guidance 118). 2012. URL: http://www.nice.org.uk/guidance/ta242 (Accessed May, 2015).

21. NICE. Trabectedin for the treatment of advanced soft tissue sarcoma. 2010. URL: https://www.nice.org.uk/guidance/ta185 (Accessed May, 2015).

22. NICE. Lenalidomide for treating myelodysplastic syndromes associated with an isolated deletion 5q cytogenetic abnormality. 2014. URL: https://www.nice.org.uk/guidance/ta322 (Accessed May, 2015).

23. NICE. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262). 2015. URL: https://www.nice.org.uk/guidance/ta329 (Accessed May, 2015).

24. NICE. Imatinib for the adjuvant treatment of gastrointestinal stromal tumours (review of NICE technology appraisal guidance 196). 2014. URL: https://www.nice.org.uk/guidance/ta326 (Accessed May, 2015).

25. NICE. Colistimethate sodium and tobramycin dry powders for inhalation for treating pseudomonas lung infection in cystic fibrosis. 2013. URL: http://www.nice.org.uk/guidance/ta276 (Accessed May, 2015).

26. NICE. Pazopanib for the first-line treatment of advanced renal cell carcinoma 2011. URL: https://www.nice.org.uk/guidance/ta215 (Accessed April, 2015).

APPENDIX

A.1 Short description of concepts used in the MEA Analysis Framework, with mathematical definitions

Decision uncertainty is typically visualised using the *cost-effectiveness plane* and the *cost-effectiveness acceptability curve* (CEAC).[18] The expected cost of decision uncertainty is composed of the probability of taking the wrong decision based on current evidence and the cost associated with that wrong decision. The expected cost of decision making can be quantified numerically via *Expected Value of Information (EVI)* methods. This is because the *Expected Value of Perfect Information (EVPI)* can be interpreted as the value of eliminating the possibility of taking the wrong decision.[18]

Here, we recognise that by making a decision under uncertainty, we may make a choice that is associated with greater risk. In the HTA context the cost associated with this risky choice can either be expressed as monetary cost, or a cost in terms of a health or life foregone. We do not know whether we are making an optimum choice (if we did, then there would be no uncertainty), but through modelling we can express the probability that any choice is optimum. The expected cost of the decision, due to the uncertainty, is known as the expected opportunity loss, or (equivalently) the EVPI.

Although the terms expected opportunity loss and EVPI already exist, we introduce a new term for the same quantity: the *Payer Uncertainty Burden (PUB).*Our rationale is that we wish to make it clear that the payer incurs an *expected* cost each time a decision is made under uncertainty, compared to a scenario in which there is no uncertainty. Another way of thinking about this is to imagine making a large number of decisions under uncertainty. Some of the choices made will be "wrong", leading to losses. The *Payer Uncertainty Burden (PUB)* is the expected value of eliminating these losses. The PUB is mathematically the same as both the Expected Value of Perfect Information, and the Expected Opportunity Loss.

The PUB is given by:

Equation 1:

$$PUB = EVPI = \left[\mathbb{E}_{\theta}\left\{\max_{d} \quad (\mathrm{NB}(\mathrm{d},\theta)\right\} - \max_{\mathrm{d}} \quad \mathbb{E}_{\theta}\{\mathrm{NB}(\mathrm{d},\theta)\}\right] \geq 0,$$

where NB is the net benefit function, d indexes decision options in some set *D*, and θ is a vector of uncertain model parameters.

Standard decision theory assumes that we will always seek to maximize expected utility (in HTA the utility associated with a decision option is usually considered to be the net benefit i.e. health outputs multiplied by the value of one unit of health, minus costs). However, there are some circumstances in which it may be perfectly reasonable to adopt a decision option that does not have the greatest expected net benefit according to our cost-effectiveness model. Under a decision *not* to adopt the optimum option, we pay a penalty (the difference between the expected net benefit of the optimum option, and the expected net benefit of our chosen cost-ineffective option). In this report we will refer to this difference as the *Payer Strategy Burden (PSB).*

The PSB is given by:

Equation 2:

$$PSB(d') = \left[\max_{d} \mathbb{E}_{\theta}\{\mathrm{NB}(\mathrm{d},\theta)\} - \mathbb{E}_{\theta}\{\mathrm{NB}(\mathrm{d}',\theta)\}\right] > 0,$$

where d' is a cost-ineffective decision option.

Similar to the loss we incur when choosing a cost-ineffective strategy (the PSB), we can accrue a gain when we switch from a cost-ineffective choice to the most cost-effective strategy, which we call the *Payer Optimality Gain (POG).* This POG is the difference between the expected net benefit of the new intervention and the expected net benefit of the intervention that was previously cost-effective. The interpretation of that optimality gain is that the payer will accrue savings with the new intervention that can be used for other purposes.

Now, if we imagine making a decision to recommend a strategy that is not expected to be cost-effective under uncertainty, we will incur two costs: the strategy burden, and (compared to a scenario in which uncertainty is eliminated) the decision uncertainty burden. We denote the sum of these as the *Payer Strategy and Uncertainty Burden (P-SUB).* Each of these quantities (PUB, PSB, P-SUB) can either be expressed in monetary units, or in health output units (for example, life years or QALYS).

The P-SUB is given by:

Equation 3:

$$PSUB(d') = PUB + PSB(d') = \left[\mathbb{E}_\theta\left\{\max_d(\mathrm{NB(d,\theta)}\right\} - \mathbb{E}_\theta\{\mathrm{NB(d',\theta)}\}\right] > PUB$$

Because the P-SUB of a cost-ineffective strategy is defined by the sum of the PUB and the PSB, the PUB equals the P-SUB of the optimal decision option where the PSB is zero. If the population size is known, the annual population PUB, PSB or P-SUB can be calculated and knowledge of the decision time horizon would enable calculating their lifetime values.

The EVPI or PUB is easily obtained using the Sheffield Accelerated Value of Information (SAVI) online tool (which can be accessed here: http://savi.shef.ac.uk/SAVI/), which automatically calculates it if parameters, costs and effects of the probabilistic sensitivity analysis are uploaded.

We furthermore might want to assess the value of a MEA in terms of its reduction of the P-SUB. For this, we introduce the residual P-SUB that remains when a MEA is adopted. The residual P-SUB will be calculated differently for price reduction and RwR schemes. For price reduction schemes, it will be necessary to perform another probability sensitivity analysis (PSA) (varying all uncertain parameters in the model and recording results for a large number of iterations) with the price rule in place and the P-SUB resulting from those PSA results will be the residual P-SUB with that price reduction MEA.

If net benefit analysis and the residual P-SUB indicate that the previously cost-ineffective decision is now optimal and that the residual P-SUB has reduced to a level that seems acceptable to the decision-maker, the analysis could be stopped here. In many cases however, the residual P-SUB may still be large due to uncertainties present in the model. In some cases, these can partly be resolved with research.

To assess RwR schemes, we first of all need to perform *partial EVPI (EVPPI)* analysis in order to identify the parameters contributing the most decision uncertainty. The EVPPI should give an indication of which parameters cause the most decision uncertainty and by how much this could potentially be reduced if all of the decision uncertainty surrounding that (group of) parameter(s) could be resolved. The EVPPI is easily calculated with SAVI that

generates EVPPI results for all individual parameters as well as groups of parameters defined by the user.

Assessing the value of a possible trial design in terms of its potential PUB reduction requires the use of another established concept, the *Expected Value of Sample Information (EVSI).*[18] The EVSI quantifies by how much a specific trial can reduce the PUB. It is easily calculated with SAVI when a simulation of resulting parameter values for the trial is available.[19] The EVSI is then simply the EVPPI of the parameter(s) modeled in the trial. The PUB is then reduced by the EVSI, yielding the residual PUB. If the strategy is still not expected to be cost-effective, the P-SUB is composed of the PUB and the new PSB; but in most cases we would expect the decision-maker to take recommend the optimum strategy.

The residual PUB for a RwR scheme is given by:

Equation 4:

$$rPUB = PUB - EVSI$$

To assess synergies between a price reduction and a RwR scheme, the above process for obtaining the residual P-SUB needs to be repeated on the results of the PSA with the price reduction scheme in place.

The residual P-SUB of a price reduction scheme may help in deciding which MEA scheme to choose if a previously cost-ineffective strategy is reversed to become a cost-effective strategy. If the strategy remains cost-ineffective despite the price reduction in place, or if the only MEA considered is a RwR scheme, the residual P-SUB only does not provide sufficient information as to which decision to take.

The question then really is whether it may be worthwhile recommending a strategy that is not expected to be cost-effective at present with the prospect of further research being conducted that may support that decision. This adds a temporal dimension and our proposed framework requires that the lifetime value of the trial in the future be greater than the PSB up to the point at which the trial reports.

Or expressed differently, this means that the value of the trial that will be accrued in the future needs to be greater than the losses we incur up to the point at which the trial reports.

The condition at which recommending a cost-ineffective strategy can yield more value in the future:

Equation 5:

$$\sum_{t=0}^{t^*-1} \frac{PSB(d')}{(1+\rho)^t} < \sum_{t=t^*}^{T} \frac{EVSI}{(1+\rho)^t} \ ,$$

where t refers to periods of time, t^* is the review time at which the trial reports, T is the overall time horizon for the decision and ρ is the discount rate.

Using the terms before, the net EVSI (that is the lifetime EVSI less the lifetime PSB) has to be greater than zero to outweigh the strategy-specific burden with research. Only if this is true may it be worthwhile recommending a cost-ineffective strategy, that has not become optimal through price reductions, under the condition of a RwR MEA. The net EVSI has to be greater than 0 in order for the value of research to offset the PSB up to the point at which the trial reports and is given by:

Equation 6:

$$net.EVSI = \sum_{t=t^*}^{T} \frac{EVSI}{(1+\rho)^t} - \sum_{t=0}^{t^*-1} \frac{PSB(d')}{(1+\rho)^t} > 0$$

A.2 End of life framework

End of life valuation was included in the analysis by developing a formal framework. This framework uses the median incremental survival gains of all comparators in order to determine the correct maximum acceptable ICER for each simulation of the PSA. To incorporate this in the EVI analysis, calculations were performed in R.

The incremental median survival gains of pazopanib, sunitinib and interferon-α were obtained through subtracting the median survival times of the different comparators in each

simulation. This led to three possible comparisons $d \in D$, for each of which the incremental median survival time was calculated using (for comparators *i* and *j*).

Equation 7:

$$\Delta t_d{}' = t_i{}' - t_j{}'$$

To reflect the situation at the time of appraisal, we assumed that both best supportive care and interferon-α only ever be valued at £20,000 per QALY. With the knowledge of the incremental survival gains of sunitinib and pazopanib against each other and against interferon-α, we could then calculate the net benefits with the appropriate maximum acceptable ICER for each simulation.

To illustrate these calculations, we present one example scenario. For each of the comparators, there were eight possible positions in the ranking of survival gain leading to 24 possible outcomes (see). This results from three possible positions for three comparators (3x2 possible outcomes) multiplied by 2 possible relationships between 2 pairs (2x2) (more than 3 months more or less than 3 months more) leading to 24 possible outcomes.

One of them may be that pazopanib exhibits a slightly larger survival gain against sunitinib (but not to the extent of 3 months) and sunitinib exhibits a survival gain of more than three months against interferon-α. This can be written as: P > S >> I, where > stands for greater survival gain but not greater than three months and >> stands for a survival gain greater than three months.

We then want to calculate the net benefit for sunitinib and reflect that it fulfils end of life criteria (the net benefit for interferon-α was assumed to be evaluated at £20,000 per QALY every time regardless of its position in the survival gain hierarchy). The net benefit for sunitinib is thus the sum of the net benefit of interferon-α valued at £20,000 per QALY and the incremental QALYs and costs of sunitinib versus interferon-α valued at the end of life weight of £50,000 per QALY.

Equation 8:

$$NB_S = Q_I\lambda_{20} - C_I + (Q_S - Q_I)\lambda_{50} - (C_S - C_I)$$

This equation reflects that only those QALYs are valued at the end of life threshold that are provided additional to the existing QALYs obtained from the incumbent technology.

In the calculation of the net benefit of pazopanib we now want to reflect that pazopanib provides a survival gain greater than sunitinib but not greater than three months. Because pazopanib implicitly provides a survival gain of more than three months against interferon-α, we need to reflect the net benefit of sunitinib and add to it the incremental costs and QALYs of pazopanib versus sunitinib valued at £20,000 per QALY.

Equation 9:

$$NB_P = Q_I\lambda_{20} - C_I + (Q_S - Q_I)\lambda_{50} - (C_S - C_I) + (Q_P - Q_S)\lambda_{20} - (C_P - C_S)$$

Both of the above equations can be simplified to yield:

Equation 10:

$$NB_S = Q_S\lambda_{50} - Q_I\lambda_{30} - C_S$$

Equation 11:

$$NB_P = Q_P\lambda_{20} + Q_S\lambda_{30} - Q_I\lambda_{30} - C_P$$

Once simplified, there are four different net benefit calculations over the 24 cases for each of the two comparators (see Table A 1). The choice of the net benefit calculation was then made specific to each PSA row depending on the survival gain in that row.

Table A 1: EoL net benefit calculations

Survival outcomes	NB_P	NB_S	NB_I	NB_B
$P \gg S > I$	$Q_P\lambda_{50} - Q_S\lambda_{30} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$P \gg S \gg I$	$Q_P\lambda_{50} - Q_I\lambda_{30} - C_P$	$Q_S\lambda_{50} - Q_I\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$P > S \gg I$	$Q_P\lambda_{20} + Q_S\lambda_{30} - Q_I\lambda_{30} - C_P$	$Q_S\lambda_{50} - Q_I\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$S \gg P > I$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{50} - Q_P\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$S \gg P \gg I$	$Q_P\lambda_{50} - Q_I\lambda_{30} - C_P$	$Q_S\lambda_{50} - Q_I\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$S > P \gg I$	$Q_P\lambda_{50} - Q_I\lambda_{30} - C_P$	$Q_S\lambda_{20} + Q_P\lambda_{30} - Q_I\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$I \gg S > P$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$I \gg S \gg P$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{50} - Q_P\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$I > S \gg P$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{50} - Q_P\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$I \gg P > S$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$I \gg P \gg S$	$Q_P\lambda_{50} - Q_S\lambda_{30} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$I > P \gg S$	$Q_P\lambda_{50} - Q_S\lambda_{30} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$S \gg I > P$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{50} - Q_P\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$S \gg I \gg P$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{50} - Q_P\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$S > I \gg P$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{50} - Q_P\lambda_{30} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$P \gg I > S$	$Q_P\lambda_{50} - Q_S\lambda_{30} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$P \gg I \gg S$	$Q_P\lambda_{50} - Q_S\lambda_{30} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$P > I \gg S$	$Q_P\lambda_{50} - Q_S\lambda_{30} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$P > I > S$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$S > I > P$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$I > P > S$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$P > S > I$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$I > S > P$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$
$S > P > I$	$Q_P\lambda_{20} - C_P$	$Q_S\lambda_{20} - C_S$	$Q_I\lambda_{20} - C_I$	$Q_B\lambda_{20} - C_B$

≫ means more than 3 months survival gain
> means greater survival gain but not more than 3 months greater

A.3 Calculations for MEA scheme A used in pazopanib appraisal

With overall survival causing the most uncertainty, it made sense to link price with the median survival time. Because at a maximum acceptable ICER of £20,000 per QALY the most cost-effective strategy was interferon-α, we proposed an alternative MEA that reduced the reimbursed price to zero in patients who do not survive as long as they could have while using interferon-α (scheme A. Money back guarantee).

To assess this MEA scheme within the present model, we ran a PSA where the price (daily cost) of pazopanib in each simulation was linked with the median survival time following from the hazard ratio of that simulation. The median survival time of pazopanib was obtained using

Equation 12:

$$t_i' = \left[-\frac{\log\left(S(t)^{\frac{1}{HR_i}}\right)}{\lambda} \right]^{\frac{1}{\gamma}}$$

with S(t)=0.5, and HR_i=1 for i=interferon-α, and HR_i assuming different values for each simulation for the other comparators, and λ and γ being constant Weibull parameters of the interferon-α survivor function.

The price, in this case the daily cost, of pazopanib was then made a function of t_P', $P = f(t_P')$. We decided this function should take the form of the following in order to reflect that the reimbursed price of pazopanib would reduce to zero when the survival gain provided fell below that associated with interferon-α, and the full price would be reimbursed in case that survival gain was larger than that of interferon-α.

Equation 13:

$$Price_P = \begin{cases} 0, & if\ t_P' < t_I' \\ P_P^f, & if\ t_P' \geq t_I' \end{cases}$$

With the obtained results for the PSA, the P-SUB with this price rule in place could be calculated and compared with the other MEA options.