THE UNIVERSITY of York

A Pilot Study of Value of Information Analysis to Support Research Recommendations for NICE

CHE Research Paper 4

A Pilot Study of Value of Information Analysis to Support Research Recommendations for the National Institute for Health and Clinical Excellence

Karl Claxton*
Simon Eggington**
Laura Ginnelly*
Susan Griffin*
Christopher McCabe**
Zoe Philips***
Paul Tappenden**
Alan Wailoo**

*Centre for Health Economics, University of York
**ScHARR, University of Sheffield
***Department of Economics, University of Nottingham

June 2005

Background

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The new CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Disclaimer

Papers published in the CHE Research Paper (RP) series are intended as a contribution to current research. Work and ideas reported in RPs may not always represent the final position and as such may sometimes need to be treated as work in progress. The material and views expressed in RPs are solely those of the authors and should not be interpreted as representing the collective views of CHE research staff or their research funders.

Further copies

Copies of this paper are freely available to download from the CHE website www.york.ac.uk/inst/che/pubs. Access to downloaded material is provided on the understanding that it is intended for personal use. Copies of downloaded papers may be distributed to third-parties subject to the proviso that the CHE publication source is properly acknowledged and that such distribution is not subject to any payment.

Printed copies are available on request at a charge of £5.00 per copy. Please contact the CHE Publications Office, email che-pub@york.ac.uk, telephone 01904 321458 for further details.

Centre for Health Economics
Alcuin College
University of York
York, UK
www.york.ac.uk/inst/che

Contents

	Page
Executive summary	iii
Chapter 1: Introduction and overview of methods	1
Chapter 2: The cost-effectiveness and value of information associated with repeat screening for age related macular degeneration	12
Chapter 3: Glycoprotein IIb/IIIa antagonists for the treatment of acute coronary syndromes	26
Chapter 4: Clopidogrel and modified-release dipyridamole in the secondary prevention of occlusive vascular events	41
Chapter 5: The cost-effectiveness and value of information associated with treatment for influenza in otherwise healthy adults	58
Chapter 6: Liquid-based cytology for cervical screening	74
Chapter 7: Disease-modifying therapies for the management of multiple Sclerosis	87
Chapter 8: Implications, methodological issues and challenges	100
Chapter 9: Implementing value of information analysis in the NICE research and development process	107
Chapter 10: Conclusions	113

Executive summary

Background

This project developed as a result of the activities of the Research Teams at the Centre for Health Economics, University of York, and ScHARR at the University of Sheffield in the methods and application of decision analysis and value of information analysis as a means of informing the research recommendations made by NICE, as part of its Guidance to the NHS in England and Wales, and informing the deliberations of the NICE Research and Development Committee.

Bayesian decision analysis and value of information analysis (DA-VOI) provides a methodological framework which explicitly considers the uncertainty surrounding the decision of a health care system to adopt a health technology. Specifically, using existing evidence, these methods focus on the likelihood of making a wrong decision if the technology is adopted. The value of additional research is based on the extent to which further information will reduce this decision uncertainty. This framework values the additional information, which may be generated by further research, in a way which is consistent with the objectives and the resource constraints of heath care provision (the cost-effectiveness threshold). This allows a comparison of the potential benefits of further research with the costs of further investigation, a comparison and prioritisation of alternative research recommendations, both within and between Technology Assessments, as well as an assessment of the value of investing resources in research or other activities, such as the provision of health service. In this sense it provides a unified and coherent framework for prioritisation of research and the use of heath care technologies.

Objectives

The specific objectives of the pilot study were to:

- Demonstrate the benefits of using appropriate decision analytic methods and value of information analysis to inform research recommendations.
- Establish the feasibility and resource implications of applying these methods in a timely way, to inform NICE.
- Identify critical issues and methodological challenges to the use of value of information methods for research recommendations (with particular regard to the new reference case as a suitable basis for this type of analysis).

The project consists of a series of case studies based on recent technology assessment reports completed by the York and Sheffield group for NICE. These included:

- Screening for age related macular degeneration (AMD)
- Glycoprotein IIb/IIIa antagonists for acute coronary syndrome (GPAs)
- Clopidogrel and dipyridamole in the secondary prevention of occlusive vascular events (CLO)
- Neurominidase inhibitors for the treatment of influenza (NIs)
- Liquid based cytology screening for cervical cancer (LBC)
- Beta interferon and glatiramer acetate in the management of MS (MS)

The purpose was to establish the feasibility and requirements of value of information analysis once submissions and Technology Assessment Reports (TARs) are conducted within the reference case specified in the recent methods guidance. Therefore case studies were selected on the basis that the existing TAR comes as close to the new reference case analysis as possible

Results

With the exception of screening for AMD all the other case studies met the original selection criteria for inclusion. Screening for AMD was not included in the original TAR for AMD. However, it was a recommendation for further research and the analysis for this case study is based on a screening model for AMD which was developed for the NCCHTA. The other five associated TARs included an

appropriate decision analytic model probabilistic analysis. In each case the existing TAR came close to the new reference case. However, any specific shortcomings are highlighted in each case study chapter. The more general issue of whether the existing references case is a sufficient basis for DA-VOI is discussed in chapter 8 and suggests that with the ongoing development of more detailed methods guidance on modelling, probabilistic analysis and evidence synthesis, a well conducted reference case analysis will provide a sufficient basis for value of information analysis. The core tasks and initial reanalysis of the case studies were completed in a timely way and within the proposed timeframe (4 weeks). It is anticipated that the additional resources required to move from a well conducted reference case analysis to full value of information analysis will be less than required within this pilot study which was based on pre reference case assessment reports.

The decision uncertainty surrounding the choice between strategies was characterised in the form of cost-effectiveness acceptability curves and frontiers. In each case the decision model was reanalysed and value of information analysis conducted. The Expected Value of Perfect Information (EVPI) surrounding each decision problem for the population of England and Wales, and the EVPI associated with particular model inputs was established using appropriate non-parametric methods

The value of research differed substantially across the 6 technology appraisals (EVPI ranged from £2.8m to £865m). In some cases the analysis indicated that the original research recommendations should not be regarded as a priority, e.g., the EVPI surrounding LBC for was low (£2.8m). In other cases it indicated that additional research should be commissioned, e.g., the EVPI surrounding CLO for stroke patients was high (£865).

The analysis indicated which comparators should be included in future research and also suggested other parameters that could be excluded. Estimates of value of information for the decision problem and for groups of parameters were also presented for relevant patient sub groups, e.g., the value of information different across the patient groups considered in the CLO and AMD case study. A number of case studies presented scenarios to explore alternative views of the relevant evidence e.g., inclusion of related and “unrelated events” in the assessment of CLO and impact of restricting consideration of evidence at 6 months in GPAs ; different structural assumptions regard mechanism of action, e.g., additive nature of information gains during screening for AMD; as well as the impact on value of information when relevant alternative may have been excluded from the original scope of the appraisal, e.g., the including the potential role of clopidigrel in the GPA case study.

The implications for the value of research in each of the areas were presented at a general level, as well as for the design of any future research in terms of features such as the relevant patient groups and comparators, and whether experimental design was likely to be required. The full reporting of the analysis conducted, and a discussion of the results for each of the case studies, can be found in Chapters 2 to 7. The case studies also highlighted a number of more general methodological issues including: consideration of all comparators, synthesising direct and indirect evidence, and considering structural as well as parameter uncertainty which are more fully discussed in chapter 8.

Conclusions and recommendations

Demonstration of benefits

The framework proved by DA-VOI was successfully implemented for each of the 6 case studies and provides the value of additional information, which may be generated by further research. This is consistent with the objectives of the heath care system (maximise the health gains of the population of England and Wales) and is based on the same resource constraints (the cost-effectiveness threshold which is used to develop guidance on use of the technology).

- For a particular assessment, this allows comparison of the potential benefits of further research (EVPI) with the costs of further investigation. If the potential benefits exceed the additional costs (including opportunity costs to patients) then further investigation maybe required to support guidance on use. The EVPI associated with the groups of parameters indicates the type of evidence that will be most valuable and therefore the type of studies that should be recommended.
- It also allows comparisons to be made across different technology assessments and prioritisation between alternative research recommendations, as well as a comparison between the value of investing resources in research or other activities such as the provision

of health service. In this sense it provides a unified and coherent framework for prioritisation of both research and the use of heath care technologies

The reference case as a sufficient basis for VOI

The results of any analyses are conditional on the use of appropriate model structure, appropriate synthesis of evidence and characterisation of other uncertainties. This is important for estimates of expected cost-effectiveness but even more important for estimates of value of information which are particularly sensitive to these issues. The existing reference case and methodological guidance requires supplementary guidance on the detailed use of methods to ensure that the adequacy of reference case submission can be assessed and that an adequate reference case analysis will provide sufficient basis for value of information analysis. This process of developing detailed methodological guidance is underway, coordinated by the National Decision Support Unit. It is important that in identifying and recommending methods the full characterisation of decision uncertainty should be a primary concern.

Critical issues and methodological challenges

It should be recognised that the key challenges for this type of analysis are not the VOI methods themselves but structuring decision problems, synthesis of evidence and the characterisation of uncertainty (required for estimating costs and effects as well as VOI). The development of methods in these areas is ongoing and will require continued support from a variety of sources. Particular issues, many of which have been highlighted in the case studies, include:

- Ensuring a sufficiently wide scope for the assessment to include all the relevant alternative strategies. This includes other technologies as well as different clinical policies (start and stop criteria) for a single technology. The exclusion of alternative strategies may not change the overall guidance on use of a technology but in some cases it may have a substantial impact on the value of information and on research recommendations.
- Dealing simultaneously with heterogeneity (variability by observed patient characteristics), variability (variability by unobserved characteristic) and uncertainty.
- Reflecting the additional uncertainty due to potential biases in the evidence, which may come from different types of study and/or suffer from publication bias.
- Modelling the exchangeability of the evidence with the parameters required in the model and reflecting any additional uncertainty.
- The inclusion or exclusion of unrelated events from the evidence and the potential role of prior elicitation from “experts”.
- The potential role of using priors elicited from “experts” within the NICE process and appropriate methods of elicitation of priors.
- Exploring and reflecting the additional uncertainty surrounding alternative but credible structural assumptions.
- Establishing efficient methods of searching for evidence on all model parameters not simply those associated with measures of effect.
- The synthesis of both direct and indirect evidence for measures of effect but also for other model parameters.

Many of these issues are being addressed though various programmes of research around the UK. However these areas of research require further development and continued support. In additional it should be recognised that these issues require multidisciplinary working with collaboration across many different centres. Infrastructure support to facilitate full collaboration across these areas of work should be sort from a variety of sources. As all these methods evolve, it should be recognised that the detail of what is required within an adequate reference case analysis will also develop over time.

Other issues specific to VOI include:

- Estimating the effective population that may benefits from additional evidence, including estimating time horizons for different technologies and incorporating this uncertainty in the estimates of value of information
- Estimating the value of information for correlated parameters
- Estimating the overall value of information based on estimates of the value of information for patient subgroups
- Presenting the value of information and the value of full implementation of guidance on use with in the same framework of analysis

Again work is currently ongoing on all of these issues but continued support from a number of sources for this methods work as well as support for an infrastructure of collaboration is needed.

Feasibility and resource implications

The pilot demonstrates that VOI is feasible within reasonable time lines, even based on pre reference case analysis. The use of VOI as part of the reference case (taking account of the recommendations made above) is for most types of models limited, not by time and resource requirements, but by the capacity to conduct this type of analysis and the dissemination of these methods. Therefore training in VOI methods should be considered as a cost-effective means of easing these capacity constraints.

However, complex and computationally expensive models (patient level simulations) make probabilistic analysis and therefore VOI potential very time and resource intensive. There are therefore 2 issues that should be addressed:

- It should be recognised that using patient level simulation is very costly in the sense that it may prevent reliable estimates of cost-and effect, and decision uncertainty as well as VOI being presented. In these circumstances it should be avoided if possible (by use of alternative structures and programming techniques). More work is required to establish those circumstances were the use of patient level simulation unavoidable.
- Where patient level simulation is required then there are techniques available to solve computationally expensive models, characterise uncertainty and estimate VOI. Indeed these have been used in NICE submissions. Further work is required to pilot their feasibility when patient level simulation is unavoidable and dissemination of appropriate methods.

Implementation to inform research recommendations

There are a range of possible options to implement DA-VOI within the NICE process, to inform research recommendations. These are more fully discussed in chapter 9. In this chapter we avoid making recommendations for implementation but outline possible options with some assessment of their strengths and weaknesses for NICE to consider. In general there a two levels at which DA-VOI could be implemented:

- DA-VOI could be implemented at the TAR stage of the process, either selectively or ultimately becoming part of the reference case for the Assessment Report. This would mean that the analysis would be available to inform the research recommendations made by the Appraisals Committee which generally to date have not been based on any formal analytic framework or evidence. The advantage of this would be that the decisions about the use of a technology and the evidence required to support the guidance could be appropriately considered at the same time.
- Alternatively DA-VOI could be implemented in a similar way to the case studies presented here: as a supplementary analysis to and existing TAR once guidance on use and research recommendations have been made. This would then provide an analysis that could inform the deliberations of the NICE Research and Development Committee in considering which of the research recommendations made should be regarded as a priority. Potential ways of identifying which of the research recommendations should be considered for DA-VOI and are outlined in chapter 9. Although this approach may reduce the resource requirements (it may not if DA-VOI is in addition to the TAR) and in the short run avoid the current capacity

problems in conducting this type of analysis there will always be a danger that some very valuable research requirements will be missed and other less valuable evidence requirements will be prioritised.

- However, these two alternatives need not be viewed as substitutes. The latter maybe regarded as the most feasible way of progressing in the short run. But, as capacity and methods develop, a move towards making DA-VOI a routine part of the TAR and the research recommendations in the guidance more firmly grounded on evidence and an explicit analysis may be achievable in the medium term.

1. Introduction and overview of methods

1.1 Introduction

This project developed as a result of the activities of the Research Teams at the Centre for Health Economics, University of York, and ScHARR at the University of Sheffield in the methods and application of decision analysis and value of information analysis (DA-VOI) as a means of informing the research recommendations made by NICE as part of its Guidance to the NHS in England and Wales and informing the deliberations of the NICE Research and Development Committee.

The specific project proposal was developed following a presentation of the potential role of DA-VOI in identifying those circumstances were additional evidence will be required to support guidance on the use of particular technologies and prioritise research recommendations made by the Appraisals Committee. In addition the framework of analysis allows decision makers to identify what type of evidence would be most valuable and the type of studies which should be conducted to better inform guidance decisions in the future.

1.2 Objectives

The specific objectives of the pilot study were to:

- Demonstrate the benefits of using appropriate decision analytic methods and value of information analysis to inform research recommendations.
- Establish the feasibility and resource implications of applying these methods in a timely way, to inform NICE.
- Identify critical issues and methodological challenges to the use of value of information methods for research recommendations (with particular regard to the new reference case as a suitable basis for this type of analysis).

The project consists of a series of case studies based on recent technology assessment reports completed by the York and Sheffield group for NICE. The purpose is to establish the feasibility and requirements of value of information analysis once submissions and Technology Assessment Reports (TARs) are conducted within the reference case specified in the recent methods guidance. Therefore case studies were selected and reported on the following basis:

- The existing TAR comes as close to the new reference case analysis as possible.
- The VOI analysis will be conducted using the case studies as they were developed and reported in the respective TARs.
- Any shortcomings with respect to the new reference case will be discussed, particularly if these have implications for the feasibility and reliability of value of information analysis.
- In each case, an assessment will be made of the suitability of the new reference case as a suitable basis of value of information analysis.
- Issues required for value of information analysis, which are not currently part of the reference case, will be highlighted.
- Selection of case studies will be made to highlight particular issues and challenges for value of information analysis.

A series of six case studies were selected based on recent technology assessment reports completed by York and Sheffield for NICE. These included:

- Screening for age related macular degeneration (AMD)
- Glycoprotein IIb/IIIa antagonists for acute coronary syndrome (GPAs)
- Clopidogrel and dipyridamole in the secondary prevention of occlusive vascular events (CLO)
- Neurominidase inhibitors for the treatment of influenza (NIs)

- Liquid based cytology screening for cervical cancer (LBC)
- Beta interferon and glatiramer acetate in the management of MS (MS)

A brief and non-technical overview of DA-VOI methods is presented below followed by brief reports on each of the case studies in chapters 2-7. Each case study chapter is intended to be a supplement to the original TAR on which the decision analytic model and probabilistic analysis is based. For full details of the analysis each chapter should be read in conjunction with the associated TAR. However, each case study chapter is intended to stand-alone and can be read independently of the rest of this report. Each case study chapter follows a common format including: the background to the original Appraisal and Guidance; a brief description of methods referencing the original TAR, a reporting of results for the adoption decision (estimates of cost-effectiveness and decision uncertainty) and for research recommendations (the value of information for the decision problem and for groups of model parameters), and a discussion of implications for research reconditions as well as some of the methodological issues raised specific to the case study.

Chapter 8 provides a general discussion of the results and the methodological issues and challenges raised during the pilot. It also includes an assessment of whether the new reference case will provide a good basis for value of information analysis and identifies some key issues in modelling such as, evidence synthesis, computation, bias, and structural uncertainty, which need to be addressed for the future application of DA-VOI. Chapter 9 provides a brief discussion of the feasibility of using DA-VOI to inform research recommendations, resource requirements and outlines possible strategies for the implementation of value of information methods to inform research recommendations within the NICE process. A summary of the key findings and conclusions are provided in chapter 10.

1.3 An overview of methods

Bayesian decision theory and value of information analysis provides an analytic framework which can be used to establish the value of acquiring additional information to inform a decision problem. These methods have firm foundations in statistical decision theory[7 8] and have been successfully used in other areas of research such as engineering and environmental risk analysis.[9 10 26] More recently these methods have been extended to setting priorities in the evaluation of health care technologies.[11-16] In addition they have been usefully applied to a number of different health technologies,[17-22] including a series of case studies taken from guidance issued by NICE.[23]

The application of these methods requires three core tasks to be completed: (i) the construction of a decision analytic model to represent the decision problem; (ii) a probabilistic analysis of this model to characterise the current decision uncertainty; and (iii) establishing the value of additional information.[19]

1.4 Decision analysis

Evaluative research is useful insofar as it informs the choice between alternative strategies for patient management. Decision analysis presents these decision problems and the key inputs to these decisions explicitly.[27 28] Decision modelling requires all of the relevant inputs to the decision to be explicitly identified, and facilitates the synthesis of data from a variety of sources.[16] Randomised trials are a crucial source of parameter estimates for decision models, particularly estimates of the magnitude of treatment effects. Other sources of data – for example, the baseline risk and resource implications of particular clinical events – may be taken from non-trial sources such as observational studies and administrative datasets. In some circumstances, where no evidence exists for particular inputs, clinical judgement may also be incorporated.

1.5 Probabilistic analysis

All decisions about the cost-effectiveness of interventions are based on uncertain information about variables such as clinical effects, health-related quality of life and resource use. Decision analytic models and methods of evidence synthesis can be used to combine evidence on each parameter to assess the extent of uncertainty in the decision.[28] The extent and the quality of the evidence available, for each of the inputs, can be reflected in probability distributions assigned to these estimates, where more uncertainty about an input (less information or information of poorer quality) is represented by assigning a more diffuse distribution. Without access to patient level data, these distributions are assigned based on secondary sources (e.g. published literature, meta- analysis and evidence synthesis). The choice of the type of distribution and its parameters for a particular model input is not arbitrary, but should be based on the existing evidence and what the type of distribution would be most appropriate. For example, probabilities should be represented by Beta distributions, which are bounded by zero and one, and their parameters can be based on either the number of observations or on mean and variance. [3 28-30]

The uncertainty surrounding the decision problem can be characterised by 'propagating' these distributions through the model using Monte Carlo simulation methods, where values for the input parameters are drawn at a random from the probability distributions which have been assigned.[3 28-30] This random sampling is repeated a large number of times. The output of these simulations provides a distribution of expected costs and outcomes for each strategy being compared. The uncertainty surrounding the cost-effectiveness of a technology, for a range of thresholds for cost-effectiveness, can be represented as a cost-effectiveness acceptability curve (CEAC).[11] Figure 1.1 illustrates an example of a CEAC where the probability that the intervention is cost effective increases as the willingness to pay for additional heath (QALY) or the threshold for cost-effectiveness increases.

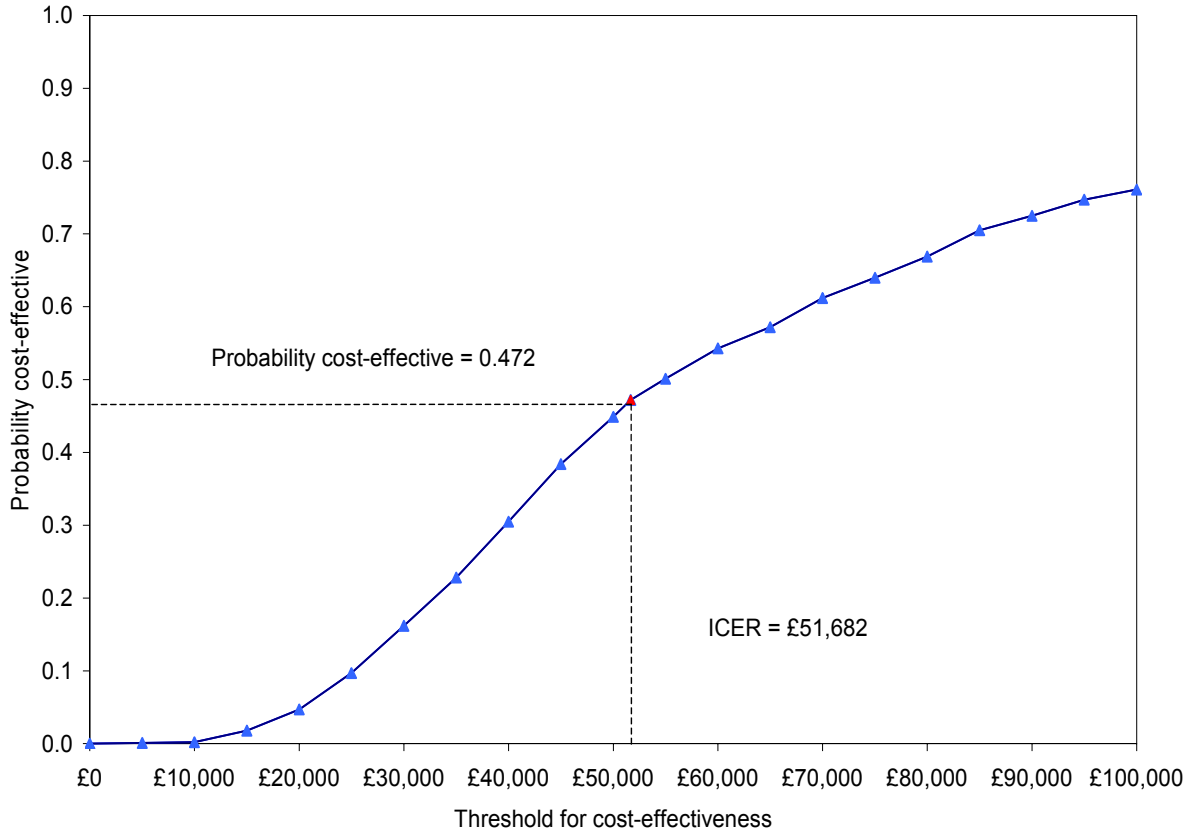


Figure 1.1: Cost-effectiveness acceptability curve example

If the objective underlying health technology assessment is to make decisions that are consistent with maximising health gains from available resources, then decisions should be based on expected cost-effectiveness given the existing information (i.e. using the mean differential costs and outcomes between the scenarios being compared). This does not necessarily mean that the intervention which has the highest probability of being cost-effective should be adopted. For example, in figure 1.1 if the threshold for cost-effectiveness was just greater than £51,682 (the ICER) then the intervention should be adopted even though the probability that it is cost-effective is less than 0.5 (0.472). This is because the distribution of the additional net benefits (where health outcomes are re-scaled in monetary terms using the cost-effectiveness threshold)[31 32]is positively skewed, with a mean greater than its median value. The adoption decision can be represented with a CEAC by including a cost-effectiveness frontier, which indicates which of the alternatives will be cost-effective.[11]

Although decisions should be based on expected cost-effectiveness given the existing information, this does not mean that adoption decisions can simply be based on little, or poor quality, evidence, as long as the decision to conduct further research to support adoption (or rejection) is made simultaneously.[12 19]

1.6 The value of information

Decisions based on existing information will be uncertain, and there will always be a chance that the wrong decision will be made. If the wrong decision is made, there will be costs in terms of health benefit and resources forgone. Therefore, the expected cost of uncertainty is determined jointly by the probability that a decision based on existing information will be wrong and the consequences of a wrong decision. The expected costs of uncertainty can be interpreted as the expected value of perfect information (EVPI), since perfect information can eliminate the possibility of making the wrong decision. If the objective of the health care system is to maximise gains in health outcome subject to a budget constraint then this is also the maximum that the health care system should be willing to pay for additional evidence to inform this decision in the future, and it places an upper bound on the value of conducting further research.[12 13 17 19 33] However, there may be other objectives of health care provision such as equity. If these other objectives can be identified and valued then these can be incorporated into the analysis and the societal value of information.[12]

This general idea is illustrated in Figure 1.2. With current information, decisions must be made before we know how the uncertainties (p(x)) will be resolved - i.e. we must make a decision now based on the expected values of the all of the model inputs (choose “std” in Figure 2). However, with perfect information we can make our decisions once we know how these uncertainties (p(x)) are resolved - i.e., we can make different decisions for different resolutions of the uncertainties (choose “std” if 1-p(x) but choose “Ex” if p(x)). The EVPI is simply the difference between the payoff (expected net benefit) with perfect and current information.[21 33]

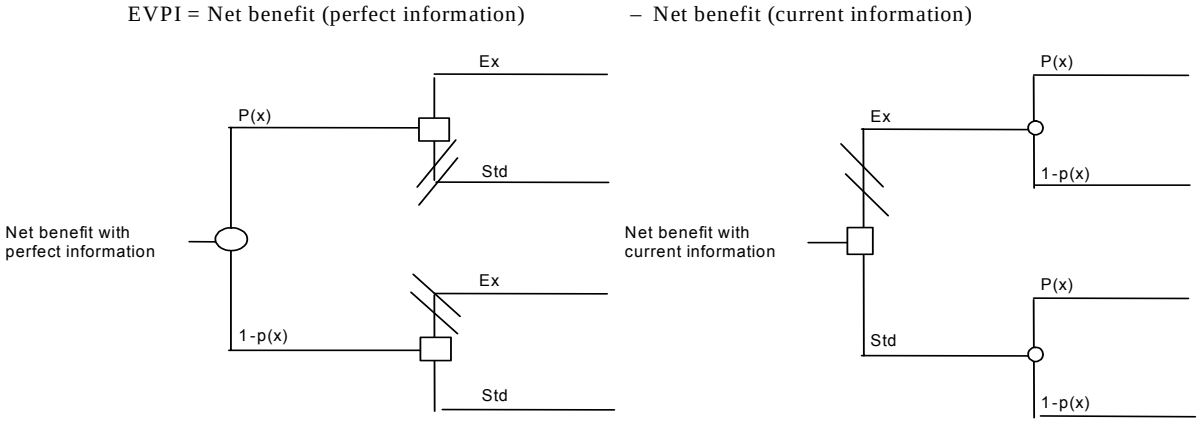


Figure 1.2: Calculating EVPI example

We can work out EVPI directly from the simulated output from our model as it relates to the individual patient.[9 21 33 34] For example, if there are two alternative interventions (t = 1,2) interventions, with unknown parameters θ. Then given the existing evidence, the optimal decision is the intervention that generates the maximum expected net benefit:

$$\max_t E_\theta B(t, \theta),$$

i.e., the maximum net benefits over all the iterations from the simulation because each iteration represents a possible future realisation of the existing uncertainty (a possible value of θ).

With perfect information, the decision-maker would know how the uncertainties would resolve (which value θ will take) before making a decision and could select the intervention that maximises the net benefit given a particular value of theta:

$$\max_t B(t, \theta).$$

However, the true values of θ are unknown (we don't know which value θ will take), Therefore, the expected value of a decision taken with perfect information is the found by averaging the maximum net benefit over the joint distribution of θ:

$$E_\theta \max_t B(t, \theta),$$

i.e., first calculate the maximum net benefit for each iteration from the simulation (for a particular value of θ), then take the average over these maximum net benefits (over the possible values of θ).

The expected value of perfect information for an individual patient is simply the difference between the expected value of the decision made with perfect information about the uncertain parameters θ, and the decision made on the basis of existing evidence:

$$E_\theta \max_t B(t, \theta) - \max_t E_\theta B(t, \theta).$$

This provides the EVPI surrounding the decision as a whole for each time this decision is made (for an individual patient or individual episode). However, once information is generated to inform the decision for an individual patient or patient episode then it is available to inform the management of all other current and future patients as well. Therefore, for research prioritisation it is important that EVPI is expressed for the total population of patients who stand to benefit from additional information over the expected lifetime of the technology. This

requires some assessment of the effective lifetime of the technology (the period over which information about the decision will be useful), estimates of incidence over this period and prevalence in the first year. The EVPI associated with future patients is discounted to provide the total EVPI for the population of current and future patients. If this population EVPI exceeds the expected costs of additional research, then it is potentially cost-effective to conduct further research.[12 13]

Figure 1.3 illustrates the population EVPI for the example used in Figure 1.1. When the threshold for cost-effectiveness (maximum value of health outcome) is low, the technology is not expected to be cost-effective and additional information is unlikely to change that decision (EVPI is low). In these circumstances the EVPI increases with the threshold because both the decision uncertainty increases (tendig to increase the EVPI) and the consequences of information changing the decision are valued more highly. Conversly, when the threshold is higher than the ICER, the intervention is expected to be cost-effective and this decision is less likely to be changed by further research as the threshold is increased. In these circumstances the decision uncertainty falls as the threshold increases (tending to reduce the EVPI), but the consequences of information changing the decision are valued more highly (tending to increase the EVPI). For higher values of the threshold the EVPI falls because in this case the reduction in decision uncertainty off sets the increased value of information changing decision. However, the EVPI will ultimately increase with very high values of the threshold because decision uncertainty will fall at a declining rate with the threshold increasing at a constant rate i.e., the effect of the value of changes in decision will ultimately off set the reduction in decision uncertainty. In this particular case, the population EVPI reaches maximum when the threshold is equal to the expected incremental cost-effectiveness ratio of this technology. In other words, the EVPI reaches a maximum when we are most uncertain about whether to adopt or reject the technology based on existing evidence.[12 13 17]

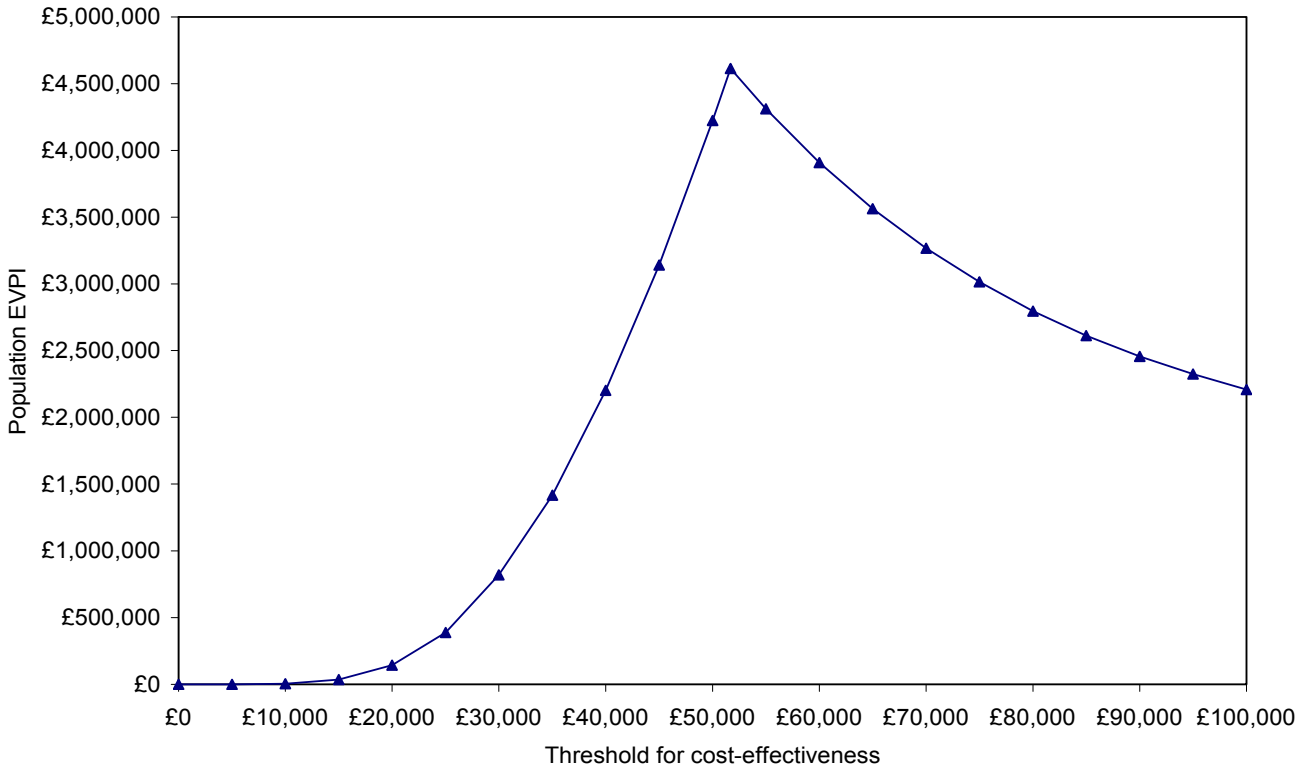


Figure 1.3: EVPI curve example

However, most decision problems involve more than two alternatives. The principles of calculating EVPI remain the same but the EVPI curve can take a variety of shapes depending on whether the alternatives being considered are cost-effective at some value of the threshold (there will be a number of peaks or a discontinuities in the EVPI curve at threshold values equal to the ICER of each of the alternatives) or if some of the alternatives are dominated or extendedly dominated (the peak or discontinuity will be in negative threshold space i.e., we would only wish to adopt the alternative if your willing to pay to reduce heath outcome).

It should be clear from this discussion of EVPI suggests that the value of further research will depend on both the uncertainty surrounding estimates of cost and effect but also on how cost-effective or cost ineffective a technology is expected to be given existing evidence, and the size of the patient population that could benefit from additional research. One implication is that it is perfectly possible that the value of further research about a new technology, which is substantially cost-effective based on existing evidence, will be very low even if there is uncertainty surrounding the parameters, i.e., there may be uncertainty in cost and outcomes but the decision uncertainty and therefore the EVPI may still be low. In these circumstances the technology should be adopted and no further research is required to support this decision.

The value of reducing the uncertainty surrounding particular input parameters in the decision model can also be established (partial EVPI). This type of analysis can be used to focus further research by identifying those inputs for which more precise estimates would be most valuable. In some circumstances, this will indicate which endpoints should be included in further experimental research. In other circumstances, it may focus research into getting more precise estimates of particular inputs which may not necessarily require experimental design and can be provided relatively quickly. The analysis of the value of information associated with each of the model inputs (partial EVPI) is, in principle, conducted in a very similar way to the EVPI for the decision as a whole.[9 17 21 33 34] In this case the expected value with perfect information is found by taking the maximum expected net benefit given perfect information only about the parameter of interest (calculating expected net benefits over all the other uncertain parameters the model) and then averaging over all the possible value of the parameter of interest. The EVPI for the parameter is again simply the difference between the expected net benefit with perfect information and the expected value with current information (the same as for decision EVPI). However, this does require substantial additional computation for models where the relationship between the inputs and expected cost and outcomes is not linear, for example in Markov models.[21 33]

Figure 1.4 illustrates the partial EVPIs associated with the decision EVPI in Figure 1.3 at a threshold of £40,000 per QALY. In this example, the EVPI associated with reduction in symptom days is relatively high and suggests that further experimental research may be worthwhile. However, other inputs with lower partial EVPI, such as the baseline probability of hospitalisation, may not require experimental research but may also be important if the costs of further investigation (resources and delay) are low. It should be noted that the partial EVPIs will not sum to the overall EVPI due to the interactions within the model structure.

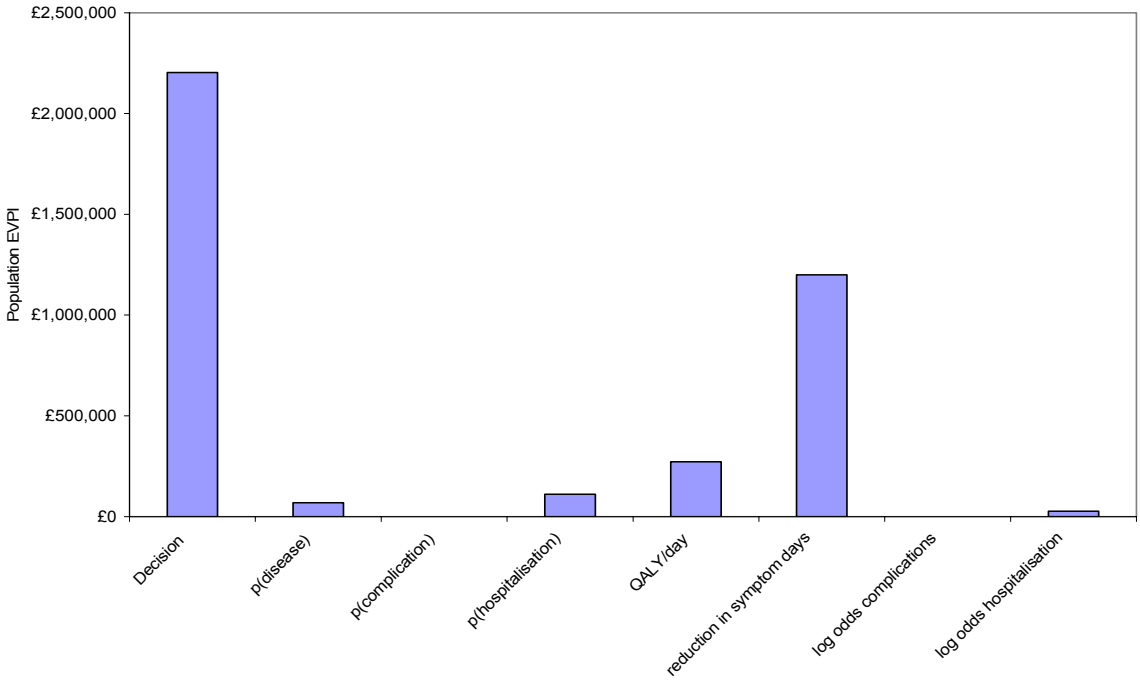


Figure 1.4: Partial EVPI example

1.7 Prioritising research recommendations

The EVPI places an upper bound on the societal returns to further investigation. The EVPI for the decision problem can be used as a first hurdle for proposed research.[12 13 17 19 21] If the costs of investigation exceed the EVPI, then the proposed research will not be cost-effective, i.e., the population EVPI can be used to rule out research recommendations which will not be worth while for a societal perspective.

For those decision problems where the EVPI exceeds the costs of research it is possible to compare EVPIs across patient groups and different technologies to prioritise research recommendations. In general, additional research will be more valuable for a patient groups or technology where the EVPI is higher. However, it should be noted that this direct comparison requires some assessment of the cost of proposed research. For example, even where the EVPI is lower research in that area may not be a lower priority if the costs of further investigation are expected to be substantially lower. In principle it would be useful to explicitly compare the marginal benefits and costs of research proposals in prioritising research. The same framework of DA-VOI can be extended to establish the expected value of sample information for particular research designs and to compare these marginal benefits of research to the marginal costs. However, this type of analysis is beyond the scope of the current pilot.[12 14 18 33]

The EVPI associated with groups of model parameters can be used to focus potentially cost-effective research on those inputs for which more precise estimates would be most valuable. This may indicate which endpoints should be included in further experimental research, or it may focus research on getting more precise estimates of particular inputs, which may not necessarily require experimental design and can be provided relatively quickly.[17 21 33]

1.8 Conclusion

Bayesian decision analysis and value of information analysis provides a methodological framework which explicitly considers the uncertainty surrounding the decision of a health care system to adopt a health technology. Specifically, using existing evidence, these methods

focus on the likelihood of making a wrong decision if the technology is adopted. The value of additional research is based on the extent to which further information will reduce this decision uncertainty. This framework values the additional information, which may be generated by further research, in a way which is consistent with the objectives and the resource constraints of heath care provision (the cost-effectiveness threshold). This allows a comparison of the potential benefits of further research with the costs of further investigation, a comparison and prioritisation of alternative research recommendations, both within and between Technology Assessments, as well as an assessment of the value of investing resources in research or other activities, such as the provision of health service. In this sense it provides a unified and coherent framework for prioritisation of research and the use of heath care technologies.

References

1. Townsend J, Buxton M. Cost-effectiveness scenario analysis for a proposed trial of hormone replacement therapy. *Health Policy* 1997;39:181-94.
2. Davies L, Drummond MF, Papanikolaou P. Prioritizing investments in health technology assessment. *International Journal of Technology Assessment in Health Care* 2000;16:73-91.
3. Buxton M, Hanney S. Assessing payback from Department of Health Research and Development: Second report. HERG Research Report 24. Uxbridge: Brunel University, 1997.
4. Phelps CE, ST P. Priority setting in medical technology and medical practice assessment. *Medical Care* 1990;28.
5. Raiffa H, Schlaifer R. *Probability and Statistics for Business Decisions*, 1959.
6. Pratt J, Raiffa H, Schlaifer R. *Statistical decision theory*. Cambridge MA, 1995.
7. Thompson KM, Evans JS. The value of improved national exposure information for perchloroethylene (perc): a case study for dry cleaners. *Risk Analysis* 1997;17:253-271.
8. Howard RA. Information value theory. IEEE Transactions on Systems Science and Cybernetics SSC2, 1966:122-26.
9. Fenwick. E, Claxton. K, Sculpher. M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics* 2001;10:779-89.
10. Claxton K. The irrelevance of inference: a decision making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics* 1999;18:341-364.
11. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. *Health Economics* 1996;5:513-524.
12. Claxton K, Walker S, Lacey L. Selecting treatments: a decision theoretic approach. *Journal of the Royal Statistical Society* 2000;163:211-225.
13. Claxton K, Thompson K. A dynamic programming approach to efficient clinical trial design. *Journal of Health Economics* 2001;20:432-448.
14. Sculpher M, Drummond M, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. *Journal of Health Services Policy* 1997;2:26-30.
15. Claxton K, Neuman PJ, Araki SS, Weinstein MC. The value of information: an application to a policy model of Alzheimer's disease. *International Journal of Technology Assessment in Health Care* 2001;17:38-55.
16. Claxton K, Neumann P, Araki S, Weinstein MC. The efficient design of clinical trials: an application to the evaluation of treatment strategies for Alzheimer's disease. *Journal of Health Economics* forthcoming.
17. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence. *Lancet* 2002;360:711-715.
18. Fenwick E, Claxton K, Sculpher M. A Bayesian analysis of pre-operative optimisation of oxygen delivery. *Medical Decision Making* 2000;20.
19. Fenwick E, Claxton K, Sculpher M, Briggs A. Improving the efficiency and relevance of health technology assessment: the role of decision analytic modelling. Centre for Health Economics Discussion Paper No 179, 2000.
20. Palmer S, Fenwick E, Claxton K, Sculpher M. Applications of value of inforamtion using Bayesain numerical methods. *Medcial Decision Making* 2001;21.
21. Claxton K. Efficient fourth hurdle regulation: an application of value of information analysis to a sample of health technologies (Abstract). *Medcial Decision Making* 2001;21.
22. Fleurence RL, Torgerson DJ. Setting priorities for research. *Health Policy* 2004;Forthcoming.
23. Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effective models; a suggested framework and example of application. *Pharmacoeconomics* 2000;17:461-477.
24. Hammitt JK, Cave JAK. Research planning for food safety: a value of information approach.: RAND publication series, 1991.
25. Hunink M, et al. *Decision making in health and medicine*, 2001.

26. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000;17:479-500.
27. Critchfield GC, Willard KE. Probabilistic analysis of decision trees using Monte Carlo simulation. *Medical Decision Making* 1986;6:85-92.
28. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probalistic sensitivity analysis using Monte Carlo simulation. *Medical Decision Making* 1985;5:157-17.
29. Phelps C, Mushlin AI. On the near equivalence of cost-effectiveness analysis. *International Journal of Technology Assessment in Health Care* 1991;17:12-21.
30. Stinnett A, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical Decision Making* 1998;18:S68-S80.
31. Ades AE. Markov Chain Monte Carlo estimation of a multi-parameter decision model: consistency of evidence, and the accurate assessment of uncertainty. Medical Decision Making 2002 in press.
32. Felli JC, Hazen GB. Sensitivity analysis and the expected value of perfect information. *Medical Decision Making* 1998;18:95-109.

2. The cost-effectiveness and value of information associated with repeat screening for age related macular degeneration

2.1 Background

2.1.1 Condition and technology

Age-related macular degeneration (AMD) is a degenerative condition of the macula. It is one of the most common causes of vision loss in people over 50. The disease varies in severity, from a slight loss in vision to near blindness. AMD is classified as either wet (neovascular) or dry (non-neovascular), it is neovascular AMD that progresses most rapidly and causes the more severe vision loss. About 10% of patients who suffer from macular degeneration have wet AMD. If one eye develops neovascular membrane, the other eye is at moderate risk of having the same problem. Neovascular AMD is further defined by its location in the choroidal neovascular vessels (subfoveal, justafoveal or extra-foveal) and by its pattern of leakage (classic, occult, mixed or recurrent).(1)

Treatments for certain types of AMD have developed over the last few years and include confluent argon laser photocoagulation, verteporfin photodynamic therapy (PDT), radiotherapy and transpupillary thermotherapy.[1]

2.1.2 Technology Assessment Review

Photodynamic therapy for age related macular degeneration has recently been appraised by the National Institute for Clinical Excellence (NICE).[2]The evidence from the Assessment Report, and the provisional guidance issued by NICE, indicated that PDT will only be potentially cost-effective for the treatment of AMD in the better seeing eye (after 1st eye involvement) and only for certain types of AMD (neovascular, predominantly classic, subfoveal) AMD can progress rapidly (declining visual acuity) and is a significant cause of blindness. Early PDT can halt or slow the decline in visual acuity. Earlier treatment with PDT at better starting visual acuities is more cost-effective, and treatment is not recommended for starting visual acuities lower than 20/100.

2.1.3 NICE guidance

In September 2003, NICE issued guidance on the use of PDT for age-related macular degeneration. The guidance recommended the following[3]

1. Photodynamic therapy (PDT) is recommended for the treatment of wet age-related macular degeneration for individuals who have a confirmed diagnosis of **classic with no occult** subfoveal choroidal neovascularisation (CNV) and best-corrected visual acuity 6/60 or better. PDT should be carried out only by retinal specialists with expertise in the use of this technology.

2. PDT is not recommended for the treatment of people with **predominantly classic** subfoveal CNV associated with wet age-related macular degeneration, except as part of ongoing or new clinical studies that are designed to generate robust and relevant outcome data, including data on optimum treatment regimens, long-term outcomes, quality of life and costs.

3. The use of PDT in **occult** CNV associated with wet age-related macular degeneration was not considered because the photosensitising agent (verteporfin) was not licensed for this indication when this appraisal began. No recommendation is made with regard to the use of this technology in people with this form of the condition.

4. Patients currently receiving treatment with PDT could experience loss of well being if their treatment is discontinued at a time they did not anticipate. Because of this, all NHS patients who have begun a course of treatment with PDT at the date of publication of this guidance should have the option of continuing to receive treatment until their clinical condition indicates that it is appropriate to stop.

2.1.4 Research recommendations

The recommendations for further research, issued as part of the NICE guidance were that:

- Several randomised controlled trials of PDT are ongoing, including two placebo-controlled trials of verteporfin PDT, one in patients with minimally classic CNV using standard or reduced laser settings, and one in patients with occult CNV.

- The Committee recommended that further research is needed on the use of PDT for individuals with predominantly classic subfoveal CNV related to ARMD. The primary objectives of this research should be to determine the optimum treatment regimen and long-term benefit of PDT, and to add to the current evidence on quality of life for this group of individuals.

- At present it is not known whether population screening for ARMD would be practical or cost effective. Research on screening for ARMD is being commissioned by the UK HTA programme.

2.2 Methods

Given that treatment with PDT is more effective the earlier it is initiated in the course of the disease, there is a *prima facie* case that screening would be cost-effective by identifying patients with AMD before their visual acuity declines. A self-screening test of central vision distortion called the Amsler grid[4] is available and it has been suggested that this could be used as a basis of screening.[2]

Following the conclusions of the Assessment Report undertaken for NICE[2] and the provisional guidance regarding the use of PDT, we have focused on the use of weekly self-screening following 1st eye involvement with neovascular AMD. This self-screening strategy is compared to two alternatives: no screen but diagnosis and treatment of eligible AMD following self-referral (due to declining visual acuity) to an ophthalmologist (this strategy is consistent with provisional NICE guidance); and a strategy of no screening and no PDT. The analysis reported here has assessed the cost-effectiveness of, and potential value of future research for, these alternative strategies. Full detail of this model can be seen in a forthcoming HTA report.(5)

Although the structure of the model developed was consistent with published evidence regarding the natural history of the disease, a number of structural assumptions were questioned after consultation with clinical experts. Alternative structural assumptions, consistent with these alternative clinical opinions, were therefore explored

2.2.1 Model structure

The structure of the decision model is illustrated in Figure 2.1. A Markov process[6] is used to model the incidence of 2nd eye neovascular AMD over 10-years and the associated decline in visual acuity following undiagnosed 2nd eye involvement.

Patients enter the model with neovascular AMD previously diagnosed in the 1st eye. Two alternative starting visual acuities are modelled, 20/40 or 20/80. The implication of this is that the worse seeing individuals would generally receive PDT or no PDT at a lower visual acuity. Each week patients can decide to self-screen (comply) using the Amsler grid, which is an A4 sheet of paper containing a series of lines, which appear distorted if a change in vision has occurred.

Patients with positive screen results (self diagnosed) will refer for a full eye exam by an ophthalmologist. Patients may also self-refer due to declining visual acuity, measured as a loss of 1 or more lines. In the base case model at a loss of 4 or more lines, all patients will have self-referred to the ophthalmologist (in the absence of data expert judgement was used to specify the probability that patients self refer on loss of visual acuity). The full eye examination will identify patients with neovascular AMD in the 2nd eye (i.e. false positives are

identified). Angiography is then undertaken in those with confirmed neovascular AMD to identify the type of neovascular disease that is present, and thus determine if the patient is eligible for PDT. Since angiography is used to identify and monitor AMD in the clinical trials of PDT it is taken to be the gold standard test in this model.

Patients with diagnosed AMD that is eligible for PDT will then either have PDT (screen + PDT and no screen + PDT strategy), or not have PDT (no screen + no PDT strategy). Costs and QALYs associated with the use or non-use of PDT are then assigned. The expected quality of life with PDT depends on the visual acuity at diagnosis, where patients with better visual acuities will experience better quality of life. The costs of PDT are constant throughout the visual acuity groups.[7]

2.2.2 Alternative structural assumptions

Currently the effect of the Amsler grid, in terms of identifying patients with AMD, can occur before AMD develops and at each stage of visual acuity loss. However it may be that there is no additional benefit from the Amsler grid after a patient has developed a visual acuity problem (after a loss of one or more lines). Also patients only refer to see an ophthalmologist once they have a decline in visual acuity, when there are noticeable changes in their eyesight. However given that patients in the model have already had first eye involvement and may therefore be expected to be more vigilant in recognising changes in their vision, patients may refer when there is no loss in visual acuity (AMD state). This may be because they are using other stationary objects to imitate the Amsler grid. These alternative structural assumptions were modelled.

A 10-year time horizon was used in the model as, during this period, almost all patients developed 2nd eye disease (96%), and this disease was diagnosed in most patients in both screen and no screen groups (92%). In other words, all patients with 2nd eye involvement will be diagnosed at some point, the question is when and at what visual acuity this happens. Given the decision problem to be addressed, an NHS perspective was used for the analysis. Health benefits are expressed in terms of QALYs.

2.2.3 The evidence

A full list of all the parameters and their sources is presented in table 2.1. The incidence of 2nd eye neovascular AMD,[8] the eligibility for PDT (sub-types of AMD),[9] the sensitivity and specificity of the Amsler grid screen[10] and compliance with self screening[(11] were all based on observational studies. Beta distributions were assigned to reflect the amount of evidence available for each of these parameters using measures of variance reported in the studies. Although two trials of PDT for AMD are available, only the TAP trial included predominantly classic AMD. The decline in visual acuity for undiagnosed 2nd eye involvement was based on the 2- year results of the control arm of the TAP trial of PDT[2] as reported in the NICE Assessment Report, with beta distributions assigned to these transition probabilities.

No evidence was available regarding the probability that patients will self-refer following each decline in visual acuity. Therefore, expert judgments from a primary care physician with specialist research interest in AMD were used with beta distributions reflecting the additional uncertainty about range of possible values.

All cause mortality was also incorporated in the model (for a males and female population aged 55-64) based on UK life tables.[(12)]

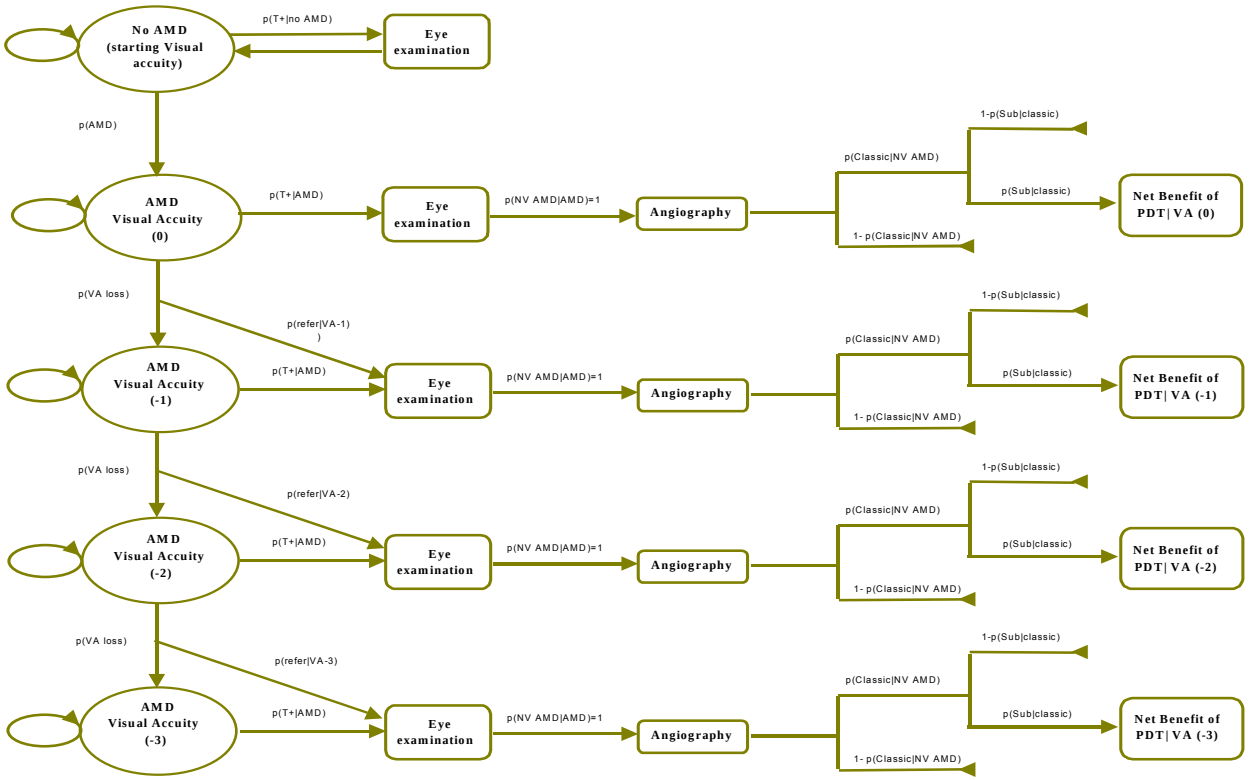


Figure 2.1: Model structure for AMD self-screening

Table 2.1: Sources of data

Parameters	Value	Distribution	Source
Incidence of 2nd eye neovascular AMD	Yr 1 = 0.00198 Yr 2 = 0.00404 Yr 3 = 0.00685 Yr 4 = 0.0124 Yr 5 = 0.0114	Beta distributions Alpha = 9, Beta = 95 Alpha = 18, Beta = 56 Alpha = 17, Beta = 36 Alpha = 11, Beta = 12 Alpha = 5, Beta = 6	(8)
Progression of visual acuity with 2nd eye involvement	VA0 to VA1 = 0.014 VA1 to VA2 = 0.062 VA2 to VA3 = 0.062 VA3 to VA4 = 0.060	Beta distributions Alpha = 2.93, Beta = 204.06 Alpha = 12.54, Beta = 189.16 Alpha = 12.44, Beta = 187.70 Alpha = 11.93, Beta = 186.19	(2)
Accuracy of Amsler grid Sensitivity Specificity	 0.59 0.04	 Beta Alpha = 65.65, Beta = 44.35 Alpha = 5.28, Beta = 104.72	(10)
Compliance with self screening	0.55	Beta Alpha = 49, Beta = 40	(11)
Eligibility for PDT (AMD sub-groups)	0.56	Beta Alpha = 17, Beta = 829	(9)
Self referral on decline in VA	VA1 = 0.2 VA2 = 0.6 VA3 = 0.8 VA4 = 1	Beta distributions Alpha = 2, Beta = 8 Alpha = 6, Beta = 4 Alpha = 8, Beta = 2 Constant	Clinical judgement
QALYs with PDT	VA0 = 2.34 VA1 = 2.26 VA2 = 2.12 VA3 = 2.10 VA4 = 2.	Gamma distributions Alpha = 193.58, Beta = 0.012 Alpha = 218.58, Beta = 0.010 Alpha = 406.31, Beta = 0.005 Alpha = 404.02, Beta = 0.005 Alpha = 370.49, Beta = 0.005	(7)
QALYs without PDT	VA0 = 2.17 VA1 = 2.12 VA2 = 2.03 VA3 = 2.01 VA4 = 1.99	Gamma distributions Alpha = 217.67, Beta = 0.010 Alpha = 178.15, Beta = 0.011 Alpha = 287.43, Beta = 0.007 Alpha = 286.02, Beta = 0.007 Alpha = 287.30, Beta = 0.006	(7)
Costs of PDT	£6475.35	Constant	(7)
Costs of diagnosis and screen	£55.88 + £112 + £108	Constant	(2, 13)

Costs and QALYs associated with the use of and non-use of PDT at the different visual acuity levels (20/40, 20/50, 20/64, 20/80, 20/100, 20/126) are taken from the output of a cost-effectiveness model of PDT developed as part of the NICE appraisal of PDT.[7] The authors used a Markov model to estimate the costs and outcomes of PDT with verteporfin using patient level data taken from the TAP trial.[14] Two-year (within-trial estimate) and 5 year time periods were used to assess cost and outcomes. Time-trade-off methods were used by Brown et al[(15)] to elicit utilities for various visual acuity levels in the model, and utility decrements from adverse events were estimated through expert panel.[15] Gamma distributions

were assigned to expected QALY gains on PDT using the reported means and variances taken from the simulated model output. Costs used in the model were taken from Meads, et al.[2] and the model output suggested that costs were constant across visual acuity states.

The costs of ophthalmologist screening and angiography (diagnosis) were based on the NICE Assessment Report.[2] Given that the Amsler grid only constitutes an A4 sheet of paper, if self-administered, the costs of self-screening are zero.

2.2.4 Probabilistic analysis

To reflect uncertainty in the parameters in the model, they were incorporated as probability distributions,[16] full details of which are available elsewhere.[5] Monte Carlo simulation was used to propagate the prior distributions assigned to model inputs and estimate the expected costs and outcomes associated with each alternative therapy and incremental cost effectiveness ratios were calculated. The simulation estimates mean expected costs and QALYs and their associated distributions.

The results of the model are presented in two ways. Firstly, mean costs and QALYs for the various comparators are presented and their cost-effectiveness compared, estimating incremental cost-effectiveness ratios as appropriate, using standard decision rules.[17] Given that mean costs and QALYs gained are estimated with uncertainty, the output from the simulations were then used to generate cost-effectiveness acceptability curves[18;19] for the 3 strategies.

2.2.5 Value of information analysis

The output of these simulations was also used to estimate the expected value of perfect information (EVPI)[20,21] for individual patients. Population EVPIs were based on the incidence of 2nd eye AMD (not differentiated by type or visual acuity) taken from the NICE Assessment Report[2] and alternative assumptions about the expected lifetime of the technology of 5, 10 and 15 years. An analysis of the parameter EVPIs associated with groups of model inputs was also conducted.

2.3 Results

2.3.1 Adoption decisions

The results for both 20/40 and 20/80 starting visual acuity patients can be seen in Table 2.2. As expected, the screen and treatment option has the highest costs for both 20/40 and 20/80 starting visual acuities, £3651 and £3662 respectively. The no screen and no treatment strategy is associated with the lowest costs for both starting visual acuities. Treatment with PDT is associated with additional QALYs, therefore the 2 strategies that involve treatment have higher QALYs than the no treatment strategy. However there are large differences between the numbers of QALYs gained in the screen + treat and no screen + treat strategies. This is because the screening allows patients be diagnosed at a better visual acuity level, that is before it has declined significantly, and those higher visual acuity groups are associated with better outcomes (QALYs).

Cost-effectiveness and decision uncertainty are extremely similar for both males and females, hence we have only reported the results for males here.

Table 2.2: Results from the base case model

	Mean QALYs	Mean costs	ICER
Starting visual acuity = 20/40			
Screen and treat	1.2136	£3,651	£12,892
No screen and treat	0.9836	£2,643	E dominated
No screen and no treatment	0.9377	£98	
Starting visual acuity = 20/80			
Screen and treat	1.0915	£3,662	£17,757
No screen and treat	0.9215	£2,644	E dominated
No screen and no treatment	0.8908	£98	

For patients with a starting visual acuity of 20/40, screening can be regarded as cost-effective when compared to no treatment, with an incremental cost per additional QALY of £12,892. The strategy of no screen but treatment on diagnosis is not cost-effective when compared to no treatment (incremental cost per additional QALY equals £54,446) and is subject to extended dominance.[17]

The cost-effectiveness of screening is, however, uncertain. The probability that each intervention is cost-effective is reported in Figure 2.2 and shows that the probability that no screen and treat will be cost-effective remains very close to zero over a range of cost-effectiveness thresholds. The probability that screening is cost-effective with a threshold for cost-effectiveness of £30,000 is 0.89.

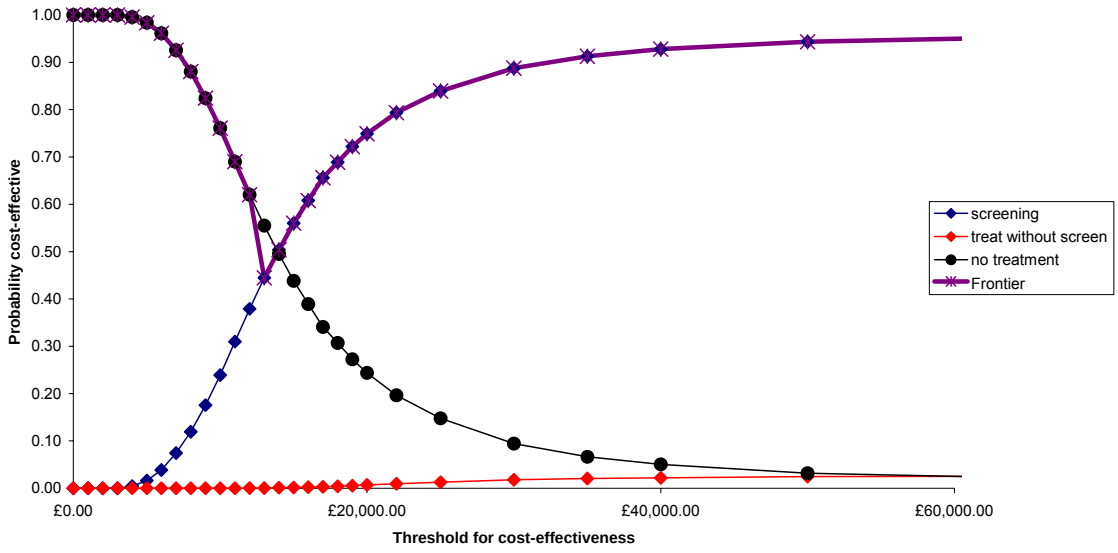


Figure 2.2: Cost-effectiveness acceptability curve for 20/40 model

For patients with a lower starting visual acuity of 20/80 screening is less likely to be considered cost-effective when compared to no treatment (Figure 2.3), although the incremental cost per additional QALY equals of £17,757 is higher than in the group with higher starting visual acuity and is more uncertain (probability that screening is cost-effective with a threshold for cost-effectiveness of £30,000 is 0.73).

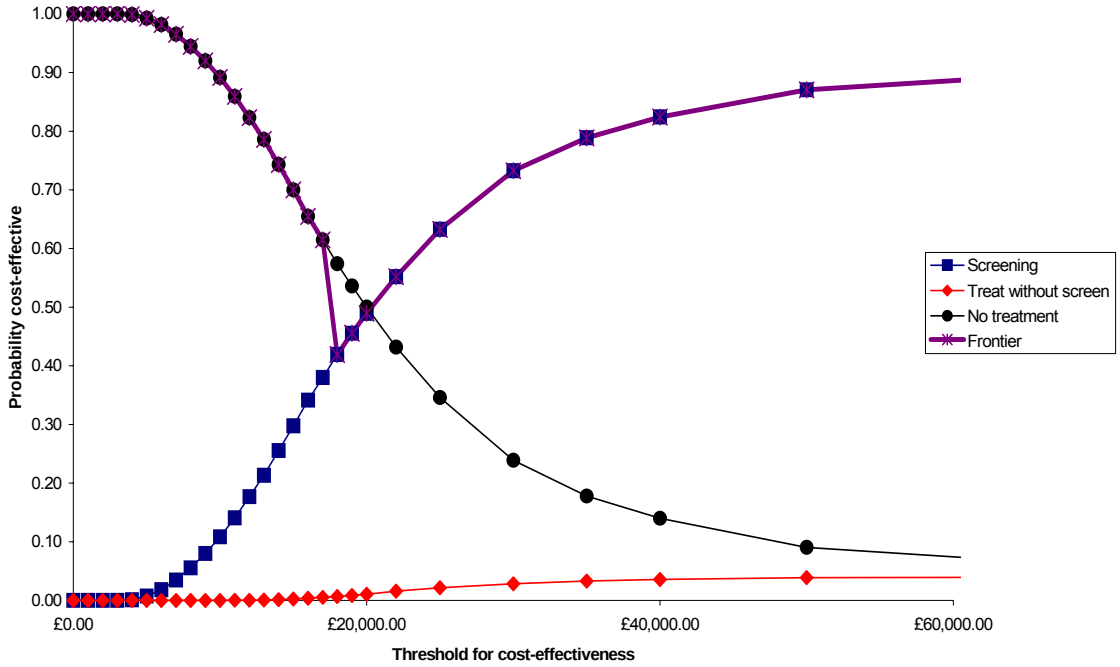


Figure 2.3: Cost-effectiveness acceptability curve for 20/80 model

2.3.2 Research recommendations

The population EVPI for both starting visual acuities is shown in Figure 2.4. At a threshold for cost-effectiveness of £30,000, the population EVPI is £6.18 million assuming a 10-year lifetime for the technology (£158 for individual patients) or £3.54 and £8.15 million assuming a lifetime of 5 and 15 years, respectively. The population EVPI with starting visual acuity of 20/80 is higher (because of increased decision uncertainty): £15.33 million assuming a 10 year life time of the technology (£393 for individual patients), or £10.24 and £20.22 million assuming a lifetime of 5 and 15 years, respectively. Estimates of EVPI for male and female populations are very similar.

The EVPI for each of the groups of model inputs for the 20/40 model is illustrated in Figure 2.5 for a threshold for cost-effectiveness of £30,000 and a 10-year lifetime for the technology. For patients with a starting visual acuity of 20/40, the value of information associated with the expected QALYs with or without PDT is £3.73 million. The other groups of model inputs, such as screening accuracy, have no value of information associated with them. At a starting visual acuity of 20/80 (Figure 2.6), the value of information associated with the expected QALY with or without PDT is £4.67 million and the value associated with the progression of visual acuity is just below £194,000. The other groups of model inputs have no value of information associated with them.

In general, individual EVPIs will not sum to the EVPI for the decision as a whole. In this case many of the groups of model inputs have no value associated with them. This does not mean that the uncertainty surrounding their values is unimportant (together, they generate the EVPI for the decision), but it does mean that more information about these inputs *individually* may not be valuable.

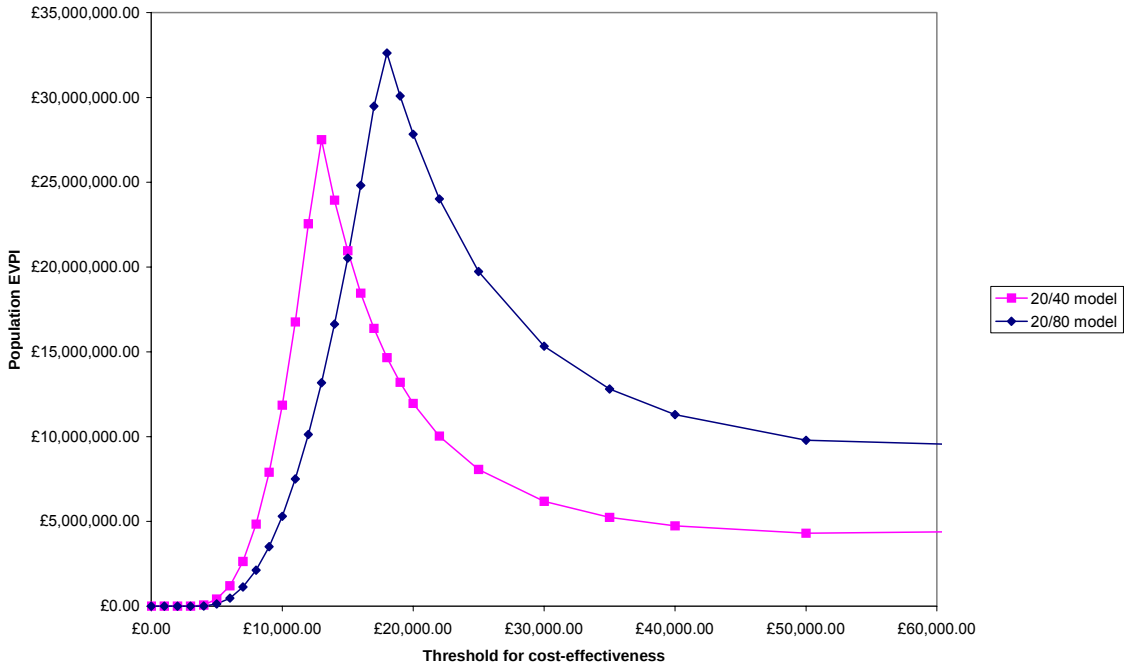


Figure 2.4: Population EVPI for 20/40 and 20/80 models

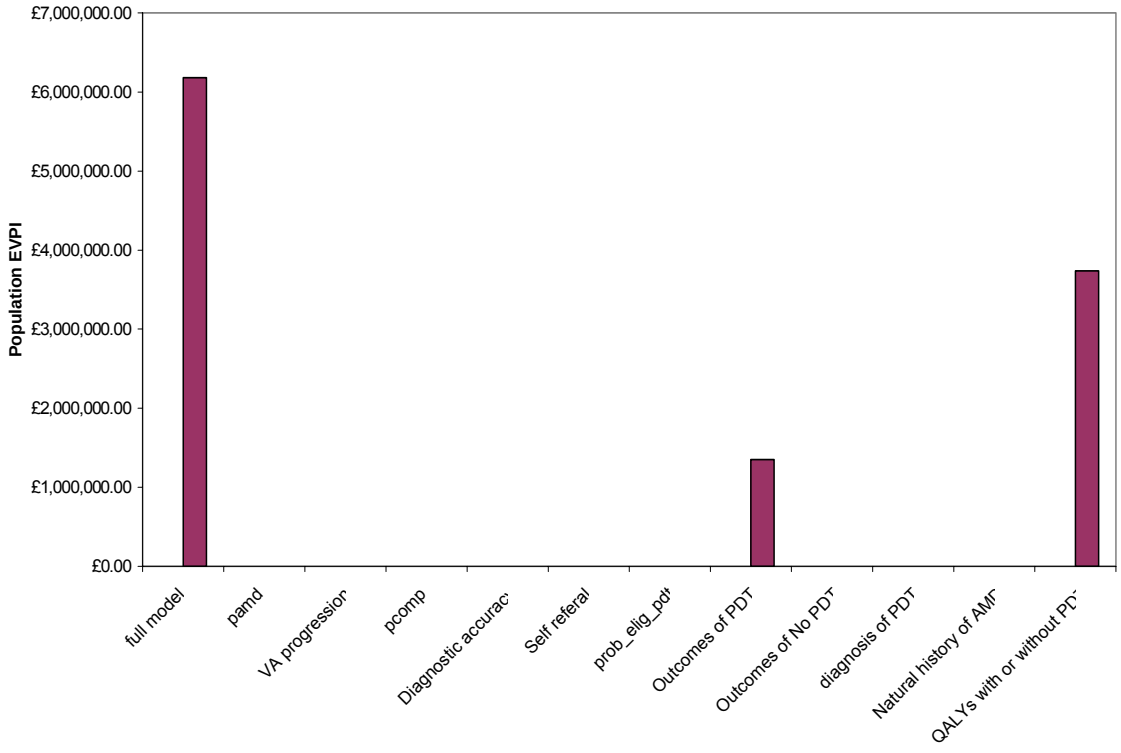


Figure 2.5: EVPI for model inputs: starting visual acuity of 20/40
(Lifespan: 10 years. Threshold: £30,000)

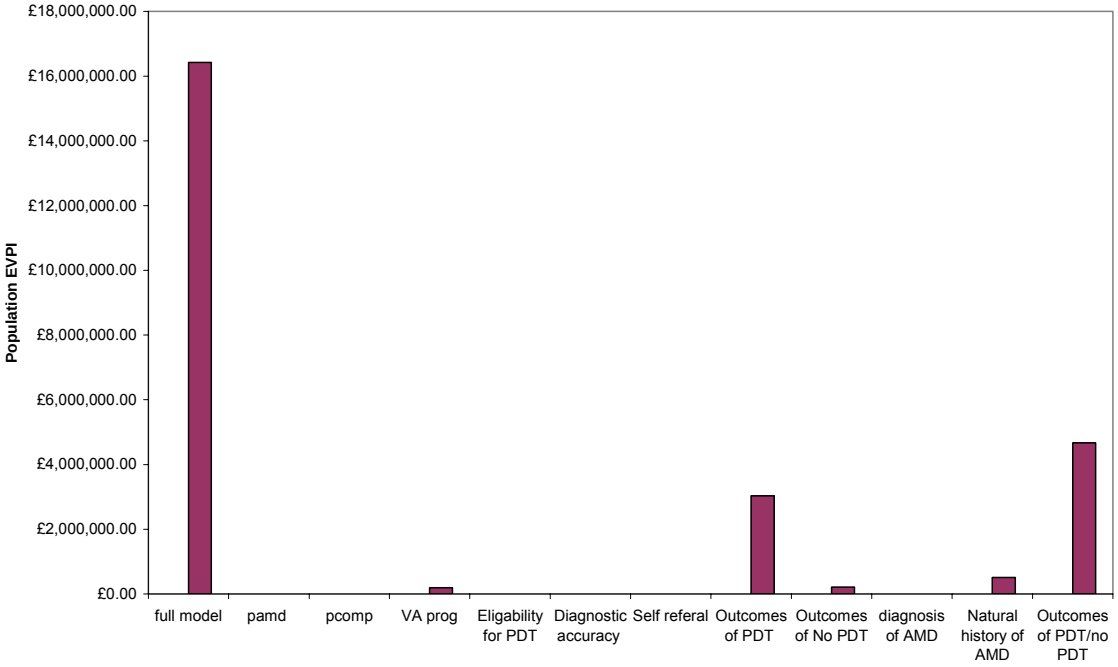


Figure 2.6: EVPI for model inputs: starting visual acuity of 20/80
(Lifespan: 10 years. Threshold: £30,000)

2.3.3 Alternative structural assumptions

Two alternative assumptions regarding the effect of the Amsler grid were explored:

Scenario 1: The Amsler grid provides no additional benefit, in terms of identifying those patients with AMD after a loss in visual acuity
Scenario 2: All patients' will self refer after a loss in visual acuity (1 line)

The results of the 2 models using these alternative assumptions are presented in Tables 2.3-2.6 below.

Table 2.3: Costs and outcomes for model assuming no additional effect of the Amsler grid after a loss of visual acuity

	Mean QALYs	**Mean Costs**	**ICER**
Starting visual acuity = 20/40			
Screen and treat	1.2128	£3,650	£12,828
No screen and treat	0.9816	£2,637	E dominated
No screen and no treatment	0.9359	£98	
Starting visual acuity = 20/80			
Screen and treat	1.0885	£3,653	£17,855
No screen and treat	0.9202	£2,640	E dominated
No screen and no treatment	0.8894	£98	

Table 2.4: Simulation results (£30,000 threshold value) for model assuming no additional effect of the Amsler grid after a loss of visual acuity

	Probability cost-effective	EVPI assuming 10-year lifetime
Starting visual acuity = 20/40		
Screen and treat	0.88	£6,686,984
No screen and treat	0.02	
No screen and no treatment	0.10	
Starting visual acuity = 20/80		
Screen and treat	0.72	£15,865,519
No screen and treat	0.03	
No screen and no treatment	0.25	

Table 2.5: Costs and outcomes for model assuming all patients' will self refer after a loss in visual acuity

	Mean QALYs	Mean Costs	ICER
Starting visual acuity = 20/80			
Screen and treat	1.2183	£3,663	£16, 178
No screen and treat	1.0648	£2,743	E dominated
No screen and no treatment	0.9980	£99	
Screen and treat	1.0908	£3,652	£23,162
No screen and treat	0.9737	£2,739	E dominated
No screen and no treatment	0.9374	£99	

Table 2.6: Simulation results (£30,000 threshold value) for model assuming all patients' will self refer after a loss in visual acuity

	Probability cost-effective	EVPI assuming 10-year lifetime
Starting visual acuity = 20/40		
Screen and treat	0.66	£30,466,013
No screen and treat	0.12	
No screen and no treatment	0.22	
Starting visual acuity = 20/80		
Screen and treat	0.52	£40,179,821
No screen and treat	0.09	
No screen and no treatment	0.39	

For both starting visual acuities changing the assumption of the additive effect of the Amsler grid had little effect on the costs and QALYs. Screening is still regarded as cost-effective when compared to no treatment. This is because the majority of patients are diagnosed through the self-screen before any loss in visual acuity. Decision uncertainty is however sensitive to structural assumptions, with the probability that screen + treat is cost effective reduced from 0.89 in the base case 20/40 model to 0.66 in scenario 2. The population EVPI for these alternative assumptions increases for both scenarios in the 20/40 and 20/80 starting visual acuity models.

2.4 Discussion

2.4.1 Interpretation

Self-screening following 1st eye neovascular AMD appears to be a potentially cost-effective intervention for patients with initial visual acuities ranging from 20/40 to 20/80. However, the cost-effectiveness of self-screening is uncertain and, at a threshold for cost-effectiveness of £30,000 per additional QALY, the value of information surrounding the decision problem is significant, particularly when patients have lower initial visual acuities. The EVPI may exceed the cost of further investigation, which suggests that further research will be potentially cost-effective. The EVPI associated with model inputs indicates that more evidence about the impact of PDT on expected quality of life, and the quality of life for those not treated with PDT,

would be most valuable and would require experimental design. It also suggests that additional evidence about other inputs individually, such as screening accuracy alone, maybe of little value. However, this does not mean that additional information about all the model inputs combined would not be valuable.

2.4.2 Alternative model structures

Structural uncertainty regarding the effect of the Amsler grid was also explored using 2 alternative structural assumptions: Amsler grid provides no additional benefit, in terms of identifying those patients with AMD after a loss in visual acuity or all patients' will self refer after a loss in visual acuity. Although the cost-effectiveness results were not sensitive to structural assumptions the EVPI was particularly sensitive to model specification, it is therefore it is crucial to determine the most appropriate model structure when using VOI analysis to inform research prioritisation.

2.4.3 Caveats

This model has focused on self-screening patients with 1st eye neovascular AMD. A policy of self-screening for AMD before first eye involvement could be considered and could be modelled. However the results of this model and previous analysis in the TAR demonstrates that such a policy would not be cost-effective for a number of reasons: treatment will not be cost-effective in the worse seeing eye as it will not have an impact on overall visual acuity; the very low incidence in this group of patients will generate very large number of false positive results and unnecessary eye examinations; and the gains in quality of life offered by this strategy will be only be realised a number of years in the future for the small number of patients who have treatable neovascular AMD in the fist eye but develop untreatable neovascular AMD in the second eye. Since this strategy will not be cost-effective the decision uncertainty and EVPI surrounding this policy would also be very low. A policy of regular (3 monthly) repeated eye examinations could also be considered but again this analysis indicates that this strategy would also not be cost-effective because it would be very costly (many more negative eye examinations) and is unlikely to be effective as VA can decline rapidly in the period between examinations.

2.4.4. Methodological issues

The value- of information analysis has demonstrated 2 important issues relating to this particular decision problem:

1. The importance of defining an appropriate model structure. Although in this case alternative model structure did not change the adoption decision they did substantially change the value of information estimates generated.

2. It is apparent from this analysis that the scope of the original NICE assessment report[(2)], as dictated by the Department of Health remit, was not wide enough, and should have included a component to identify patients with AMD.

2.4.5. Relationship with the research recommendations provided by NICE

The EVPI analysis provided here specifically addresses one of the research recommendations specified by NICE. It is apparent from this analysis that screening a population at risk for AMD can offer benefits in terms of early identification and hence treatment, and that policies to identify patients at risk must be evaluated simultaneously with policies for appropriate treatment.

Appendix. Reference case evaluation: Repeat screening for age related macular degeneration

Element of health technology assessment	Reference case	Criteria met by assessment?	Comment
Defining the decision problem	The scope developed by the Institute	Yes	
Comparator	Alternative therapies routinely used in the NHS	Yes	
Perspective on costs	NHS and PSS	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Synthesis of evidence on outcomes	Based on a systematic review	No	Review of literature we comprehensive but not systematic
Measure of health benefits	Quality adjusted life years (QALYs)	Yes	
Description of health states for calculation of QALYs	Health states described using a standardised and validated generic instrument	No	Outcomes of treatment taken from a previous study, submitted as part of the NICE assessment
Method of preference elicitation for health state valuation	Choice-based method, for example, time trade-off, standard gamble (not rating scale)	No	
Source of preference data	Representative sample of the public	No	
Discount rate	An annual rate of 3.5% on both costs and health effects	No	The discount rates for costs and health outcomes recommended at the time of the NICE assessment were used.
Equity position	An additional QALY has the same weight, regardless of the other characteristics of the individuals receiving the health benefit	Yes	

References

1. Royal College of Ophthalmologists. *Age Related Macular Degeneration:Guid*elines; February 2000.
2. Meads C, Salas C RT, Moore D, Fry-Smith A, C. H. Clinical effectiveness and cost utility of photodynamic therapy for wet age-related macular degeneration. Birmingham: West Midlands Health Technology Assessment Group, University of Birmingham; January 2002,.
3. National Institute for Clinical Excellence. Guidance on the use of photodynamic therapy for age-related macular degeneration: *Technology Appraisal 68*; September 2003.
4. AMD Alliance International. Amsler Grid: *Check your eyesight*. In; 2002.
5. Claxton K, Ginnelly L, Sculpher M, Philips Z, Palmer S. A pilot study on the use of decision theory and value of information analysis as part of the NHS health technology assessment programme.: *Health Technology Assessment*; Forthcoming.
6. Briggs A, Sculpher M. An introduction to markov modelling for economic evaluation. *Pharmacoeconomics* 1998;13:397-409.
7. Smith D, Fenn P, M D. Cost and Effectiveness of Photodynamic Therapy with Verteporfin For Age-Related Macular Degeneration: *The UK Case*; 2002.
8. Gregor Z, Bird AC, IH C. Senile disciform macular degeneration in the second eye. *British Journal of Ophthalmology* 1977;61:141-147.
9. Margherio RR, Margherio AR, DeSantis ME. Laser treatments with verteporfin therapy and its potential impact on retinal practices. *Retina 2000*;20:325-30.
10. Schuchard R. Validity and interpretation of Amsler grid reports. *Archives of Ophthalmology* 1993;111:776-780.
11. Fine AM, Elmen MJ, et al. Earliest symptoms caused by neovascular membranes in the macula. *Archives of ophthalmology* 1986;104:513-514.
12. Office for National Statistics. Census 2001 population in England and Wales; 2002.
13. The Chartered Institute of Public Finance and Accountancy (CIPFA). *Health Service Financial Database 2001*; 2001.
14. Bressler N, M. Photodynamic therapy of subfoveal choroidal neovascularization in age- related macular degeneration with verteporfin: two-year results of 2 randomized clinical trials-tap report 2. *Archives of Ophthalmology* 2001;119:198-207.
15. Brown M, Brown G, Sharma S, Kistlet J, Brown H. Utility values associated with blindness in an adult population. *British Journal of Ophthalmology* 2001;85:327-331.
16. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000;17:479-500.
17. Johannesson. M, Weinstein. MC. On the decision rules of cost-effectiveness analysis. *Journal of Health Economics* 1993;12:459-467.
18. Van Hout BAA, M. J.; Gordon, G. S., et al. . Costs, effects and c/e-ratios alongside a clinical trial. *Health Economics* 1994;3:309-19.
19. Fenwick. E, Claxton. K, Sculpher. M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics* 2001;10:779-89.
20. Claxton K. The irrelevance of inference: a decision making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics* 1999;18:341-364.
21. Fenwick E, Claxton K, Sculpher M, Briggs A. Improving the efficiency and relevance of health technology assessment: the role of decision analytic modelling. Centre for Health Economics *Discussion Paper No 179*; 2000.

3. Glycoprotein IIb/IIIa antagonists for the treatment of acute coronary syndromes

3.1 Background

3.1.1 Condition and technology

Acute Coronary Syndromes (ACS) is a term that includes a range of patients with a similar underlying pathology. At one end of the spectrum are those patients with evidence of ST elevation on a resting electrocardiogram (ECG) who are eligible for treatment with thrombolysis and who may subsequently develop Q-wave on their ECG – this is a full myocardial infarction (MI). At the other, are patients who are classified as having either unstable angina or non-Q-wave MI. Non-Q-wave MI is the term used when the cardiac enzymes are elevated to the range indicating that MI has occurred, but a Q-wave does not develop on ECG tracings. Unstable angina represents a spectrum of clinical states that fall between stable angina and acute MI and includes new onset angina and angina occurring >24 hours post-MI.

Not only is unstable angina an unspecific diagnostic category, but patients present with varying degrees of atherosclerosis (stenosis size, location and plaque fragility), thrombus formation (low or high platelet content) and vasospasm. Each of these contributes to the morbidity and mortality of the disease and, therefore, represents a potential target for intervention with medical therapy. Aspirin and heparin are typically used to reduce thrombus formation, and nitrates are used to help reduce vasospasm and cardiac oxygen requirements. Interventional therapy typically involves percutaneous coronary intervention (PCI), such as angioplasty, or coronary artery bypass surgery. NHS Hospital Episode Statistics suggest the incidence of unstable angina is around 1000 cases per million total population per year, or about 10 per acute hospital per week [1].

Glycoprotein IIb/IIIa antagonists (GPA) are a class of drugs to prevent platelet aggregation in the acute treatment of patients with non-ST-elevation ACS. The aim of these drugs is to reduce the risk of cardiac death and acute MI. Two broad groups of GPA are licensed in the UK: Abciximab (ReoPro®, Eli Lilly) is a monoclonal antibody targeted at the receptor (also known as a 'large molecule' GPA); while Eptifabtide (Intergrilin®, Schering Plough) and Tirofiban (Aggrastat®, MSD) are more conventional pharmacological receptor antagonists (also known as 'small molecule' GPAs).

GPAs are used in two general ways to manage ACS patients. Firstly, as an adjunct to PCI; for those patients who undergo such a procedure, Abciximab is the GPA which is mainly used for this purpose. GPAs can be also used as a form of medical management for non-ST-elevation ACS patients regardless of whether or not they subsequently go on to have a PCI. Tirofiban and Eptifabtide are mainly used for this indication.

3.1.2 Technology Assessment Review

In 2000, NICE commissioned two rapid reviews on the use of GPAs in cardiology. One focussed on the role of GPAs in ACS [2]; the other considered the use of GPAs alongside PCI [3]. Whilst the first of these reviews concluded that there was a small benefit in the use of GPAs as part of medical management for non ST elevation ACS, the second demonstrated a consistent benefit in the use of GPAs alongside PCI. In 2002 a systematic review update was commissioned [4]. The main conclusions relating to clinical effectiveness were that:

- The effectiveness of GPAs as adjuncts to PCI was confirmed further.
- Evidence for the use of GPAs in situations where PCI is not undertaken (i.e. alongside medical management) was weakened.
- There is no evidence for the clinical superiority of Tirofiban or Eptifabtide over Abciximab.

In terms of the cost effectiveness of the use of GPAs in the UK, there were serious limitations to the published evidence [5]. For example, the majority of RCTs evaluating the efficacy of GPAs had been undertaken outside the UK where clinical practice is quite different; the RCTs had short follow-up, typically of just 30 days or 6 months and none had directly compared the various ways in which GPAs could be used in ACS patients in the UK. Because of this, a separate model was developed to assess the long term cost effectiveness of the use of GPAs in the UK NHS. The model evaluated three GPA-based strategies in comparison with usual care:

- GPAs as part of initial medical management (Strategy 1);
- GPAs in patients with planned PCI where GPAs are started once a decision to undertake PCI has been made (Strategy 2);
- GPAs as an adjunct to PCI where the agent is used at the time of PCI or is started up to 1 hour before the procedure (Strategy 3); and
- No use of GPAs (Strategy 4).

The conclusion of the report was that GPAs used as part of medical management (Strategy 1) was the most cost effective use of resources, with ICERs ranging from £4,605 to £11,671 (in comparison with usual care (Strategy 4)). The strategy of using GPAs as an adjunct to PCI was found to be economically inferior to medical management under all scenarios. Further details of the model are provided in Section 3.2.

3.1.3 NICE guidance

In September 2002, NICE issued guidance on the use of GPAs for the treatment of ACS [6]. The guidance recommended the following:

1. That GPAs should be considered as part of medical management for patients with unstable angina or non ST elevation MI and that the management pathway should also include other pharmacological interventions and, where appropriate, early coronary angioplasty with a view to revascularisation either by PCI or coronary artery bypasses graft surgery.
2. Intravenous use of GPAs is recommended as part of initial medical management in patients who are at high risk of subsequent MI or death.
3. Where PCI does not occur or is not immediately available, medical management with GPAs is still recommended.
4. In determining risk clinicians should take into account a combination of risk factors including clinical investigations such as ECG changes.
5. Cardiac troponin testing is useful for diagnosing ACS, but that GPA treatment can be initiated before the results of a Troponin test is known.
6. If PCI is indicated but is delayed beyond the medical management phase, GPAs are recommended as an adjunct to the PCI.
7. GPAs should be considered as an adjunct to PCI for all patients with diabetes and for those patients undergoing complex procedures. In procedurally uncomplicated elective PCI, where the risk of adverse sequalae is low, the use of GPAs is not recommended.
8. GPAs are not currently licensed for use as an adjunct to thrombolytic therapy.

3.1.4 Research recommendations

The recommendations for further research, issued as part of the NICE guidance were that:

- The effects of GPAs in current UK practice should be investigated in carefully designed research to assess their benefits in non-ST elevation ACS patients who are not scheduled for PCI.
- Research should be carried out to investigate the efficacy of GPAs in subgroups such as women.

- The results of the CURE trial may lead to a consideration of the use of Clopidogrel for the management of patients with ACS. Research is required to establish the relative roles of GPAs and Clopidogrel in the short term management of ACS patients.
- Research is required to establish the statistical relationship between clinical risk factors and Troponin levels.

3.2 Methods

3.2.1 Model structure

The model developed as part of the assessment review process, and used here, took the perspective of the UK NHS for costing and adopted a lifetime horizon. Health outcomes were measured in terms of QALYs. It consisted of a short term component and a long term extrapolation.

Short term model
The short term component of the model characterised the period up to 6 months following an episode of ACS. Three mutually exclusive outcomes were modelled: non-fatal myocardial infarction (MI), death and ischemic heart disease (IHD) without MI during the short term period. Figure 3.1 provides the short term model structure.

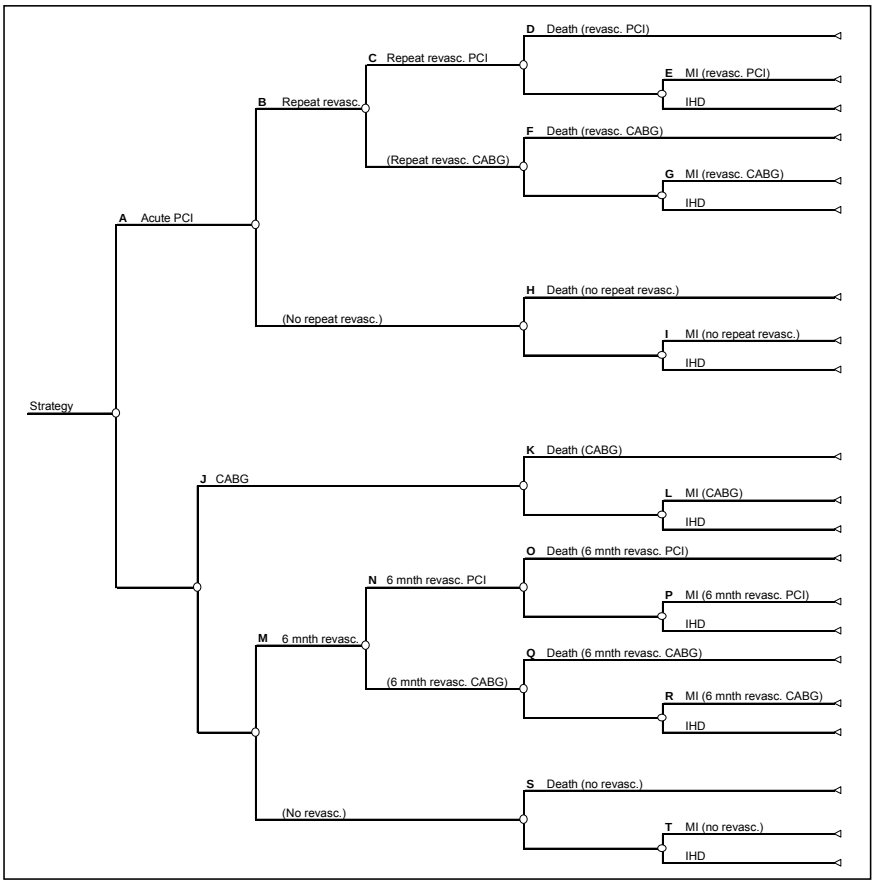


Figure 3.1 Short term component of the GPA Model

Baseline event rates were taken from a UK source, PRAIS-UK. This is an observational cohort registry of 1046 patients admitted to 56 UK hospitals with acute coronary syndromes during 1998-9 [7]. These data provided the path probabilities for Strategy 4 (see Table *.1) and the resource use estimates. Further details regarding the input data can be found in the main modelling report [5].

Table 3.1 Baseline probabilities used in the short term model (The node refers to the labels in Figure 3.1)

Node	Description	Probability	***Parameters of the beta distribution***	
			α	β
A	Acute PCI	0.05	53	980
B	Repeat revasc.	0.048	8	157
C	Repeat revasc. PCI	1.00	-	-
D	Death (revasc. PCI)	0.00	0.01	7.99
E	MI (revasc. PCI)	0.13	1	7
F	Death (revasc. CABG)	0.00	-	-
G	MI (revasc. CABG)	0.00	-	-
H	Death (no repeat revasc.)	0.03	5	152
I	MI (no repeat revasc.)	0.03	5	147
J	CABG	0.05	47	933
K	Death (CABG)	0.11	5	42
L	MI (CABG)	0.07	3	39
M	6 month revasc	0.05	48	885
N	6 month revasc PCI	0.48	23	25
O	Death (6 month revasc. PCI)	0.09	2	21
P	MI (6-month revasc. PCI)	0.10	2	19
Q	Death (6-month revasc. CABG)	0.00	0.01	24.99
R	MI (6-month revasc. CABG)	0.16	4	21
S	Death (no revasc.)	0.08	68	817
T	MI (no revasc.)	0.05	40	777
	Baseline risk of gastrointestinal bleeding:			
	(i) Undergoing PCI in acute period	0.00	0.01	52.99
	(ii) Undergoing CABG in acute period	0.02	1	46
	(iii) No initial revasc.	0.01	12	921

To model the effect of GPAs (Strategies 1-3), the baseline event probabilities were augmented using the relative risks associated with GPAs (compared to standard care). For the base case model, the estimates of the relative risk of events for each strategy, that were incorporated into the model, were based on a random effects Meta analyses of all of the available trial evidence relevant for each strategy. This required three key considerations:

- Firstly, it was necessary to address the question of whether the relative risks associated with GPAs, which were estimated within the clinical trials, should be adjusted to reflect differences in UK practice. To inform this decision, a meta-regression analysis was undertaken to establish whether, across published trials and taking each strategy separately, the relative risk in a trial was related to the absolute baseline risk in that study. No statistically significant association was found, but this may reflect the small number of trials in the analysis. For this reason, the relative risks from the trials were incorporated into the model without adjustment, which is equivalent to assuming that relative risks are transportable across health care systems whilst the baseline risks in those studies are not.

- Secondly, not all trials reported their end-points over the required period of follow-up (i.e. six months); some simply reported end-points at 30 days. In the base-case analysis, in the absence of 6-month data, the relative risk reductions reported at 30 days were assumed to apply at 6 months. The use of an alternative assumption was explored whereby 30-day relative risks were extrapolated to six months assuming a constant hazard ratio. This produced very similar results to the assumption of

constant relative risks, so the latter was used in the base-case analysis due to its relative simplicity.

- Finally, a further complexity of the trial evidence was that the three GPAs licensed in the UK had been used in more than one way in the trials. This was the case despite the fact that the drugs are licensed for more specific purposes (see above). Despite this mismatch it was decided to include all trials in the meta-analyses. The principle behind this was that the most reliable estimate of overall treatment effects of GPAs would be generated by using as much experimental evidence as was available in ACS patients.

Table 3.2 details the pooled relative risk estimates used in the base case model. Further details relating to the individual trials from which the pooled estimates are derived can be found in the modelling report [5].

Table 3.2 Relative risk data incorporated into the model

	Relative risks (95% confidence intervals)					
	Non-fatal MI	Death	Revasc[a]	PCI[a]	CABG[a]	GI Bleed
S1	0.94 (0.87, 1.02)	0.81 (0.70, 0.93)	0.94 (0.87, 1.02)	0.97 (0.91, 1.03)	1.00 (0.94, 1.06)	1.37 (1.00, 1.87)
S2	0.70 (0.48, 1.03)	1.22 (0.61, 2.46)	1.02 (0.84, 1.24)	1.04 (0.84, 1.29)	0.76 (0.49, 1.17)	2.02 (1.02, 4.00)
S3	0.67 (0.57, 0.79)	0.77 (0.57, 1.05)	0.87 (0.76, 0.98)	0.82 (0.73, 0.93)	0.87 (0.72, 1.05)	1.29 (0.92, 1.91)

S: Strategy. a: Repeat revascularisation rate for Strategies 2 and 3

Long term extrapolation

If patients survived the 6 months following ACS, their long-term costs and QALYs were estimated using a simple 4-State Markov model populated using probability and resource use data from two cohorts of the Nottingham Heart Attack Register (NHAR) [8, 9]. Figure 3.2 details the structure of the model.

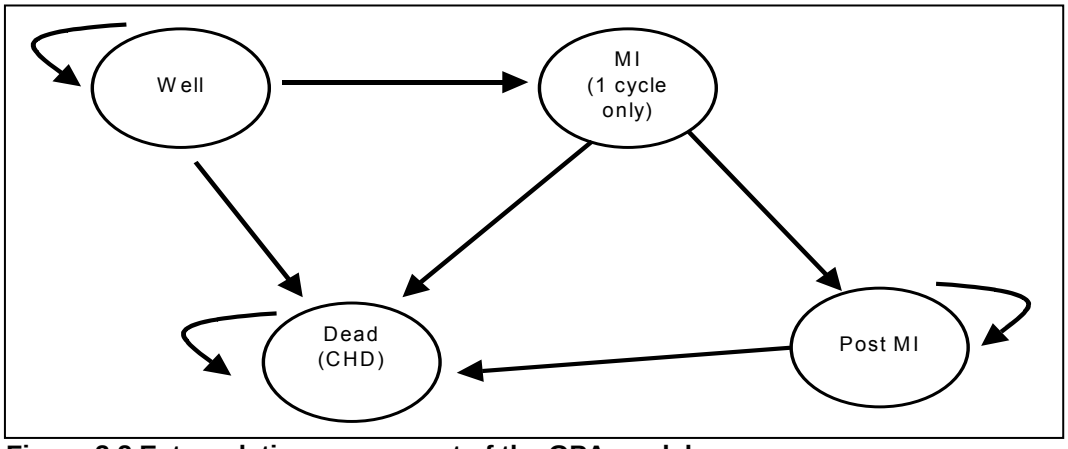


Figure 3.2 Extrapolation component of the GPA model

The cycle time of the model was 1 year. Transition probabilities were estimated using survival analysis techniques, and costs were based on the mean annual resource use within the most recent cohort of the NHAR. Table 3.3 details the transition probabilities used in the model (+ 95% confidence intervals) and the mean (+ SD) annual cost associated with each Health State. In addition to these costs, a transition cost was applied when patients moved into the 'Death' State. Future costs were discounted at a rate of 6%, and benefits at 2%. The quality adjustment of life years was undertaken assuming a single utility for all living patients based on data in the published literature [10]. The value used in the model was 0.8 with a standard deviation of 0.09.

Table 3.3 Transition Probabilities

From state:	To state: IHD	Non-fatal MI	Post-MI	Dead
IHD £1421 (£944)[a]	0.9049 (0.8896, 0.9186)	0.0186 (0.0133, 0.0254)	-	0.0765 (0.0643, 0.0904)
Non-fatal MI £3966 (£1722) [a]	-	-	0.7900 (0.7177, 0.8471)	0.2100 (0.1529, 0.2822)
Post-MI £1587 (£1091) [a]	-	-	0.9266 (0.9024, 0.9466)	0.0734 (0.0534, 0.0976)
Dead	-	-	-	1 -

[a] Derived from the Monte Carlo simulation

3.2.2 Probabilistic analysis

The model was fully probabilistic and incorporated distributions to reflect the uncertainty associated with the input data. Baseline event probabilities were modelled as Beta distributions. For resource use estimates, the probability of a particular resource use was characterised by a beta distribution, and length of stay data were incorporated as lognormal distributions. Relative risks, transition probabilities and utility estimates were also incorporated as lognormal distributions.

3.2.3 Sensitivity analyses

Detailed sensitivity analyses were undertaken in three key areas in order to determine the robustness of the base case model. These related to:

- Variations in the sources of data used to populate the base-case model;
- Variations in the baseline event rates using non-UK specific sources of data;
- The inclusion of additional strategies to those considered in the base-case model.

However, for the purpose of this report, we focus on just two sensitivity analyses to illustrate two methodological issues. The first relates to the use of available evidence. This is demonstrated by focusing on the data used to establish the relative risks associated with GPAs. The second focuses on the issue of appropriate comparators; this is demonstrated by including an additional strategy to the model. Specifically, two alternative scenarios are reported here:

1. In establishing the relative risk of events in the short term model, only data from trials reporting at six-months were incorporated into the Meta analyses.
2. Clopidogrel was considered as an additional Strategy, in addition to the three GPA-based strategies and usual care. Relative risk data for Clopidogrel was taken from the published results of the CURE trial [11].

The results of the remainder of the sensitivity analyses can be found in the modelling report [5].

3.2.4 Value of information analysis

Value of information analysis was undertaken for the base case model and for the two scenario analyses described above. In each case, the EVPI was estimated for the full model and for groups of parameters within it. For groups of parameters with positive EVPI, the EVPI on individual parameters was estimated. The parameters were grouped as follows:

- Baseline probabilities
- Relative risk data for each Strategy under consideration
- Short term costs
- Long term net benefits (incorporating QALYs and long term costs)

As described above, NHS Hospital Episode Statistics suggest the incidence of UA is around 1000 cases per million total population per year. Based on current estimates of the UK population, this implies an annual incidence of around 59,756 [12]. This figure was used to estimate the population EVPI for this decision problem. A 6% annual rate of discount was applied.

3.3 Results

3.3.1 Adoption decisions

Table 3.4 details the expected cost and QALYs, and ICERs for the base case model and for the two scenarios modelled under sensitivity analyses. ICERs were estimated using standard decision rules [13]. In all cases, strategy 2 is ruled out because it is more costly and less effective than Strategy 3 and Strategy 3 is ruled out because of extended dominance. Moreover, in the second sensitivity analysis, Clopidogrel is also ruled out due to extended dominance. Strategy 1 (GPAs used as part of medical management), results in an ICER of between £5,736 and £8,750 compared with Strategy 4 (usual care)

Table 3.4 Expected costs and QALYs for each Strategy under alternative scenarios

Strategy	Average Cost	Average QALYs	ICER
Base case model			
1	£12,688	7.7875	£5,736[a]
2	£12,207	7.6839	D
3	£12,188	7.6910	ED (£25,556 [b])
4	£12,119	7.6883	
Scenario 1: Relative risk data taken from 6-month trials only			
1	£12,611	7.7352	£8,750[a]
2	£12,194	7.6725	D
3	£12,179	7.6787	ED (£41,579 [b])
4	£12,100	7.6768	
Scenario 2: Clopidogrel as a fifth strategy			
1	£12,790	7.7630	£5,769[a]
5	£12,594	7.7173	ED (£7,026 [c])
2	£12,307	7.6591	D
3	£12,287	7.6662	ED (£26,296 [b])
4	£12,216	7.6635	

D: Dominated ED: Extended dominance
[a] ICER Strategy 1 versus Strategy 4 [b] ICER Strategy 3 versus Strategy 4 [c] ICER Strategy 5 versus Strategy 4

Figure 3.3 details the cost-effectiveness acceptability curves (CEACs) and the cost effectiveness frontier for the base case model. These incorporate the uncertainty within the model in relation to both the estimates of expected cost and QALYs, and the threshold for cost-effectiveness (maximum willingness to pay for an additional QALY). The CEACs detail the probability that each Strategy is cost-effective over a range of threshold values, and the frontier details the probability that the optimum strategy is cost effective. Strategy 4 has the highest probability of cost-effectiveness at low values of lambda. At values over £5,738, Strategy 1 is optimum. Consequently, the results from the base-case analysis demonstrate that if the health service is prepared to pay over £5,738 per QALY then Strategy 1 is always the optimal decision. Similar patterns result from the two scenarios reported as sensitivity analyses.

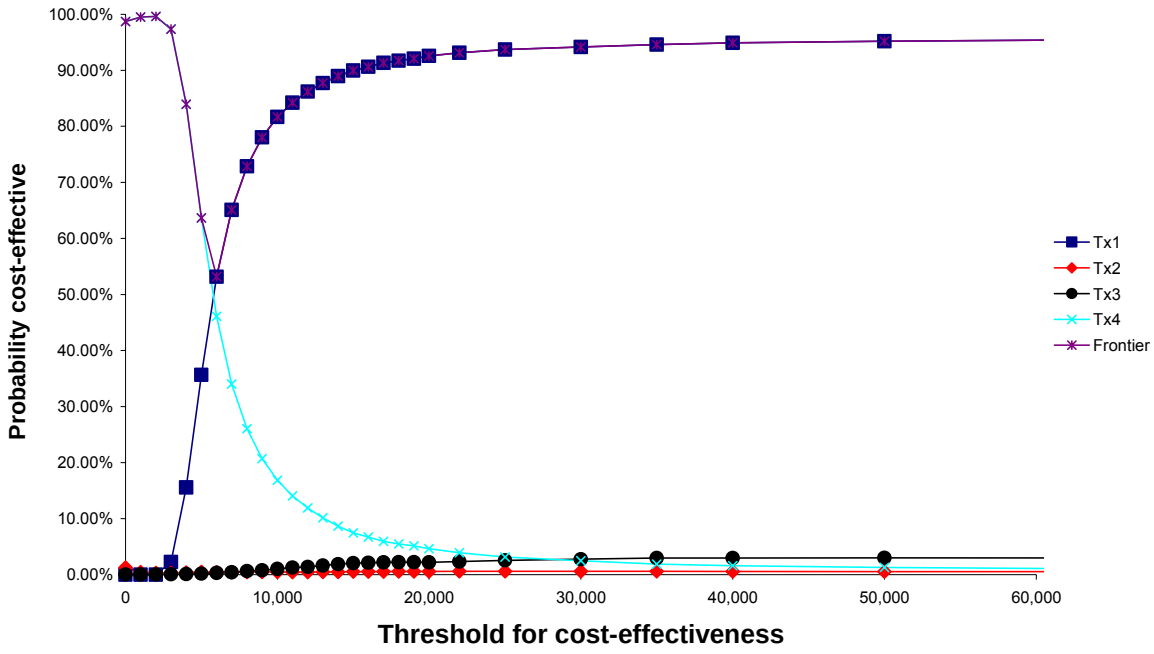


Figure 3.3 CEACs and Frontier for the base case model

3.3.2 Research recommendations

EVPI is between £42.97 and £57.81 per patient for threshold values between £10,000 and £50,000. The population EVPI for the base case model ranges from £11,464,710 (5 years, threshold = 30K) to £35,561,534 (15 years, threshold = 50K) depending on assumptions regarding the lifespan of the technology and the value of lambda (see Figure 3.4). If the lifespan of the technology is ten years and the threshold for cost-effectiveness is £30,000, the EVPI for the full model is £20,031,809 (EVPI per episode at a threshold of £30,000 is £42.97).

Figure 3.5 details the base case EVPI results for groups of parameters, assuming a ten year life span and that the threshold is £30,000. All of the uncertainty in the model is encapsulated within the relative risks associated with Strategy 1. Isolating the EVPI for the individual relative risk parameters identified that the relative risk of death for patients not undergoing an acute PCI procedure was the only parameter associated with positive EVPI. This does not imply that there is no value in obtaining further information on other parameters; it means that there is no value on obtaining further information on those parameters individually. In other words, whilst there may be value in obtaining information on all of the model parameters combined (as together there is positive EVPI for the decision), there is no value in obtaining individual information on those model inputs that have zero EVPI. For example, there is no value in obtaining further information on the relative risk of death in patients who receive, or who are scheduled to receive, GPAs at the time of a PCI procedure (i.e. Strategies 2 and 3) if these data are obtained separately from further information associated with the relative risk of death in patients who receive GPAs as part of medical management who do not undergo an acute PCI procedure (i.e. Strategy 1).

The pattern is the same across all assumed lifespan of the technology. When the threshold for cost-effectiveness is lowered to £10,000, there is some positive EVPI is the parameters associated with long term outcomes, but these are small compared with the relative risk data (e.g. for a lifespan of 10 years and threshold at £10,000, the EVPI for the long term net benefit is £480,373 compared with a value of £20,047,419 for the relative risks associated with Strategy 1).

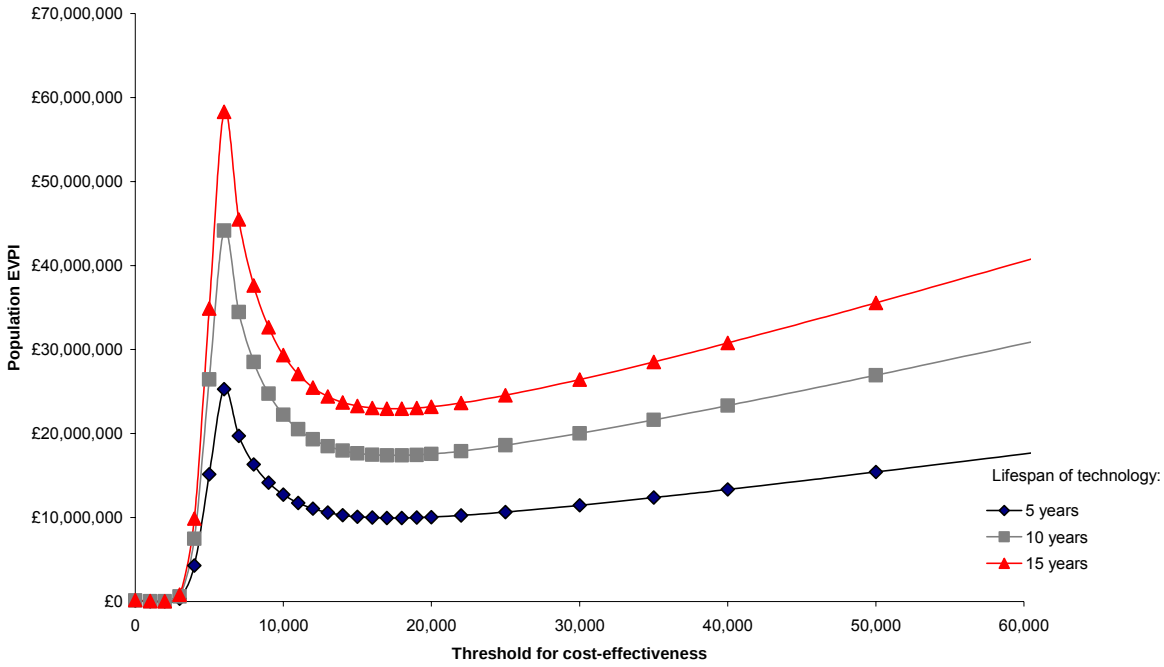
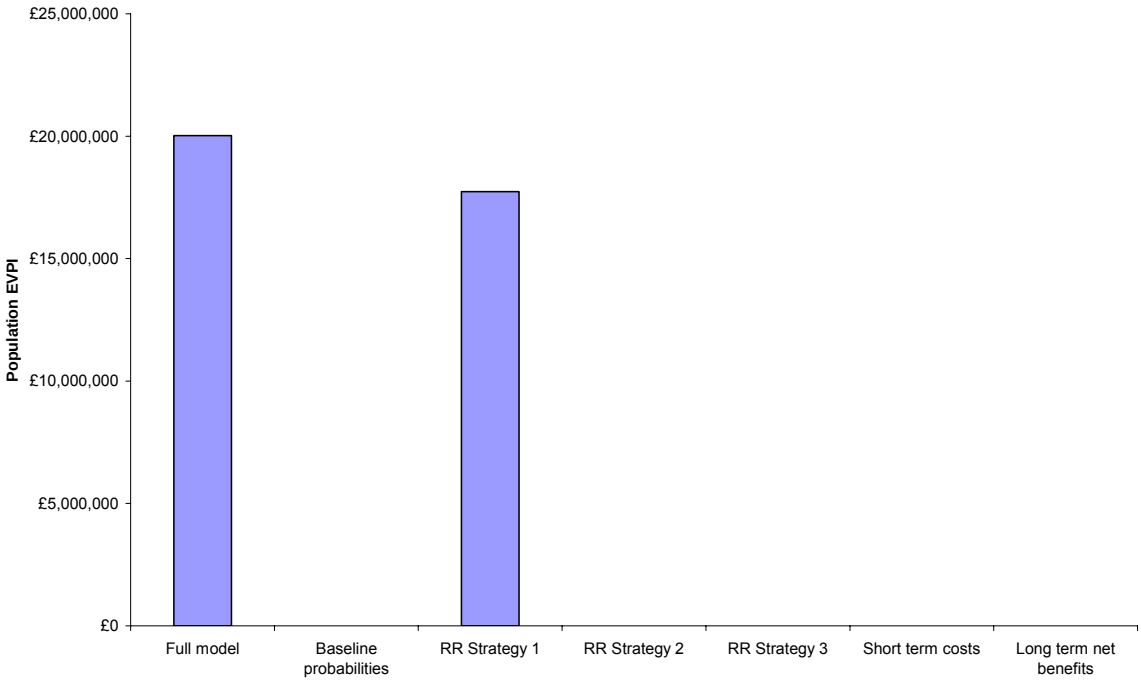


Figure 3.4 Population EVPI for the base case model

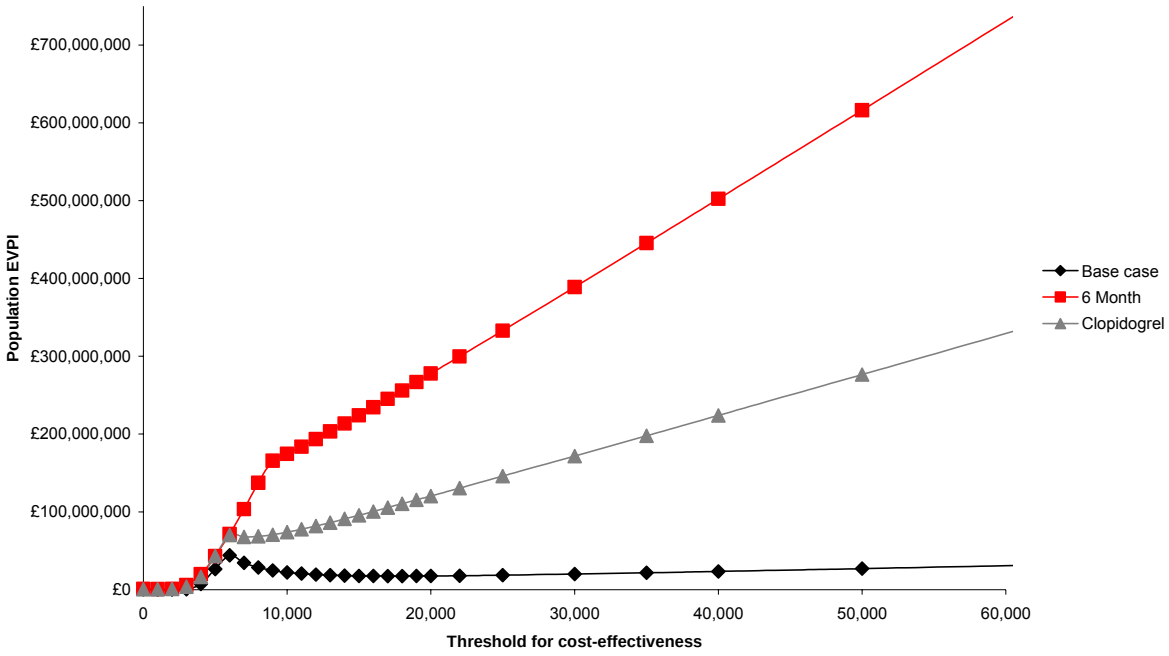


(Lifespan: 10 years. Threshold: £30,000)

Figure 3.5 EVPI for groups of parameters in the base case model

3.3.3 Scenario analyses

Figure 3.6 details the total population EVPI for the base case model and the two scenario analyses over a range of values for lambda, under an assumed lifespan for the technology of ten years. The total EVPI increases in both of the scenarios investigated though sensitivity analysis. This is because, in both cases, the probability that the optimum strategy is cost effective falls in comparison with the base case model. For example, under Scenario 1, where the relative risk data are derived from a smaller number of trials, the probability that the optimum strategy is cost-effective is below 0.7 for all Lambda values above the ICER.



Lifespan of technology: 10 years

Figure 3.6 Population EVPI under alternative assumptions

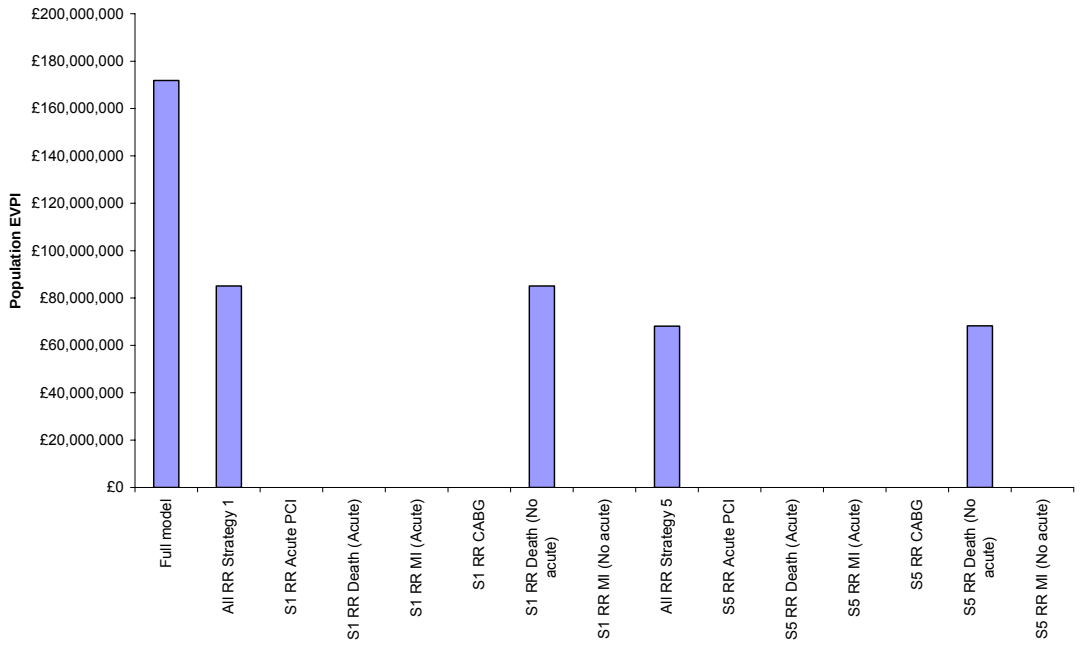
Table 3.5 details the EVPI values (assuming a lifespan of 10 years and a threshold of £30,000) for the sensitivity analysis which used only those trials reporting at 6 months to derive the relative risk estimates for the three GPA strategies. The parameters with positive EVPI follow the same pattern as the base case model, but the value of the decision uncertainty increases 8-fold. The EVPI for the relative risk of death in patients not undergoing an acute PCI under Strategy 1 is marginally greater that the EVPI for all of the relative risks for Strategy 1 combined. This is likely to be due to small correlations between values of the input variables, occurring as a result of the Monte Carlo process.

Table 3.5 EVPI under alternative relative risk estimate assumptions

	Episode EVPI	Population EVPI
Full model	834.33	388,963,186
Groups of parameters		
Baseline probabilities	0	0
Relative risks S1	799.13	372,553,107
Relative risks S2	0	0
Relative risks S3	0	0
Short term costs	0	0
Long term net benefits	1.82	848,934
Individual parameters (S1)		
RR PCI	0	0
RR Death (acute)	0	0
RR MI (acute)	0	0
RR CABG	0	0
RR Death (non acute)	799.95	372,934,171
RR MI (non acute)	0	0

Lifespan: 10 years. Threshold: £30,000

Under the second Scenario analysis, the scope of the model is widened to include an additional strategy. When Clopidogrel is evaluated alongside the use of GPAs, although the optimal decision remains the same (i.e. the use of GPAs as medical management), the uncertainty associated with that decision increases, hence EVPI increases (at an assumed lifespan of 10 years and a threshold of £30,000, the population EVPI is £171,382,025). Figure 3.7 details the population level EVPI values for the full model, groups of parameters and individual parameters. Both the relative risks associated with Strategy 1 (GPAs as medical management) and Strategy 5 (Clopidogrel) are associated with positive EVPI (RR S1: £85,041,201, RR S5: £68,136,978), and for both cases, it is the relative risk associated with death in patients who do not undergo an acute PCI that drives the uncertainty in the model. As with the first scenario analysis, the EVPI on the individual parameters is marginally greater than on the EVPI for the group of parameters containing it.



(Lifespan: 10 years. Threshold: £30,000)

Figure 3.7 The effect of the inclusion of an additional strategy on EVPI values

3.4 Discussion

3.4.1 Interpretation

The EVPI for the base case model indicates there is potentially considerable value in further research to reduce the uncertainty associated with the a-priori decision. Using base-case assumptions, EVPI is between £47.71 and £57.81 per patient for threshold values between £10,000 and £50,000. Translating this to a population figure, the EVPI is between £20 million and £26.9 million, assuming the lifespan of the technology is ten years.

EVPI is driven almost exclusively by the relative risk of death in patients not undergoing an initial PCI in Strategy 1. This would suggest that future research should be directed toward reducing the uncertainty associated with the relative risk of death in patients who are prescribed GPAs who have ACS and do not undergo a PCI procedure in the acute phase.

When the data used to derive relative risk estimates are restricted to those trials reporting at 6 months, the uncertainty associated with the optimal decision increases, increasing the EVPI 8-fold. This is not surprising; by using fewer data to derive relative risk estimates, the uncertainty associated with those estimates will inevitably increase due to the smaller number of trials included in the Meta analyses. However, this result places further emphasis on the need to obtain further research evidence on the risk reduction associated with the use of GPAs as part of medical management, for patients who do no undergo an acute PCI. Given that the data used in this model suggested that 95% of patients admitted to hospital with acute coronary syndrome do no undergo a PCI procedure, this represents a large proportion of the patient group.

When the scope of the model is widened to include Clopidogrel as a treatment option Strategy 1 remains the optimal decision but EVPI increases. This is because there are only small differences in cost and outcome between the Clopidogrel option and Strategy 1. Although Strategy 1 remains the optimal decision, there is around a one third probability that Clopidogrel is cost-effective. This uncertainty explains the high EVPI. The results of this scenario, suggest that further research to identify the relative benefits of Clopidogrel and GPAs as part of medical management, and compared with the current service provisions would be of benefit.

3.4.2 Caveats

Although it is clear from the partial EVPI analysis that the relative risk of death in patients not undergoing an acute PCI who are given GPAs (or Clopidogrel) as part of their medical management is the key parameter driving decision uncertainty, as cautionary note is required regarding the way in which risk was handled within the model. As described in the introduction, ACS includes a range of patients with important different characteristics which are likely to affect prognosis. As such, the trials from which the relative risk data are derived include patients with a variety of characteristics. For example, the medical management trials, which provide the relative risks for Strategy 1 in the model, include patients of a variety of ages, with and without ST depression and with and without troponin positivity[1]. Moreover, the PRAIS-UK data, which provided the basline risks, and the NHAR data, which was used in the long-term extrapolation included patients with a variety of different risk factors. The model therefore reflects an 'average risk' patient across all risk groups. This hetrogeniety must be bourne in mind in interpreting the EVPI results as it is likely that the value of further research will be assocaited with the use of GPAs in patients who are defined, a priori, as high risk.

A further consideration relates to the way in which the PRAIS-UK data reflects current practice in the UK. Although these data do provide the best estimates of current practice, the rate of PCI is on the increase in the UK. At present, the model assumes that, on average, 5% of patients have an acute PCI. Consequently it is the relative risk of death in patients not undergoing PCI (i.e. the remaining 95% of patients) driving the uncertainty that is associated

[1] These are examples of risk factors for ACS patients.

with Strategy 1. As the proportion of patients receiving acute PCI increases, there may be value in obtaining further information on both acute and non acute PCI patients.
Finally, it must be noted that the data used to model the relative cost-effectiveness of Clopidogrel was not identified systematically and were taken from a single trial. However, this analysis was undertaken to illustrate a particular methodological issue as opposed to establishing cost-effectiveness *per se*.

3.4.3 Methodological issues

Notwithstanding the caveats outlined above, the value-of-information analysis carried out here has demonstrated two important points:

1. That the data incorporated into a model has to reflect all available evidence. If selected data are omitted as inputs into the model, then the resultant EVPI statistics will be artificially high. This is shown in this chapter when the EVPI statistics associated with Scenario 1 are compared with the base case model results. To exclude evidence will result in an overestimate of the value of additional research and will ultimately bias research priority setting.

2. The original scope of the model did not take account of the use of Clopidogrel as an alternative treatment for these patients. When this is included, whilst the optimal decision doesn't change, the decision uncertainty does. Therefore, it is crucial that the scope of the work is defined appropriately from the outset, as this will have massive implications for the results of the EVPI analysis, and hence the need for future research in the area.

3.4.4 Relationship with the research recommendations provided by NICE

The NICE guidance provided four specific future research recommendations and the EVPI analysis provided here lends support to two of these. From this analysis it is clear that there is a need to establish the benefits of GPAs in patients not scheduled for PCI and that in establishing the optimal use of resources for the treatment of ACS patients, competing interventions, such as GPAs versus Clopidogrel, should be evaluated together. As described above, a limitation of this model was the way in which risk was incorporated. For this reason, the recommendations relating to the identification of risk factors and the efficacy GPAs in specific subgroups cannot be explicitly addressed here.

Appendix. Reference case evaluation: Glycoprotein IIb/IIIa antagonists for the treatment of Acute Coronary Syndromes

Element of health technology assessment	Reference case	Criteria met by assessment?	Comment
Defining the decision problem	The scope developed by the Institute	Yes	
Comparator	Alternative therapies routinely used in the NHS	Yes	
Perspective on costs	NHS and PSS	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Synthesis of evidence on outcomes	Based on a systematic review	Yes	Systematic review undertaken previously on behalf of NICE[4].
Measure of health benefits	Quality adjusted life years (QALYs)	Yes	
Description of health states for calculation of QALYs	Health states described using a standardised and validated generic instrument	No	Health state preference values were taken from a published source [10].
Method of preference elicitation for health state valuation	Choice-based method, for example, time trade-off, standard gamble (not rating scale)	No	
Source of preference data	Representative sample of the public	No	
Discount rate	An annual rate of 3.5% on both costs and health effects	No	The discount rates for costs and health outcomes recommended at the time of the NICE assessment were used.
Equity position	An additional QALY has the same weight, regardless of the other characteristics of the individuals receiving the health benefit	Yes	

References

1. Department of Health, *Hospital Episode Statistics*. 1998, OPCS: London.
2. McDough, et al., *A rapid and systematic review of the clinical effectiveness and cost effectiveness of glycoprotein IIb/IIIa antagonists in the medical management of unstable angina.* Health Technology Assessment, 2000. 4(30).
3. Fischer, A. and e. al., *Clinical and cost effectiveness of glycoprotein IIb/IIIa inhibitors in association with Percutaneous Coronary Intervention (PCI).* 2000, National Institute for Clinical Excellence: London.
4. Robinson, M., et al., *A systematic review update of the clinical effectiveness and cost effectiveness of glycoprotein IIb/IIIa antagonists.* Health Technology Assessment, 2002. 6(25).
5. Palmer, S., et al., *A cost effectiveness model comparing alternative management strategies for the use of glycoprotein IIb/IIIa antagonists in non ST elevation acute coronary syndrome.* 2002, Report commission by NICE.
6. NICE, *Technology Appraisal Guidance - No 47. Guidance on the use of Glycoprotein IIb/IIIa inhibitors in the treatment of acute coronary syndromes*. September 2002, National Institute for Clinical Excellence: London.
7. Collinson, J., et al., *Clinical outcomes, risk stratification and practice patterns of unstable angina and myocardial infarction without ST elevation: Prospective Registry of Acute Ischaemic Syndromes in the UK (PRAIS-UK).* European Heart Journal, 2001. 21: p. 1450-1457.
8. Gray, D., et al., *Impact of hospital thrombolysis policy on out-of-hospital response to suspected myocardial infarction.* Lancet, 1993. 341: p. 654-657.
9. Brown, N., et al., *Inpatient deaths from acute myocardial infarction, 1982-92: analysis of data in the Nottingham Heart Attach Register.* British Medical Journal, 1997. 315(159-164).
10. Fryback, D., J. Chinnis, and J. Ulviva, *Bayesian cost-effectiveness analysis: an example using the GUSTO trial.* International Journal of Technology Assessment in Health Care, 2001. 17: p. 83-97.
11. The Clopidogrel in Unstable Angina to Prevent Recurrent Events Trial (CURE) Investigators. *Effects of Clopidogrel in addition to aspirin in patients with acute coronary syndromes without ST-segment elevation.* New England Journal of Medicine, 2001. 345(494-502).
12. Office, N.S., *Mid 2000 estimates for UK population. www.statistics.gov.uk.* 2002.
13. Johannsson and Weinstein., *On the decision rules of cost-effectiveness analysis. 1993.* Journal of Health Economics, 1993. 12: p. 459-467.

4. Clopidogrel and modified-release dipyridamole in the secondary prevention of occlusive vascular events

4.1 Background

4.1.1 Condition and technology

It is widely accepted that atherothrombosis is the most important cause of occlusive vascular events (OVEs). The clinical manifestations of atherothrombosis include transient ischaemic attack (TIA), ischaemic stroke, unstable angina, myocardial infarction (MI) and intermittent claudication, a symptom associated with peripheral arterial disease (PAD). Patients with symptomatic disease in one vascular bed are also likely to have diffuse disease, placing them at risk of subsequent events in additional vascular territories. This is demonstrated in individuals with asymptomatic PAD who are twice as likely as normal subjects to suffer from concomitant coronary artery disease.[1]

There are approximately 237,000 myocardial infarctions per year in England and Wales. The annual incidence rate for men aged 30-69 is around 600 per 100,000 population and for women the equivalent rate is 200 per 100,000 population. Estimates from the National Service Frameworks indicate that there are around 110,000 new cases of stroke in England and Wales each year. The economic burden from CHD in terms of direct health care costs and indirect costs (including informal care costs and loss of productivity) is high. Overall the total annual cost of all CHD-related burdens equated to over £7 billion in 1999, the highest of all diseases in the UK for which comparable analyses have been undertaken.[1]

Aspirin is the most widely prescribed antiplatelet agent, and in secondary prevention it reduces the risk of MI, stroke and vascular death by about 25%. Two alternative antiplatelet agents, clopidogrel and modified-release (MR) dipyridamole, are licensed for the secondary prevention of OVEs. Clopidogrel, a thienopyridine antiplatelet drug is unrelated to aspirin and therefore can be used in patients who show a genuine intolerance to aspirin; it is licensed for the secondary prevention of ischaemic stroke, MI and in patients with PAD. Dipyridamole, an adenosine reuptake inhibitor and phosphodiesterase inhibitor, has both antiplatelet and vasodilating properties. Modified-release dipyridamole is licensed for the secondary prevention of ischaemic stroke and TIAs either alone or in combination with aspirin.

4.1.2 Technology Assessment Review

NICE was commissioned by the Department of Health and the Welsh Assembly Government to conduct an appraisal of the clinical effectiveness and cost-effectiveness of two alternative antiplatelet agents, clopidogrel and modified-release dipyridamole, relative to prophylactic doses of aspirin for the secondary prevention of OVEs.

The analysis considered four relevant subgroups: patients who present with non-haemorrhagic stroke, patients who present with TIA, patients who present with MI and patients with symptomatic PAD. Stroke, TIA, MI and PAD are all atherothrombotic manifestations, and it is clinically and biologically plausible that all patients with atherothrombosis receive similar benefits from treatment. However, the evidence shows that the prognosis and baseline risk of recurrent OVEs differs according to the presenting event. Furthermore, licensing of antiplatelet drugs is limited according to the presenting event of patients in the clinical trials, and so each drug is not available in every subgroup. For these reasons it was necessary to conduct separate analyses by subgroup. The agents under comparison were aspirin, clopidogrel, MR-dipyridamole or the combination of aspirin and MR-dipyridamole (ASA-MR-dipyridamole). The analysis considered two options for treatment duration: lifetime treatment with each agent, or 2-year treatment with each agent followed by treatment with aspirin for the remainder of the patients' lifetime. The 2-year treatment period was chosen to reflect the length of the relevant clinical trials, and not as a result of selecting medically optimal treatment duration. Full detail of the clinical and cost-effectiveness analysis can be seen in a forthcoming HTA report.

The review of the economic evidence from the literature and manufacturers' submissions highlighted a number of potential limitations in existing studies assessing the cost-effectiveness of clopidogrel and MR-dipyridamole in the secondary prevention of OVEs. In order to overcome these limitations, it was necessary to undertake further modelling. An 'extended' probabilistic decision-analytic model was created, which used those models submitted by the manufacturers but employed a range of further analyses. The model was developed to estimate costs from the perspective of the UK NHS, and health outcomes in terms of quality-adjusted life-years (QALYs). The extended model was adapted from the model submitted to NICE by Sanofi-Synthelabo Ltd and Bristol-Myers Squib.

4.1.3 NICE guidance

In February 2004, NICE published the Appraisal Consultation Document[2] (ACD)[2] on the use of clopidogrel and MR-dipyridamole in the secondary prevention of occlusive vascular events. The provisional guidance recommended the following for aspirin tolerant patients:

1. For patients with stroke or TIA, ASA-MR-dipyridamole for 2 years, after which patients should revert to standard care, which includes aspirin monotherapy.
2. For patients with MI or PAD, the recommended treatment is aspirin.

4.1.4 Research recommendations

The recommendations for further research contained in the ACD[2] were:

- For patients at high risk of recurrent occlusive vascular events, with diabetes or who have had coronary surgery, further research is recommended into the effectiveness of clopidogrel.
- For patients with stroke who are intolerant to aspirin, further research is recommended into comparing the effectiveness of MR-dipyridamole with clopidogrel.

4.2 Methods

4.2.1 Model structure

The structure of the decision model for the stroke and TIA subgroups is shown in Figure 4.1. A cohort of patients (60 years of age) suffering from first-ever stroke or first-ever TIA are entered into the model which estimates the number of repeat strokes, fatal vascular events, and fatal non-vascular events over 40-years (lifetime). The cycle length is one-year, and in the first year following a repeat event, the risk of further strokes or vascular death is higher than if the patient survives 1-year event-free. The structure of the decision model for the MI and PAD subgroups is shown in Figure 4.2. A cohort of patients (60 years of age) suffering from first-ever MI or from symptomatic PAD are entered into the model which estimates the number of repeat MIs, repeat strokes, fatal vascular events, and fatal non-vascular events over 40-years (lifetime).

Figures 4.1 and 4.2 illustrate the fact that the model separates mortality into vascular and non-vascular causes. One reason for doing this is to separate out related and unrelated events, and to make full use of baseline and treatment effect data relating to those events. The risk of vascular or non-vascular death (NVD), and more importantly, the proportion of deaths due to vascular or non-vascular events, changes with age and time from non-fatal vascular events. To take the treatment relative risks on all-cause death observed during the trials, and to apply these to an age-related baseline risk of all-cause death, extrapolated over the lifetime of the patient cohort, would be assuming that the proportion of vascular to NVDs remained constant over that period, whereas the evidence shows that it does not. The separation of deaths in this way is also problematic. Both of the included randomised controlled trials (RCTs) had a follow-up period of 2 years. Over this period, in a cohort of patients with an average age of around 60, the number of observed death events will not be large, and hence the number of observed vascular or NVDs will be even smaller. This results

[2] The guidance issued in the ACD is provisional and subject to change following the consultation procedure.

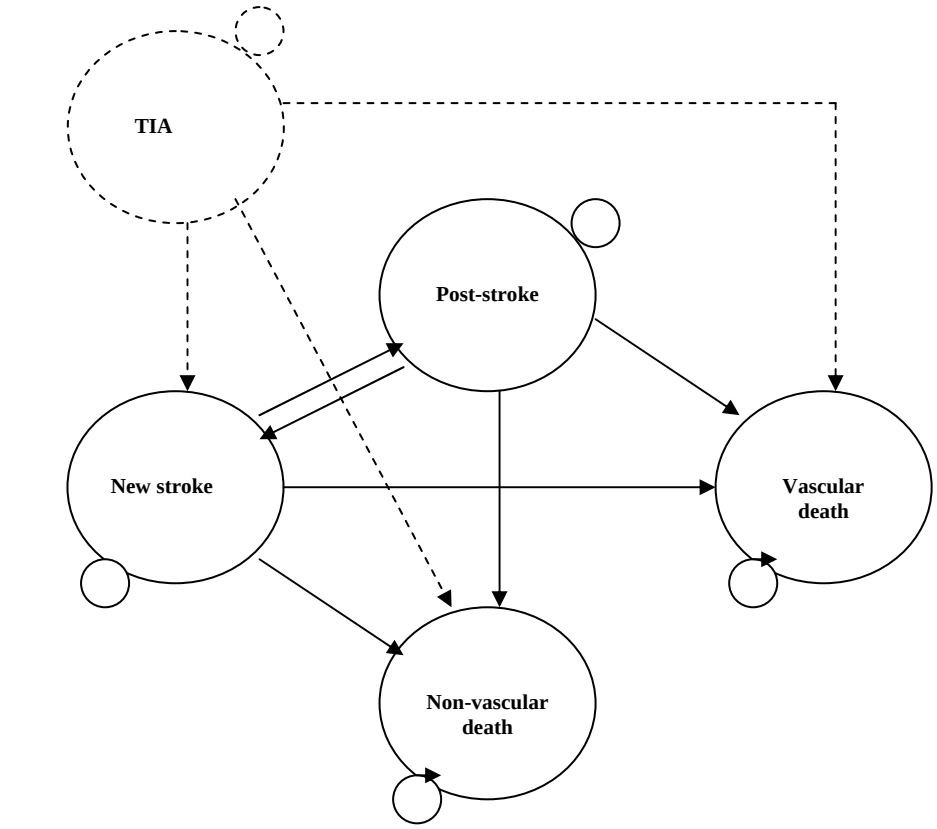


Figure 4.1: Structure of the extended model by University of York to assess cost-effectiveness of clopidogrel and MR-dipyridamole in the secondary prevention of occlusive vascular events in patients who have experienced stroke or TIA.

in greater uncertainty around the treatment effects on mortality. Secondly, the separation introduces the issue of whether or not to include treatment effects on NVD, which would not exist if all-cause mortality had been used instead. The inclusion of treatment effects on vascular mortality was expected as standard, however there was some question over the validity of including treatment effects on NVD. Separate analyses were therefore undertaken to explore the impact of including and excluding the relative risk reductions reported in the trials for NVD.

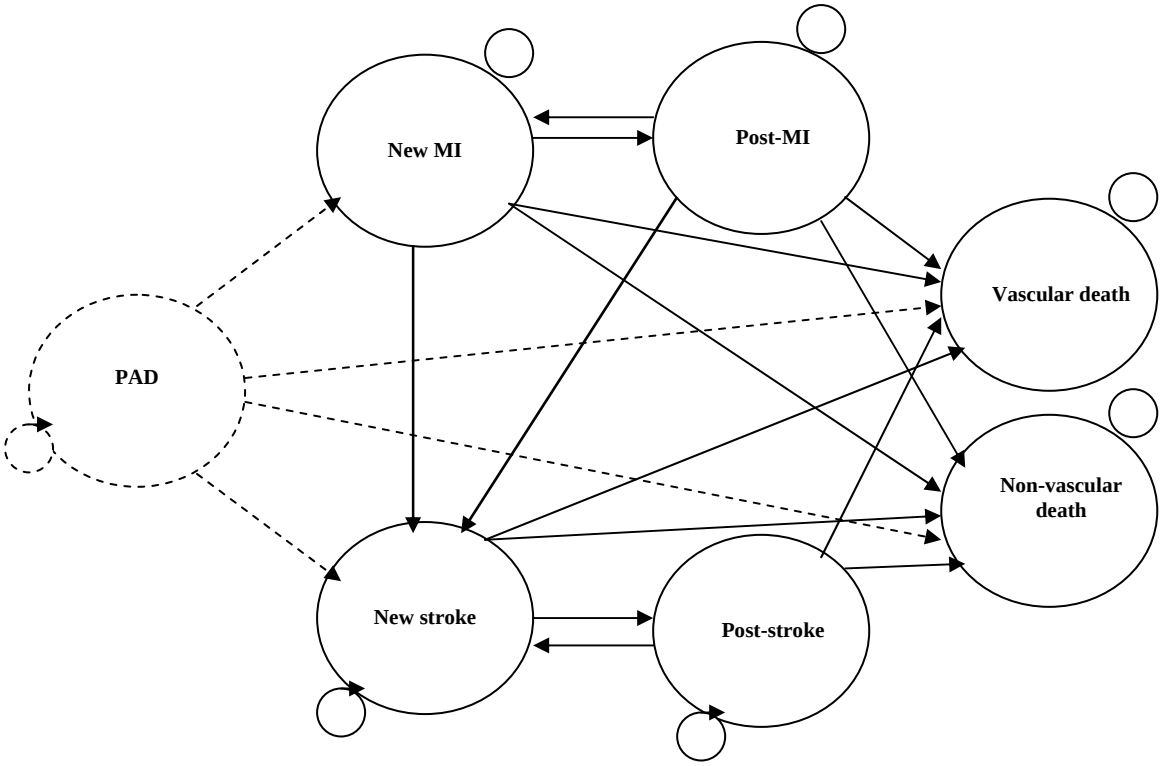


Figure 4.2: Structure of the extended model by University of York to assess cost-effectiveness of clopidogrel and MR-dipyridamole in the secondary prevention of occlusive vascular events in patients who have experienced MI or have been diagnosed with PAD.

A full list of all the parameters and their sources is available elsewhere.[1] Tables 4.1 and 4.2 provide details of the relative risks and cost data employed in the model. The baseline was assumed to represent the risk of further events while receiving treatment with aspirin, which was current practice for patients with atherothrombosis at the time to which the evidence relates. UK specific costs and utility estimated were incorporated into the model, and discounted at a rate of 6% and 1.5% per annum respectively. The perspective taken was that of the UK NHS.

Table 4.1: Relative risk parameters used in the model

	Clopidogrel		ASA-MR-dipyridamole		MR-dipyridamole	
Event	**RR**	**SE(ln(RR))**	**RR**	**SE(ln(RR))**	**RR**	**SE(ln(RR))**
Non-fatal MI[3, 4]	0.808	*0.0932*	1.058	*0.3364*	1.935	*0.2965*
Non-fatal stroke[3, 4]	0.948	*0.0682*	0.736	*0.1071*	0.981	*0.0981*
Vascular death[4, 5]	0.925	*0.0728*	0.991	*0.1257*	1.056	*0.1235*
Non-vascular death[4, 5]	1.087	*0.1057*	1.062	*0.1706*	0.981	*0.1740*
Fatal bleed[6]	0.999	*0.4470*	1.749	*0.6258*	0.499	*0.8653*
Non-fatal bleed[6]	0.938	*0.1444*	1.437	*0.3237*	0.374	*0.4774*

RR = relative risk compared to aspirin; SE = standard error; ln = natural logarithm

Table 4.2: Unit cost estimates used in the model

Item of resource use	Mean	Source
Stroke:		
Year 1 following non-disabling event	£5,907.34	Youman et al[7]
Subsequent years following 1-year event free	£1,052.17	Youman et al[7]
Year 1 following disabling event	£13,213.60	Youman et al[7]
Subsequent years following 1-year event free	£3976.61	Youman et al[7]
MI:		
Year 1 following event	£1,587.00	Palmer et al[8]
Subsequent years following 1-year event free	£3,966.00	Palmer et al[8]
PAD yearly cost	£1,000	Submission by SSBMS
Bleeding event	£2,377.24	Lightowlers and McGuire[9]
1 year supply:		
Aspirin	£3.47	British National Formulary[10]
Clopidogrel	£460.29	British National Formulary [10]
ASA-MR-dipyridamole	£117.00	British National Formulary [10]
MR-dipyridamole	£117.00	British National Formulary [10]

Cost of presenting stroke calculated by assuming 30.9% disabling events; cost of subsequent strokes calculated by assuming 35.6% disabling events (source ESPS-2). Cost of TIA equal to non-disabling stroke.

Due to separate indications and licences for clopidogrel and MR-dipyridamole, different combinations of treatment strategies are compared in each patient subgroup. The full range of licensed agents is considered in all analyses. Sensitivity analysis was used to address the assumption about treatment duration. The first analysis considered lifetime treatment duration with each of the agents, and a second analysis considered only 2-year treatment duration for clopidogrel, MR-dipyridamole and ASA-MR-dipyridamole, with all patients subsequently receiving aspirin for the remainder of their lifetime. These 2-year treatment strategies are detailed below:

- Strategy 1: Treatment with aspirin for the remainder of the patient's lifetime.
- Strategy 2: Treatment with clopidogrel for 2 years followed by treatment with aspirin for the remainder of the patient's lifetime.
- Strategy 3: Treatment with ASA-MR-dipyridamole for 2 years followed by treatment with aspirin for the remainder of the patient's lifetime.
- Strategy 4: Treatment with MR-dipyridamole for 2 years followed by treatment with aspirin for the remainder of the patient's lifetime.

The combinations of strategies considered for each patient subgroup are as follows.

- Stroke: Strategy 1 (aspirin), Strategy 2 (clopidogrel), Strategy 3 (ASA-MR-dipyridamole) and Strategy 4 (MR-dipyridamole).
- MI: Strategy 1 (aspirin) and Strategy 2 (clopidogrel).
- PAD: Strategy 1 (aspirin) and Strategy 2 (clopidogrel).
- TIA: Strategy 1 (aspirin), Strategy 3 (ASA-MR-dipyridamole) and Strategy 4 (MR-dipyridamole).

4.2.2 Probabilistic analysis

The model was fully probabilistic and the uncertainty around model parameters was characterised using appropriate distributions. Baseline risks were estimated with a multinomial regression, and the coefficients from the regression were characterised with normal distributions. Uncertainty around relative risks was characterised by lognormal distributions, and were correlated via an indirect comparison with aspirin. Full details of this are available elsewhere.[1]

4.2.3 Value of information analysis

Value of information (VOI) analysis was not undertaken in the initial report. Using the simulations that provided the cost-effectiveness estimates for the original assessment report, we estimated the expected value of perfect information (EVPI) for individual patients. Population EVPIs were then calculated using the incidence of stroke, TIA, MI and PAD as reported in the ACD.[2] Due to the similarities in the model for stroke and TIA patients, and in the model for MI and PAD patients, consideration of parameter EVPI is confined to only the stroke and MI subgroups.

4.3 Results

4.3.1 Adoption decisions

Following the provisional guidance issued by NICE regarding the use of antiplatelets for the secondary prevention of OVEs[2], we focus on the analysis concerning a 2-year treatment duration followed by aspirin for the remainder of patients' lifetimes (scenarios III and IV). The issue of treatment duration will be addressed further in the discussion.

Tables 4.3 to 4.6 display the mean costs and QALYs associated with each treatment strategy in each of the four subgroups, with incremental cost-effectiveness ratios estimated in the standard manner.[11] They also provide an estimate of the uncertainty surrounding the cost-effectiveness of each treatment strategy for different thresholds for cost-effectiveness, and finally they provide an estimate of the per patient decision EVPI.

Table 4.3: Estimated mean costs, effects and cost-effectiveness of each treatment strategy and decision EVPI for individual patients: stroke subgroup.

Strategy	Treatment duration	Cost	QALY	ICER	Threshold for cost-effectiveness:		
					£10,000	£30,000	£50,000
Scenario III: 2 year treatment duration, excluding treatment effects on non-vascular death							
Aspirin	2 years	£30,680	9.77	-	0.26	0.14	0.10
Clopidogrel	2 years	£31,648	9.81	D	0	0.12	0.18
ASA-MR-dipyridamole	2 years	£30,940	9.82	£5,500	0.62	0.62	0.60
MR-dipyridamole	2 years	£30,758	9.73	D	0.12	0.13	0.12
		Per patient EVPI			£188	£691	£1,268
Scenario IV: 2 year treatment duration, including treatment effects on non-vascular death							
Aspirin	2 years	£30,544	9.75	-	0.35	0.22	0.18
Clopidogrel	2 years	£31,481	9.78	D	0	0.10	0.16
ASA-MR-dipyridamole	2 years	£30,751	9.78	£7,968	0.52	0.53	0.52
MR-dipyridamole	2 years	£30,621	9.71	D	0.14	0.15	0.15
		Per patient EVPI			£286	£1,008	£1,793

Table 4.4: Estimated mean costs, effects and cost-effectiveness of each treatment strategy and decision EVPI for individual patients: TIA subgroup (with baseline event rates set to 80% stroke).

Strategy	Treatment duration	Cost	QALY	ICER	Threshold for cost-effectiveness:		
					£10,000	£30,000	£50,000
Scenario III: 2 year treatment duration, excluding treatment effects on non-vascular death							
Aspirin	2 years	£21,908	11.65	-	0.29	0.17	0.13
Clopidogrel	2 years	£22,811	11.69	£46,949	0	0.19	0.28
ASA-MR-dipyridamole	2 years	£21,956	11.68	£2,241	0.58	0.51	0.47
MR-dipyridamole	2 years	£22,052	11.62	D	0.13	0.14	0.12
		Per patient EVPI			£244	£932	£1,670
Scenario IV: 2 year treatment duration, including treatment effects on non-vascular death							
Aspirin	2 years	£22,085	11.70	-	0.35	0.25	0.21
Clopidogrel	2 years	£22,963	11.72	£52,339	0	0.12	0.43
ASA-MR-dipyridamole	2 years	£22,111	11.71	£4,266	0.49	0.45	0.18
MR-dipyridamole	2 years	£22,220	11.66	D	0.16	0.19	0.19
		Per patient EVPI			£348	£1,265	£2,255

Table 4.5: Estimated mean costs, effects and cost-effectiveness of each treatment strategy and decision EVPI for individual patients: MI subgroup.

Strategy	Treatment duration	Cost	QALY	ICER	Threshold for cost-effectiveness:		
					£10,000	£30,000	£50,000
Scenario III: 2 year treatment duration, excluding treatment effects on non-vascular death							
Aspirin	2 years	£18,284	8.90	-	0.83	0.29	0.22
Clopidogrel	2 years	£19,202	8.95	£17,081	0.17	0.71	0.78
		Per patient EVPI			£31	£274	£322
Scenario IV: 2 year treatment duration, including treatment effects on non-vascular death							
Aspirin	2 years	£18,182	8.87	-	0.88	0.39	0.31
Clopidogrel	2 years	£19,078	8.91	£21,448	0.12	0.61	0.70
		Per patient EVPI			£23	£387	£459

Table 4.6: Estimated mean costs, effects and cost-effectiveness of each treatment strategy and decision EVPI for individual patients: PAD subgroup.

Strategy	Treatment duration	Cost	QALY	ICER	Threshold for cost-effectiveness:		
					£10,000	£30,000	£50,000
Scenario III: 2 year treatment duration, excluding treatment effects on non-vascular death							
Aspirin	2 years	£15,180	11.04	-	0.96	0.30	0.17
Clopidogrel	2 years	£16,041	11.08	£20,733	0.04	0.70	0.83
		Per patient EVPI			£5	£146	£116
Scenario IV: 2 year treatment duration, including treatment effects on non-vascular death							
Aspirin	2 years	£15,279	11.03	-	0.96	0.52	0.36
Clopidogrel	2 years	£16,123	11.05	£31,300	0.04	0.48	0.64
		Per patient EVPI			£4	£346	£393

Treatment with clopidogrel is consistently the highest cost strategy, and treatment with aspirin is consistently the lowest cost strategy. This is due almost entirely to the huge difference in acquisition costs for each the drugs. The annual drug cost is £460.29 per patient for clopidogrel, compared to £117 for ASA-MR-dipyridamole and MR-dipyridamole, and only £3.47 for aspirin. Where MR-dipyridamole is a relevant treatment strategy, it is consistently associated with the lowest number of QALYs, and is dominated in every analysis. At a cost-effectiveness threshold of £30,000 per QALY gained, it would appear that treatment with ASA-MR-dipyridamole is the most cost-effective strategy for patients in the stroke or TIA subgroup, and that treatment with clopidogrel is the most cost-effective treatment strategy for patients in the MI subgroup. The most cost-effective treatment strategy for patients with PAD is not clear, as the inclusion of treatment effects on NVD causes the ICER for clopidogrel to exceed £30,000.

The cost-effectiveness of each treatment strategy is uncertain, and the cost-effectiveness acceptability curves (CEACs)[12] in Figures 4.3 and 4.4 illustrate this uncertainty in the stroke subgroup, and the fact that none of the treatment strategies are ruled out entirely by the analysis. The cost-effectiveness frontiers show the probability that the optimum strategy is cost effective.

The inclusion of treatment effects on NVD increases the ICER associated with each strategy, and also increases the uncertainty surrounding the most optimal strategy. This is to be expected as the trial data show that clopidogrel and ASA-MR-dipyridamole exhibit small, and statistically insignificant relative risk increases for NVD when compared to aspirin.

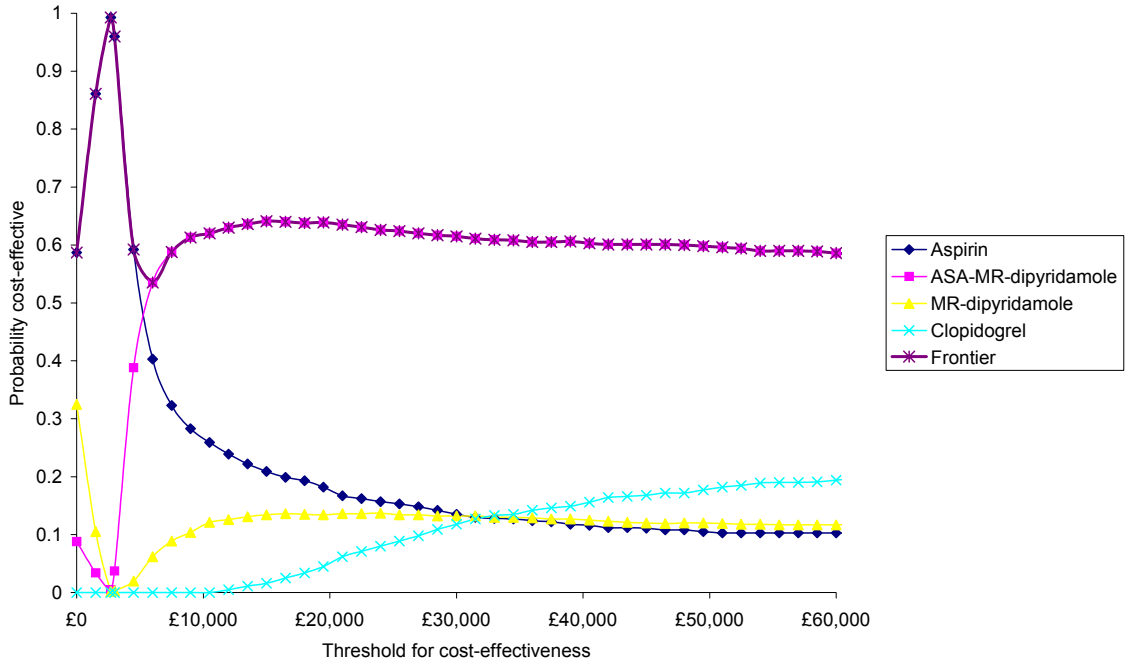


Figure 4.3: CEACs and frontier for stroke subgroup, model excluding treatment effects on NVD

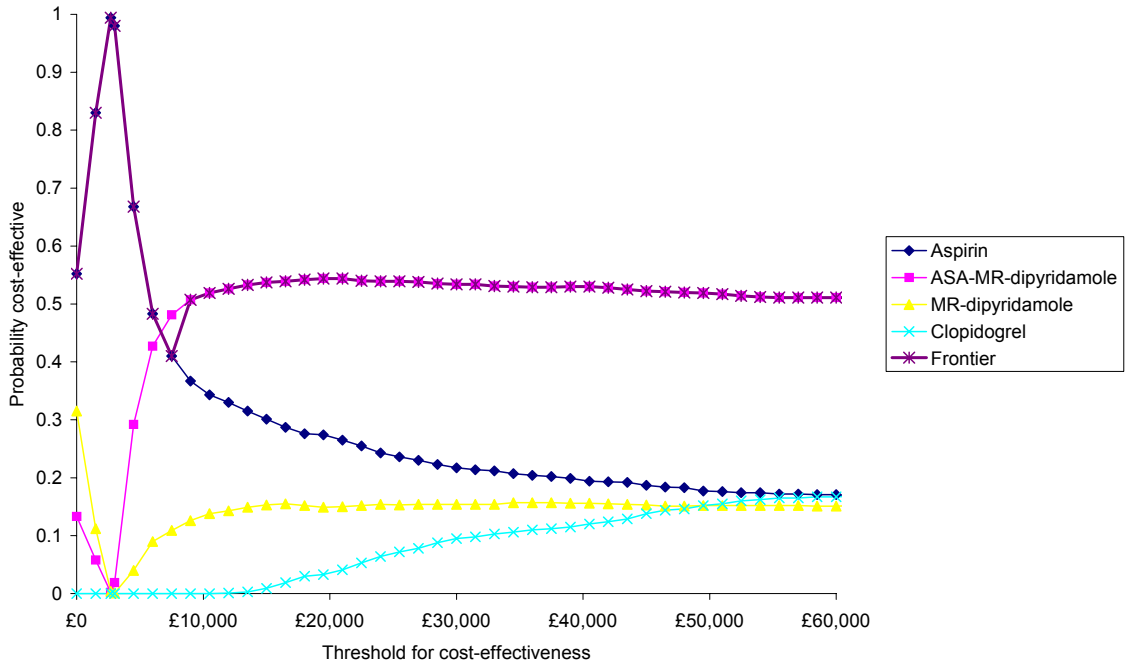


Figure 4.4: CEACs and frontier for stroke subgroup, model including treatment effects on NVD

4.3.2 Research recommendations

Population EVPI is large in each subgroup, ranging from £116mn to £865mn for an effective lifetime of the technology of 10 years and a cost-effectiveness threshold of £30,000. Figures 4.5 and 4.6 show that it increases with the threshold for cost-effectiveness for most of the range in each patient subgroup.

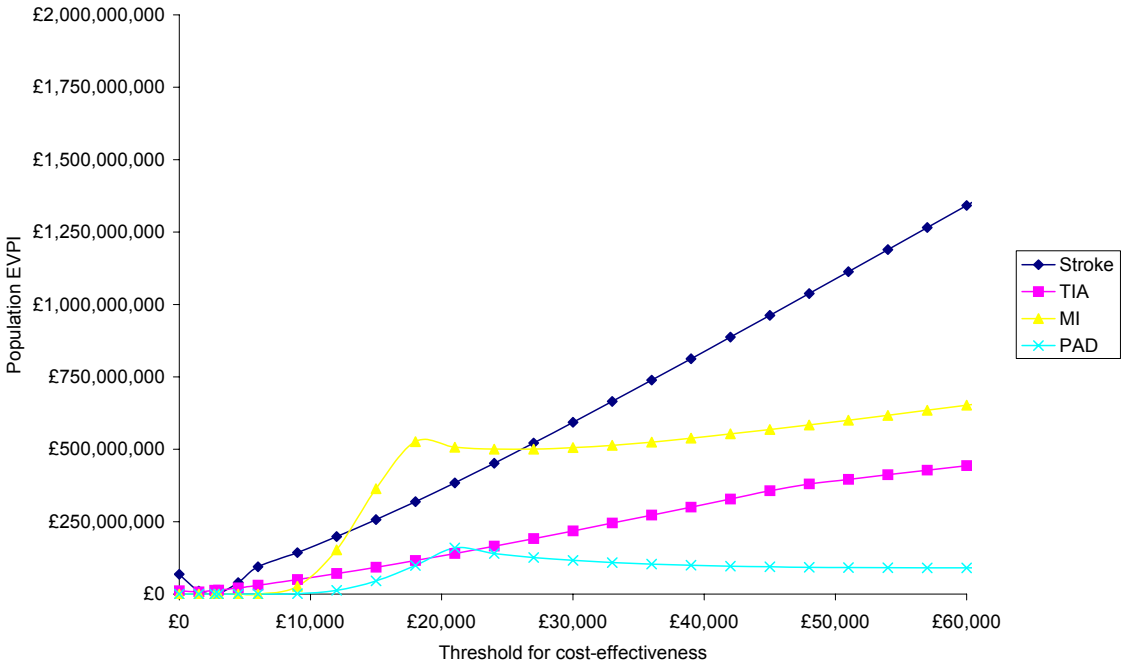


Figure 4.5: Population EVPI, model excluding treatment effects on NVD

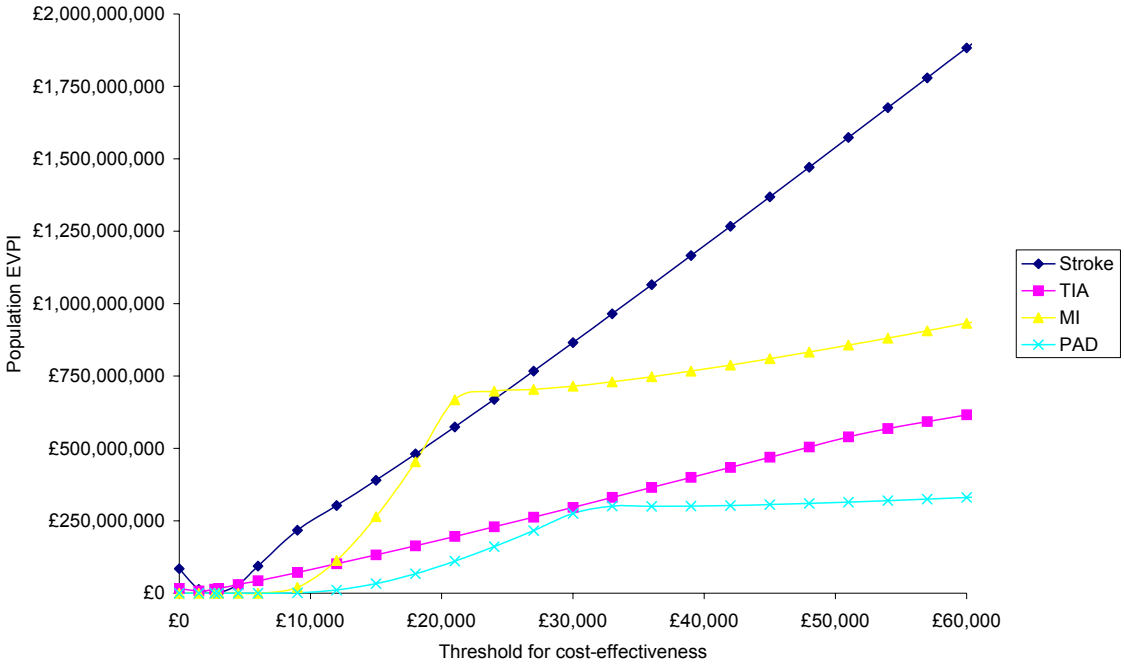


Figure 4.6: Population EVPI, model including treatment effects on NVD

The population EVPI is very high in each subgroup, and this is partly because such a large population of patients stand to benefit from any future research. It is clear that the population EVPI is much higher in the stroke subgroup for most of the range, especially when compared to MI, which is more common. This is in part due to the increase in treatment options for patients with stroke: four, compared to two for patients with MI. Secondly, the treatment comparisons in the MI subgroup are taken directly from a RCT. In contrast, the treatment comparisons in the stroke subgroup were only possible via an indirect comparison of clopidogrel and dipyridamole, as there was no trial evidence comparing both agents directly.

This indirect comparison inflates the variance around the relative risk parameters, thereby increasing the uncertainty around the optimal strategy.
The EVPI is larger when the model includes treatment effects on the risk of NVD. There are two levels of uncertainty around treatment effects on NVD. Firstly, there is uncertainty as to their existence, and secondly, if they do exist there is great uncertainty around their magnitude and direction given that observed difference in numbers of events is extremely small. EVPI reflects the latter of these. Given a difference in numbers of 1 or 4 events, the possibility that such a difference is observed solely by chance or random variation is very high. The problem is compounded when, as in the case of death, the consequence of even a small difference in numbers of events is large.

4.3.3 EVPI for individual parameters

To estimate EVPI for individual parameters, we first grouped the parameters into logical sets: baseline risk estimates, treatment relative risk estimates, estimates of utility and estimates of costs. Of these groups, for every patient subgroup, only the estimates of the relative risks associated with each treatment had positive values of information, so we then explored the relative risk parameters individually. The results of this analysis are shown in Figures 4.7 to 4.10. They reveal that the parameters accounting for almost all of the VOI were the treatment effects with respect to vascular mortality, and when they are included the treatment effects with respect to NVD.

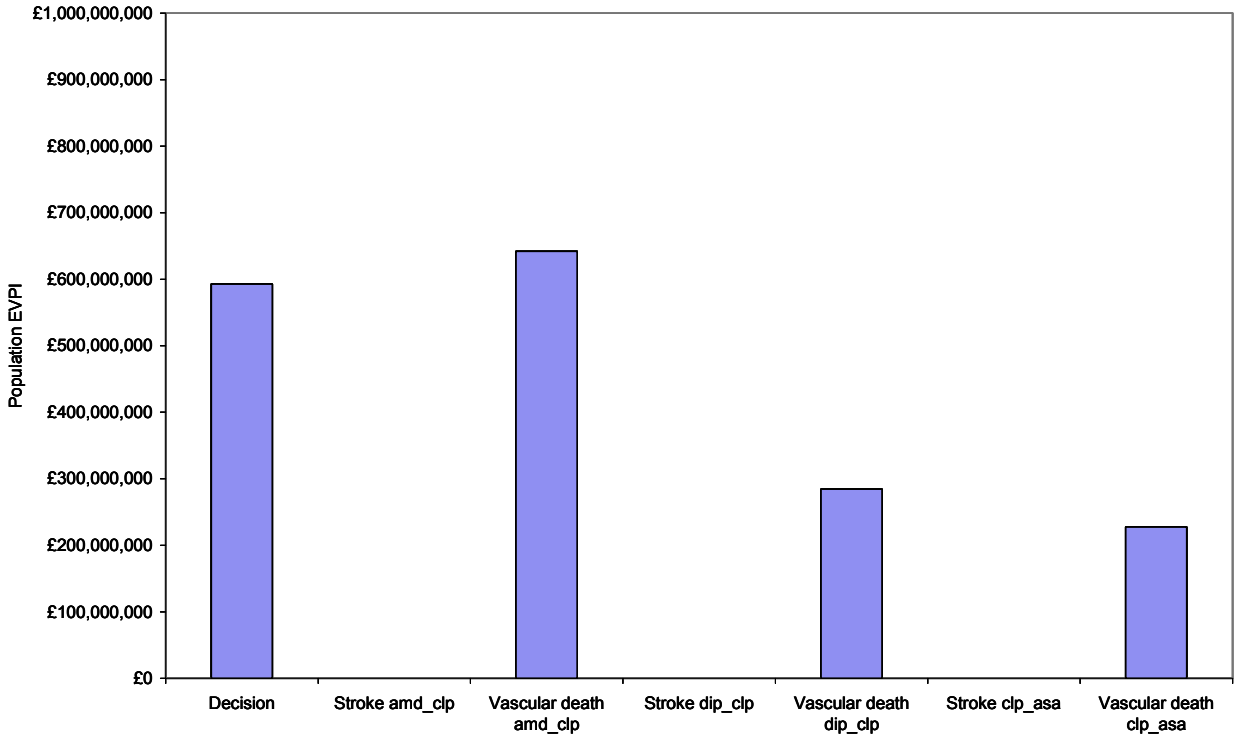


Figure 4.7: Partial EVPI on oindividual relative risks: stroke subgroup, excluding treatment effects on NVD

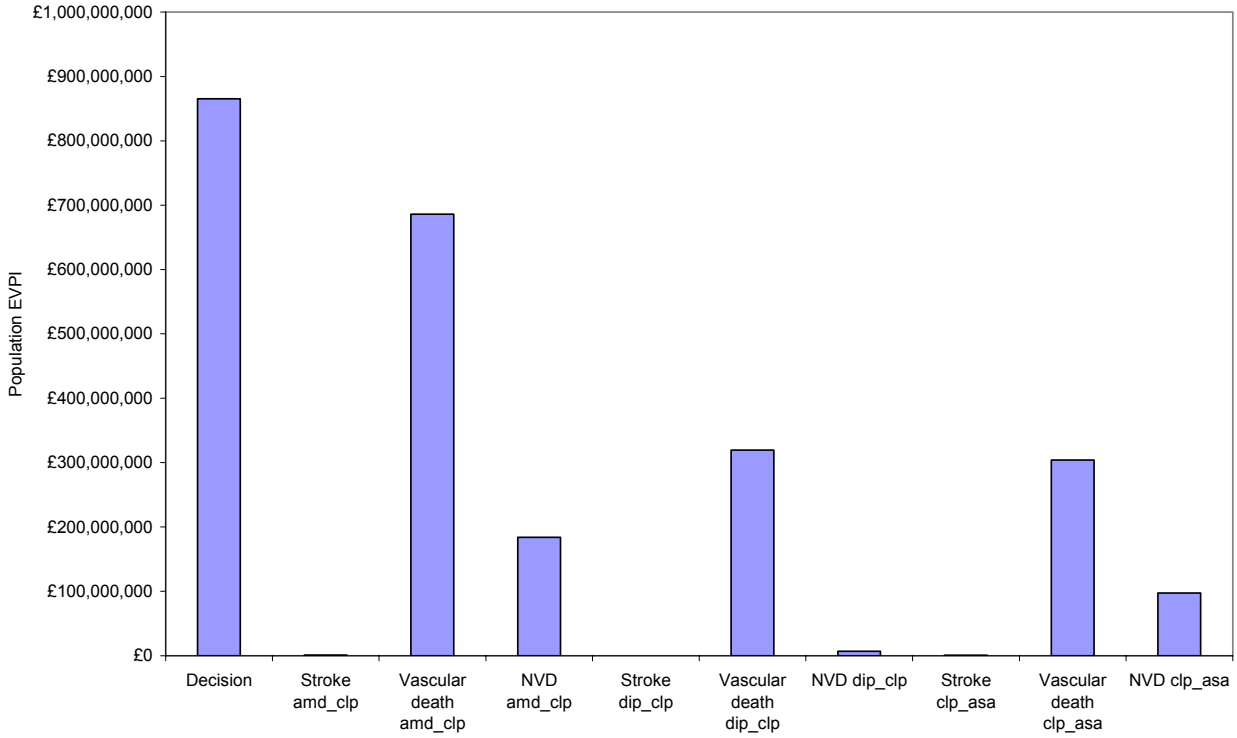


Figure 4.8: Partial EVPI on individual relative risks: stroke subgroup, including treatment effects on NVD

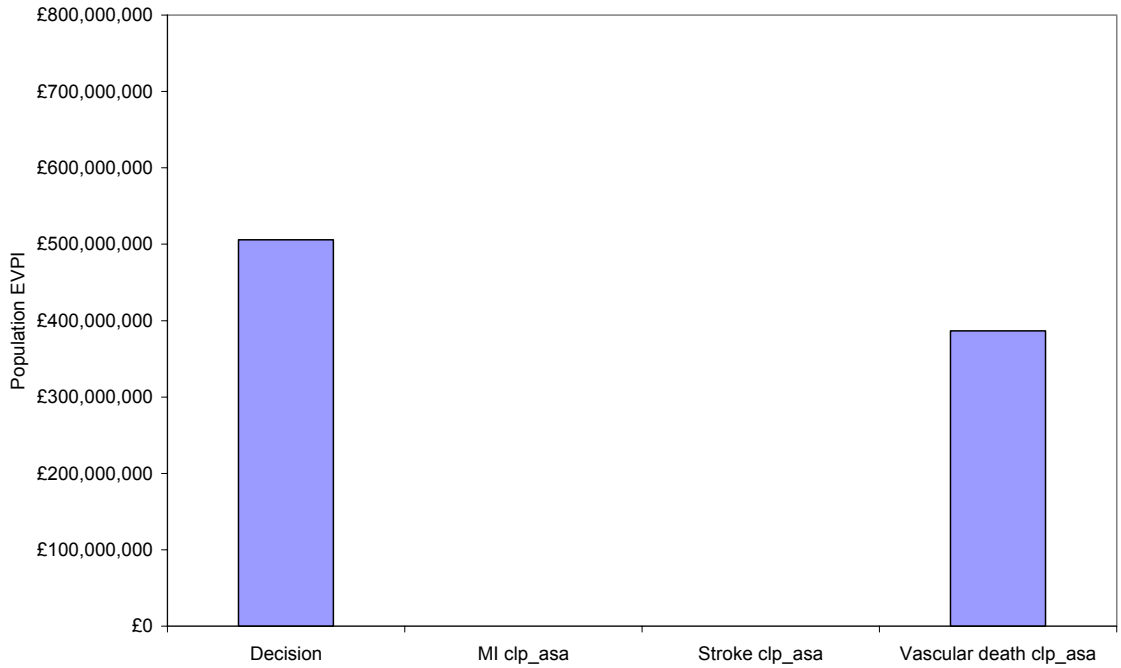


Figure 4.9: Partial EVPI on individual relative risks: MI subgroup, excluding treatment effects on NVD

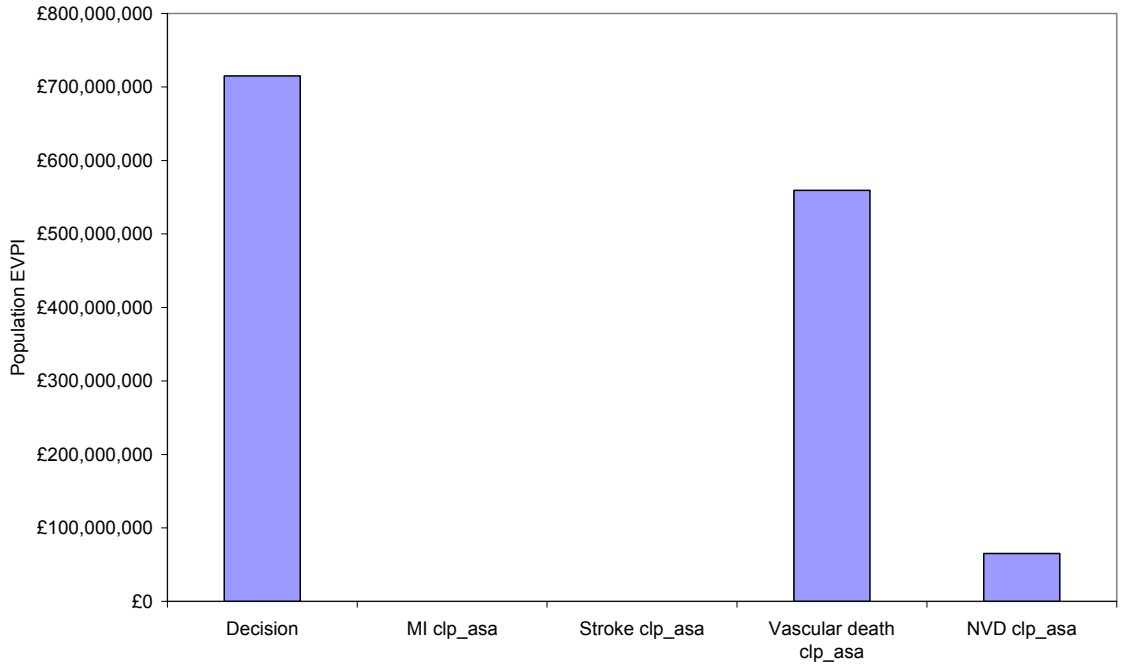


Figure 4.10: Partial EVPI on individual relative risks: MI subgroup, including treatment effects on NVD

Key to Figures 4.7-4.10: clp_asa = relative risk of clopidogrel compared to aspirin; amd_clp = relative risk of ASA-MR-dipyridamole compared to clopidogrel (indirect); dip_clp = relative risk of MR-dipyridamole compared to clopidogrel (indirect)

4.3.4 Discussion

The initial guidance issued by NICE was based on both scenarios (i.e. both the model that included treatment effects on non-vascular mortality, and the one that excluded them), demonstrating the uncertainty over the relevance of treatment effects on non-vascular mortality. This uncertainty is not captured in the model, or in the estimates of EVPI. There are two related reasons for uncertainty over the inclusion of such treatment effects. One such reason is the issue of related and unrelated events, and whether we might realistically expect these agents to have a treatment effect on non-vascular mortality. A related uncertainty over the existence of these relative risk treatment effects is the lack of evidence, shown by the wide confidence intervals that include 1, and encompass both large relative risk reductions, as well as large relative risk increases. One way to include this uncertainty in the model would be to use Bayesian methods. The prior belief that antiplatelet agents are not expected to have a differential effect non-vascular mortality could be reflected with a prior distribution centred on a relative risk of 1. The evidence from the trials could then be used to update this prior, and the small amount of evidence would probably mean that the prior would not be overwhelmed. Another function of the prior could be to reduce the likelihood of observing large relative risk increases (or decreases), which would have the effect of contracting the confidence interval.

An alternative way of handling this uncertainty is to follow the method used in the original analysis, and present different scenarios including and excluding the points of controversy. This allows decision makers to separate areas of modelling uncertainty from areas of evidence uncertainty. In this example, both models for the stroke subgroup supported the use of ASA-MR-dipyridamole as the optimal treatment strategy. Conversely, the two alternate models for the PAD subgroup gave conflicting results for a cost-effectiveness threshold of £30,000 per QALY, illustrating that if it were correct to include treatment effects on non-

vascular mortality, a decision based on a model excluding them could be wrong. On the other hand, if antiplatelets do not have a differential effect on non-vascular deaths, the inclusion of trial data with respect to treatment effects would be incorrect. Their inclusion would unnecessarily inflate the uncertainty around the optimal strategy, and potentially change the optimal strategy. The parameter EVPI around the relative risk of non-vascular death suggests that further research would be valuable to reduce the uncertainty around this parameter. However, the use of only the trial data, and the failure to make use of other data sources available at the time, namely prior beliefs based on a range of evidence and experience, means that the model that includes treatment effects of non-vascular death probably overestimates the decision and parameter uncertainty. Presenting the two models separately places a burden on the decision maker to incorporate their prior beliefs on the validity of the two structures, in order to make their decision based on a weighted average of the two results. Thus presenting alternative scenarios reduces the transparency of the decision-making process, as opposed to the use of Bayesian methods to incorporate prior beliefs explicitly so that they are open to scrutiny.

The relative risks of OVEs in figures 4.7 to 4.10 are based on the comparisons of clopidogrel to aspirin, ASA-MR-dipyridamole to clopidogrel, and MR-dipyridamole to clopidogrel. The relative risks were incorporated in this manner because it had been necessary to make an indirect comparison between clopidogrel and dipyridamole. To calculate a relative risk from an adjusted indirect comparison, one can use the following formulae:

$$RR_{BC} = RR_{BA} - RR_{CA}; \quad SE(RR_{BC})=\sqrt{(SE(RR_{BA})^2+SE(RR_{CA})^2)}$$

Where RR_{BC} is the (log) relative risk of treatment B compared to treatment C, and *SE* is the standard error.[13] Thus it was possible to calculate the relative risk of dipyridamole compared to clopidogrel, where treatment B is represented by ASA-MR-dipyridamole or MR-dipyridamole, treatment C by clopidogrel and treatment A by aspirin. By first calculating this indirect comparison, and then incorporating it into the model so that the relative risk of dipyridamole compared to aspirin was calculated via the relative risk of clopidogrel compared to aspirin, the treatment effects of each agent are correlated. However, the use of an indirect comparison increases the uncertainty around the relevant treatment effects, compared to a direct comparison of the same treatments. When incorporated in the manner of this model, the inflation of variance occurs around the treatment effects of ASA-MR-dipyridamole and MR-dipyridamole. By incorporating the extra uncertainty caused by the indirect comparison in an uneven way, the model may overestimate the decision uncertainty around the optimal treatment strategy of ASA-MR-dipyridamole for the stroke subgroup. More sophisticated modelling techniques involving the use of the variance-covariance matrix can be used to 'spread' the increase in uncertainty due to the indirect comparison evenly across each treatment strategy. Nevertheless, the uncertainty around the relevant treatment strategies could be reduced by a direct comparison of all agents.

Finally, we return to the issue of treatment duration. The use of a 2-year treatment strategy merely reflected the length of the included RCTs, and in practice patients prescribed clopidogrel or dipyridamole often continue on those treatments for longer than 2 years. The lack of trial evidence on the long-term effects of these drugs does not prevent the modelling of longer treatment duration, and an analysis of lifetime treatment was presented by assuming that the treatment effect remained constant while the patient was on therapy. This was presented as a separate analysis, allowing decision-makers to separate out areas of modelling uncertainty, but in fact lifetime treatment duration represents a relevant alternative strategy to 2-year treatment duration, and they need to be compared simultaneously in the model. This is because the decision uncertainty about 2-year treatment strategies compared to lifetime treatment strategies cannot be informed by the separate models evaluating subsets of the full range of strategies. The inclusion of more treatment strategies would likely increase the decision EVPI; also uncertainty in the treatment effects would be extrapolated further through the model for longer treatment durations, which would likely increase the value of acquiring more information on those effects. The current model does not characterise the uncertainty around the optimal treatment duration, and therefore cannot inform about the value of acquiring information on the longer-term effects of clopidogrel and dipyridamole. Again, the decision over what treatment duration to model can be augmented by evidence

and experience outside of RCTs, which are typically not designed to ascertain the optimal duration.

4.3.5 Summary

At a cost-effectiveness threshold of £30,000 per QALY gained, it seems likely that EVPI will exceed the cost of further research. For an estimated 5-year lifetime of the technology, the EVPI in each subgroup is in the range £66mn to £495mn. Individual parameter EVPI reveals that a large proportion of this value is accounted for by the individual treatment effects on vascular mortality, and would seem that further investigation into the effects of clopidogrel, MR-dipyridamole and ASA-MR-dipyridamole on vascular death would be cost-effective, in particular a direct comparison of each agent in this area. The VOI around the other individual model parameters is small or zero, but this does not mean that they may not have significant VOI in combination. The results of further research into the effects of each agent on vascular mortality will almost certainly provide more evidence of the treatment effects on NVD, but would be unlikely to provide conclusive evidence as to whether they are related events that need to be included in the model. The decision over which events are related to the treatments under consideration is a matter of judgement, which will be in part informed by trial evidence. Simply to include any event recorded in the trial without considering whether it is related to the treatments in question will inflate the uncertainty in the model. Where there is uncertainty over whether particular events should be fully excluded prior to analysis, by omitting them from the decision model, the event and the uncertainty over its relevance or existence can be explicitly included in the model with the use of Bayesian methods. Finally, the research recommendations based on this model are confined to the 2-year analyses, and ignore uncertainty around the longer-term effects of clopidogrel and dipyridamole. If the model had compared alternative treatment durations directly, the VOI around the treatment effects may have increased further, and the model could have informed about the most cost-effective treatment duration, rather than basing that decision on the effect-driven and budget-limited RCTs.

Appendix: Clopidogrel and modified-release dipyridamole in the secondary prevention of occlusive vascular events.

Element of health technology assessment	Reference case	Criteria met by assessment?	Comment
Defining the decision problem	The scope developed by the Institute	Yes	
Comparator	Alternative therapies routinely used in the NHS	Yes	
Perspective on costs	NHS and PSS	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Synthesis of evidence on outcomes	Based on a systematic review	Yes	
Measure of health benefits	Quality adjusted life years (QALYs)	Yes	
Description of health states for calculation of QALYs	Health states described using a standardised and validated generic instrument	Yes	Health state preference values were taken from published sources, all of which used the EQ-5D to elicit public-based preferences from the relevant patient group in the UK.
Method of preference elicitation for health state valuation	Choice-based method, for example, time trade-off, standard gamble (not rating scale)	N/a	
Source of preference data	Representative sample of the public	N/a	
Discount rate	An annual rate of 3.5% on both costs and health effects	No	The discount rates of 6% per annum for costs and 1.5% per annum for health outcomes recommended at the time of the NICE assessment were used.
Equity position	An additional QALY has the same weight, regardless of the other characteristics of the individuals receiving the health benefit	Yes	

References

1. Jones, L., et al., Assessment report: A rapid and systematic review of the clinical effectiveness and cost effectiveness of clopidogrel and modified-release dipyridamole in the secondary prevention of occlusive vascular events. 2004.
2. NICE, Appraisal Consultation Document: Clopidogrel and modified-release dipyridamole in the secondary prevention of occlusive vascular events. 2004.
3. Easton, J.D., Benefit of clopidogrel in patients with evidence of cerebrovascular disease. *Neurology*, 1998. 50: p. P03081.
4. Diener, H.C., et al., European stroke prevention study 2: Efficacy and safety data. *Journal of the Neurological Sciences*, 1997. 151: p. S1-77.
5. Committee, C.S., A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). *The Lancet*, 1996. 348: p. 1329-1339.
6. Collaboration, A.T., Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *British Medical Journal*, 2002. 324: p. 71-86.
7. Youman, P., et al., The Economic Burden of Stroke in the United Kingdom. *Pharmacoeconomics*, 2003. 21: p. 43-50.
8. Palmer, S., et al., A cost effectiveness model comparing alternative management strategies for the use of glycoprotein IIb/IIIa antagonists in non ST elevation acute coronary syndrome. 2002, Report commission by NICE.
9. Lightowlers, S. and A. McGuire, Cost-Effectiveness of Anticoagulation in Nonrheumatic Atrial Fibrillation in the Primary Prevention of Ischemic Stroke. *Stroke*, 1998. 29: p. 1827-1832.
10. British National Formulary (www.BNF.org, accessed October 2001). 2001, British Medical Association and Royal Pharmaceutical Society of Great Britain: London.
11. Van Hout, B.A., et al., Costs, effects and c/e-ratios alongside a clinical trial. *Health Economics*, 1994. 3: p. 309-319.
12. Fenwick, E., K. Claxton, and M. Sculpher, Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics*, 2001. 10: p. 779-789.
13. Song, F., et al., Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *British Medical Journal*, 2003. 326: p. 472-477.

5. The cost-effectiveness and value of information associated with treatment for influenza in otherwise healthy adults

5.1 Background

5.1.1 .Condition and technology

Of the three main influenza virus strains (A, B and C), only types A and B are known to cause significant morbidity in humans. Influenza A occurs more frequently and is more virulent than influenza B; there have been significant outbreaks of influenza A in England and Wales in nine of the 11 years from 1990 to 2000. There are several subtypes of influenza A, of which H3N2 and H1N1 have co-circulated in humans since 1978. Outbreaks of influenza B have occurred in England and Wales four times in the 11 years from 1990 to 2000. There have been outbreaks of both influenza A and B in two of these years. Since 1990, about 74% of influenza has been caused by the influenza A virus, but this has varied from 20% to 97% from season to season.

Influenza epidemics of varying intensity occur most winters. The condition is usually self-limiting in people who are relatively healthy, with typical symptoms such as headache, fever, sore throat, cough and aching muscles and joints lasting several days. However, more severe, predominantly respiratory complications such as pneumonia and bronchitis, are the source of substantial morbidity and increased mortality associated with influenza epidemics. In England and Wales, an estimated 6,200 to 29,600 people died during each of the epidemics between 1975-76 and 1989-90[3]; about ten times the actual number of death certifications for influenza, suggesting that influenza is responsible for many hidden deaths.

5.1.2 Technology Assessment Review

The principal component of public health strategies aimed at controlling the burden of influenza is vaccination. In the UK NHS vaccination is offered to 'high-risk' groups with uptake levels running at 69% in 2002/3[i]. Two adamantanes (amantadine (Lysovir®) and rimantadine (Flumadine®)) have been produced since the mid-1960's for both treatment and prevention although the latter does not have a UK license and clinical uptake of the former has been limited due to concerns over adverse events, resistance and its limited spectrum of activity (adamantanes operate only against the replication of influenza A).

Neuraminidase inhibitors are a relatively new class of antiviral drugs that provide additional potential strategies for the control of influenza. Zanamivir (Relenza®) for the treatment of influenza is administered by means of a diskhaler® and was the subject of one of the first technology appraisals undertaken by the National Institute for Clinical Excellence (NICE)[ii], in the UK. Guidance issued in 2000 recommended NHS use should be limited to 'high-risk' groups [iii]. The launch of oseltamivir (Tamiflu®), taken orally and licensed for both prophylaxis and treatment, prompted further review of that guidance.

5.1.3 NICE guidance

In February 2003, NICE issued guidance on the use of zanamivir, oseltamivir and amantadine for the treatment of influenza[iv]. The guidance pertains only to periods when it is known that influenza A or B is circulating, via surveillance schemes such as those operated by the Royal College of General Practitioners and the Public health Laboratory Service. It does not cover periods of pandemic influenza, impending pandemic or widespread epidemic of a new influenza strain.

In relation to otherwise healthy adults, neither zanamivir, oseltamivir or amantadine were recommended for the treatment of influenza.

5.1.4 Research recommendations

The recommendations for further research, issued as part of the NICE guidance in relation to otherwise healthy adults were that:

- More needs to be known about the quality-of-life measurement of people with influenza.
- A systematic evaluation of the feasibility of near-patient testing for influenza and the type of influenza is strongly recommended.
- Modelling to include epidemiological and health-economic aspects of the treatment and prophylaxis of influenza would aid future considerations of the cost effectiveness of the anti-influenza drugs.
- Monitoring the development of viral resistance to oseltamivir and zanamivir would be valuable.
- The systematic collection of more information on hospitalisation rates and the frequency of complications for children with influenza would be highly desirable.
- Any randomised trial of oseltamivir or zanamivir should include an amantadine arm as well as a placebo arm.

5.1.5 Additions to the appraisal report

The work presented here discusses only the cost-effectiveness of otherwise healthy adults. Neither those considered to be at elevated risk of influenza complications either due to age (over 65 years) or concommittant disease, or children are considered in this chapter.

QALY estimates associated with influenza morbidity were based on three trials of oseltamivir in the original appraisal report [v] (WV15670, WV15671 and WV15730) comprising over approximately 600 patients in total. Data from one additional trial have been made available since that time. M76001 comprises in excess of 300 placebo and 600 oseltamivir patients and have been incorporated here. The mean difference in Quality Adjusted Life Days (QALDs) between placebo and oseltamivir patients in the original three trials was 0.56. That difference decreases to 0.37 with the inclusion of M76001. This is a key driver in the cost-effectiveness estimates and therefore the results presented here differ significantly from those presented in the original TAR.

5.2 Methods

5.2.1 Description of the model

A decision tree model was developed to assess cost-effectiveness of competing influenza treatments in terms of additional cost to the UK NHS per additional Quality Adjusted Life Year (QALY) gained and the principal features of this model are described in figure 5.1.

The decision model is described in four separate stages. The decision problem for the UK National Health Service (NHS) is described at stage one, namely which of four alternative influenza treatment strategies should be adopted; amantadine (100mg daily); zanamivir (10mg twice daily); oseltamivir (75mg daily); or no drug treatment. Each treatment course lasts five days. The decision is relevant to a patient population that consists of those with influenza like illness who decide whether to consult with a GP (stage 2). This distinction is included in the model due to the likelihood that the proportion of individuals that consult will be dependent on the decision taken at stage 1, that is, more individuals will consult the GP if an effective drug is available. However, variations in the value of this parameter between the competing treatment strategies are only included in sensitivity analyses due to data limitations. It should be noted that the importance of this issue was identified by one-way sensitivity analysis in the original appraisal report[v] and was cited as one of the reasons for rejection of oseltamivir in this patient group[iv].

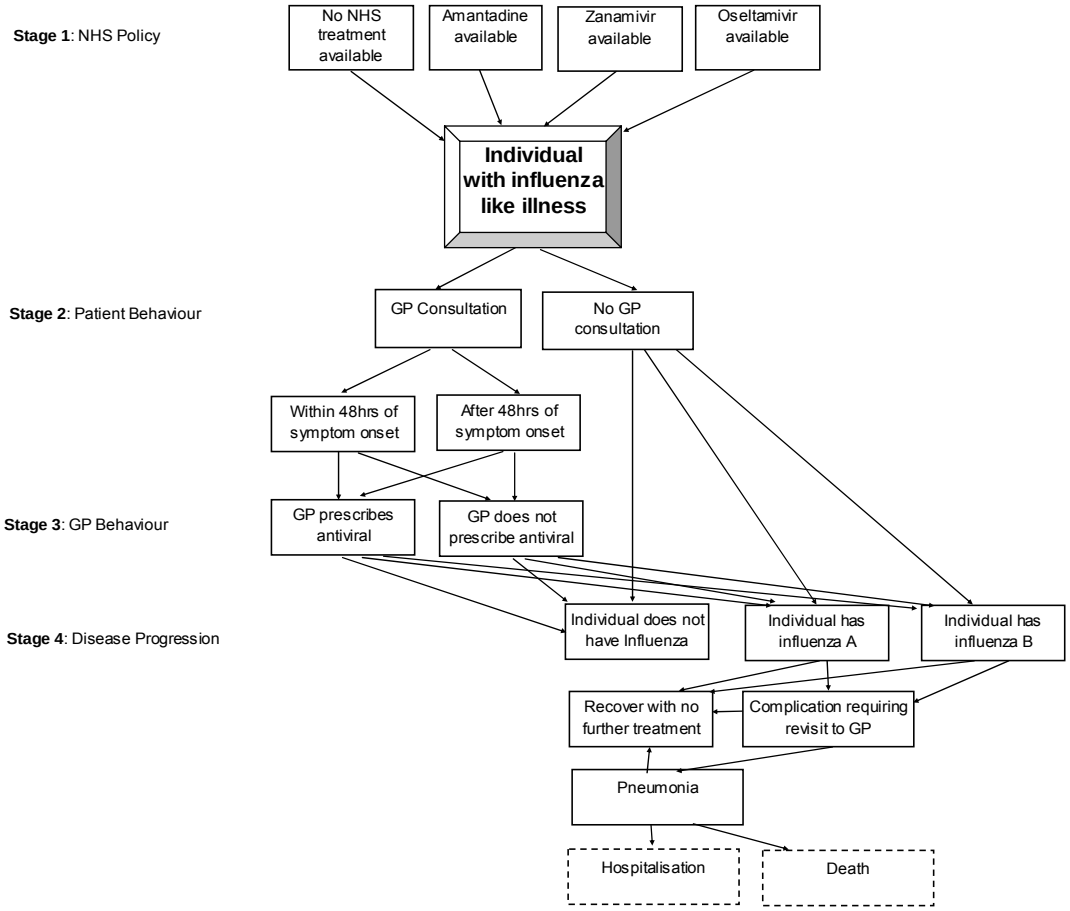


Figure 5.1: Model Schematic

It is assumed that treatment with any antiviral must be made within 48hours of symptom onset but the model allows for individuals to receive (ineffective) treatment after that time period since symptom onset may be insidious and difficult to accurately recall (stage 3). Prescription of antibiotics is often made at this initial GP consultation. Disease progression is described at stage 4. The first distinction made here is between those individuals that have genuine influenza rather than clinically indistinguishable conditions such as bacterial infections or other viruses (for example, respiratory syncytial virus [RSV]). Influenza may be either strain A or B and the distinction is important since amantadine is effective only for the treatment of influenza A. True influenza cases may experience complications which require a revisit to the GP and which may become more serious. The base case analysis includes only one serious influenza complication (pneumonia) since existing data were of limited value for a UK context in relation to both hospitalisations and mortality. This is unsurprising given that such complications are relatively rare and clinical trials are not powered to detect such differences. Several previous studies have extrapolated from intermediate datavi,vii. Results presented here do not include any such extrapolation.

5.2.2 The evidence

Parameter values are recorded in Tables 5.1 and 5.2. Data are primarily drawn from meta-analysis of randomised controlled trial (RCT) data but where no such sources were available, or were considered either inappropriate or insufficient, they were supplemented or replaced with alternative data, often from multiple sources [v].

The duration of symptoms for influenza positive persons receiving either no treatment or ineffective treatment (antiviral after 48 hours of symptom onset, amantadine for influenza B) was drawn from pooled analysis of patients receiving placebo in trials of both oseltamivir and zanamivir [v]. For both oseltamivir and zanamivir the mean reduction in symptom days was based on meta-analysis of trial data [v]. Meta analysis of trial data for amantadine was based

on the more limited outcome measure of duration of fever. Therefore, a meta-regression was constructed from oseltamivir RCTs which allowed the relationship between symptom and fever duration to be estimated [v] which was then applied to the observed mean reduction in fever from RCTs of amantadine at a dose of 100mg [viii].

Quality Adjusted Life Days (QALDs) were estimated to reflect the expected duration and severity of influenza illness for each of the four treatment strategies[4]. Four RCTs of oseltamivir versus placebo included patients' own reports of quality of life (WV15670[ix], WV15671[x], WV 15730[xi], M76001[xii]) over a period of 21 days. The incorporation of uncertainty inherent in both quality of life data and length of illness is an important component of the model and an alternative specification is explored as an additional structural sensitivity analysis.

Other health related quality of life considerations included in the model were pneumonia [xiii], adverse events associated with amantadine (by assumption) and, in sensitivity analysis only, avoided deaths [xiv].

Cost data, also shown in table 1 and expressed in 2001 prices, were drawn predominantly from UK published sources [xv,xvi,xvii xviii]. All drug costs were inflated to include pharmacy prescribing fees and container allowances [xix].

Data from the Royal College of General Practitioners' (RCGP) network of sentinel practices was used to estimate the probability that ILI is influenza[xx] and the strain of influenza[xxi]. The probability that an individual will consult the GP is derived from estimates of the number of excess GP consultations in the UK [xxii,xxiii], the size of the UK population [xxiv] and the attack rate [v]. The probability of attending within 48 hours of symptom onset [xxv] was adjusted by the probability of rapid onset of symptoms based on random effects meta-analysis of published studies [v] to estimate the proportion of individuals that truly consult the GP within the time frame required in order to benefit from any antiviral treatment. Evidence exists that some patients receive antibiotics at this initial GP consultation [xxiii] but it was assumed that this would substantially reduce where antivirals were prescribed. RCGP data were used to estimate the proportion of patients that develop influenza complications requiring a repeat GP consultation [xxvi, xxvii] and this figure was adjusted by the relative risk of antibiotic use in several RCTs of zanamivir and oseltamivir [v]. This measure was considered the most realistic proxy measure for repeat GP consultations in NHS practice available from RCTs since repeat consultations themselves were often part of the study protocol. Similar meta-analysis of RCT data was used in relation to reduced incidence of pneumonia for the two neuraminidase inhibitors. No evidence of reductions in complications of any type was identified for amantadine. This chapter does not include models which extrapolate from intermediate endpoints, such as antibiotic use or pneumonia rates, to either hospitalisations or mortality.

All costs and benefits accrue over a single influenza season with the exception of health benefits associated with avoided deaths. A discount rate of 1.5% is applied to these benefits.

5.2.3 Probabilistic analysis

Probability distributions were assigned to model parameters as indicated in tables 1 and 2. Cost and outcome parameters were assigned normal or lognormal distributions with the exception of quality of life scores. Beta distributions were assigned to the QALY score for no treatment and oseltamivir (which in turn are use to derive QALYs for amantadine and zanamivir), although QALDs are reported for ease of interpretation. A number of different distributions were applied to parameters which were used to derive probabilities in the model. Normal distributions of logged odds ratios and relative risks were assigned. Other parameters were variously beta or lognormal as appropriate. Monte Carlo simulation was used to reflect parameter uncertainty in the model via cost effectiveness acceptability curves (CEACs) and associated frontiers[xxviii].

[4] QALDs are reported here rather than QALYs for ease of interpretation. A QALY is simply 1/365 of a QALD. Cost utility ratios are reported in terms of cost per QALY.

Table 5.1: Cost and outcome parameter values and probability distributions.

Parameter description	Mean Value	(95% Confidence Interval)	Probability distribution	Source
Mean length of illness - no treatment	7.690	6.856 *to* 8.596	Lognormal	[v]
Mean reduction in length of illness - amantadine	1.320	*(0.112 to 2.852)**	Derived from other variables	[v]
Mean reduction in length of fever - amantadine	1.007	0.097 *to* 1.925	derived from other variables	[viii]
Mean reduction in length of illness - zanamivir	1.683	0.812 *to* 2.575	Normal	[v]
Mean reduction in length of illness - oseltamivir	1.919	0.939 *to* 2.918	Normal	[v]
QALDs - no treatment	15.132	12.297 *to* 17.967	Beta**	WV15670[ix], WV15671[x], WV15730[xi], M76001[xii]
QALDs - amantadine Tx	15.390		derived from other variables	
QALDs - zanamivir Tx	15.456		derived from other variables	
QALDs - oseltamivir Tx	15.502	12.788 *to* 18.215	Beta**	WV15670[ix], WV15671[x], WV15730[xi], M76001[xii]
QALD loss amantadine adverse events	0.95		None	Assumption
QALY loss from death	18.987		None	[xiv]
QALD loss pneumonia	6.348		None	[xiii]
Costs				
GP visit	21.38		None	[xv]
Antibiotic	4.05		None	[xvi,xix]
Amantadine	3.38		None	[xvi,xix]
Zanamivir	24.98		None	[xvi,xix]
Oseltamivir	19.16		None	[xvi,xix]
Hospitalisation	3503.03	2012 *to* 5743	Lognormal	[xvii]

***2.5th to 97.5th percentiles, ** Note that the Beta was applied to the QALY score although QALDs are reported for ease of interpretation.**

Table 5.2: Probability parameter values and probability distributions

Parameter description	Mean Value	(95% Confidence Interval)	Probability distribution	Source
ILI is influenza	0.46	0.425 to 0.494	Normal on Log of odds	[xxi]
Influenza is influenza A	0.684	0.484 to 0.884	Beta	[xxii]
Consulting the GP	0.282	*(0.145 to 0.533)**		
Annual excess GP consultations due to influenza	623,520	388779 to 858263	Lognormal	[xxiii,xxiv]
Size of UK population	33,743,500			[xxv]
Attack rate	0.066	0.037 to 0.112	Normal on log of odds	[v]
Consulting GP within 48hours (conditional on consulting)	0.178	*(0.150 to 0.212)**		
Consult on Day 1	0.109	0.088 to 0.135	Normal on log of odds	[xxvi]
Consult on day 2	0.093	0.073 to 0.117	Normal on log of odds	[xxvi]
Rapid onset of symptoms	0.511	0.305 to 0.714	Normal on log of odds	[v]
Receive drug if after 48 hours	0.028	*(0.015 to 0.045)**		
dependent on same variables as previous parameter				
Receive drug if before 48 hours	0.952	0.829 to 1.075	Beta	Assumption
Antibiotic at first visit - no treatment strategy	0.42	0.417 to 0.423	Normal on Log of odds	[xxiv]
Antibiotic at first visit - antiviral strategies	0.048	0.000 to 0.171	Beta	Assumption
Complication requiring return GP visit - no treatment/amantadine	0.371	0.362 to 0.380	Normal on Log of odds	[xxvi,xxvii]
Relative risk of complication - zanamivir	0.741	0.575 to 0.954	Normal on log of relative risk	[v]
Relative risk of complication - oseltamivir	0.423	0.160 to 0.930	Normal on log of relative risk	[v]
Pneumonia - no treatment/amantadine	0.034	0.008 to 0.020	Normal on log of odds	[v]
Relative risk of pneumonia - zanamivir	0.35	0.112 to 1.094	Normal on log of relative risk	[v]
Relative risk of pneumonia - oseltamivir	0.15	0.060 to 0.720	Normal on log of relative risk	[v]
Antibiotic if complication develops	0.814	0.806 to 0.822	Normal on Log of odds	

*** 2.5th to 97.5th percentiles**

5.2.4 Sensitivity analyses

The sampled QALY values used in the base case model for all four strategies are reliant upon the observed values in four RCTs for oseltamivir versus placebo. Despite the fact that oseltamivir has demonstrated a significant reduction in the length of influenza illness (at customary levels of statistical significance) compared to placebo [xxix] the quality of life data are not as clear. In over 45% of sampled values from the beta probability distributions, the quality of life score for oseltamivir is lower than that for no treatment. Several reasons for this are possible. Firstly, the alleviation of symptoms was defined in trials of oseltamivir by the absence or mild experience of feverishness, headache, sore throat, cough, myalgia, fatigue, and congestion [v]. However, quality of life may depend on symptoms other than those measured, for example 'high temperature' is also included in some trials of zanamivir, and may be affected by mild symptoms or minor complications of influenza not otherwise accounted for. Secondly, oseltamivir is associated with adverse events such as gastrointestinal disorder and vomiting which, whilst mild, may contribute to the observed values and which are not accounted for separately in the model.

Notwithstanding these concerns, an alternative specification of the model (structural sensitivity analysis 1) used data on symptom days to estimate QALDs and thereby reduce the uncertainty inherent in the model, using the following relationship:

$$Q_{i1} = Q_{i0} + \left[\left(\overline{Q_1} - \overline{Q_0}\right)\left(\overline{D_1} / D_{i1}\right)\right]$$

Where Q_i refers to the QALD score for the *i*th sample, D_i the difference in length of illness between intervention and no treatment for the ith sample. Subscript $_0$ denotes the no treatment option whilst subscript $_1$ refers to the active intervention (amantadine, zanamivir or oseltamivir).

Sensitivity analyses 2 explores the impact of potential increases in the propensity for individuals to consult the GP. For otherwise healthy adults at low risk of influenza complications, relatively few individuals would make such consultations given the absence of an effective intervention. However, the availability of neuraminidase inhibitors increases the expected benefits of a GP consultation and therefore the proportion of individuals that make such consultations. This is a potentially important parameter value for the cost-effectiveness model since these additional consultations increase costs by a much larger proportion than benefits (all individuals will generate additional GP costs whilst only a small proportion will have any capacity to benefit). A beta distribution was used to increase sample values for the baseline probability of presenting to the GP. This was based on an assumed mean increase of 10% (2.5^{th} and 97.5^{th} percentiles, 0.28 to 34.76) and was applied to the neuraminidase inhibitor treatment strategies.

5.2.5 Value of information analysis

The Monte Carlo simulation output was also used to estimate the expected value of perfect information per person for the model as a whole, for individual parameters and groups of parameters. Parameters are grouped together as five potential study types. "GP surveillance" includes eight probability parameters; ILI is influenza; influenza is strain A; present to the GP within 48 hours of symptom onset; receive treatment if presenting before 48 hours; receive treatment if presenting after 48 hours; receive antibiotic at first GP consultation (which differs by strategy). "QALY study" refers to the two QoL parameters for no treatment and oseltamivir. The three individual trial options include the relevant parameters on complications requiring repeat GP consultations, rates of pneumonia and, for amantadine, adverse event profile.

Population EVPI figures are calculated by estimating the size of the population and the expected period over which any decision will remain relevant with values for future years. We estimate the number of influenza like illnesses in the otherwise healthy adult population as 4.1 million per annum[xxxxxxi] and the relevant time period as 5, 10 and 15 years. Figures are discounted to present values at 6% per annum.

5.3 Results

Table 5.3: Base case expected costs and benefits.

Strategy	Mean Cost	Mean QALD	Cost per QALY- compared to No treatment	Cost per QALY (ICER)
No Treatment	£10.96	6.917	-	
Amantadine	£11.06	6.920	£13,122	£13,122
Zanamivir	£12.24	6.925	£53,383	dominated
Oseltamivir	£11.85	6.927	£30,868	£37,795

5.3.1 The adoption decision

Table 5.3 presents the mean costs, benefits and cost-effectiveness ratios for an individual in the community with influenza like illness. Differences in benefits (QALDs) between strategies are relatively small since for each strategy only a small proportion of patients actually receive effective antiviral treatment (either they do not visit the GP at all, do so after 48 hours, do not have influenza, or do not receive antiviral treatment). The cost effectiveness ratios indicate that for each additional QALY generated by the amantadine strategy compared to no treatment, an additional cost of £13k is incurred. Both of the neuraminidase inhibitors are more effective and more costly than amantadine. Oseltamivir dominates zanamivir and generates additional QALYs at £31k compared to no treatment and £38k compared to the next most effective treatment that is not dominated (amantadine).

The base case CEACs shown in figure 5.2 indicate that where the maximum acceptable incremental cost effectiveness ration (MAICER) exceeds £10k, the degree of uncertainty associated with the optimal strategy is substantial and this is particularly the case for the range across which amantadine is optimal. In fact, where λ lies between £25k and £35k amantadine is the least likely of all four strategies to be cost-effective, yet is the optimal strategy. Just 13% of simulations are cost-effective at λ=£30k.

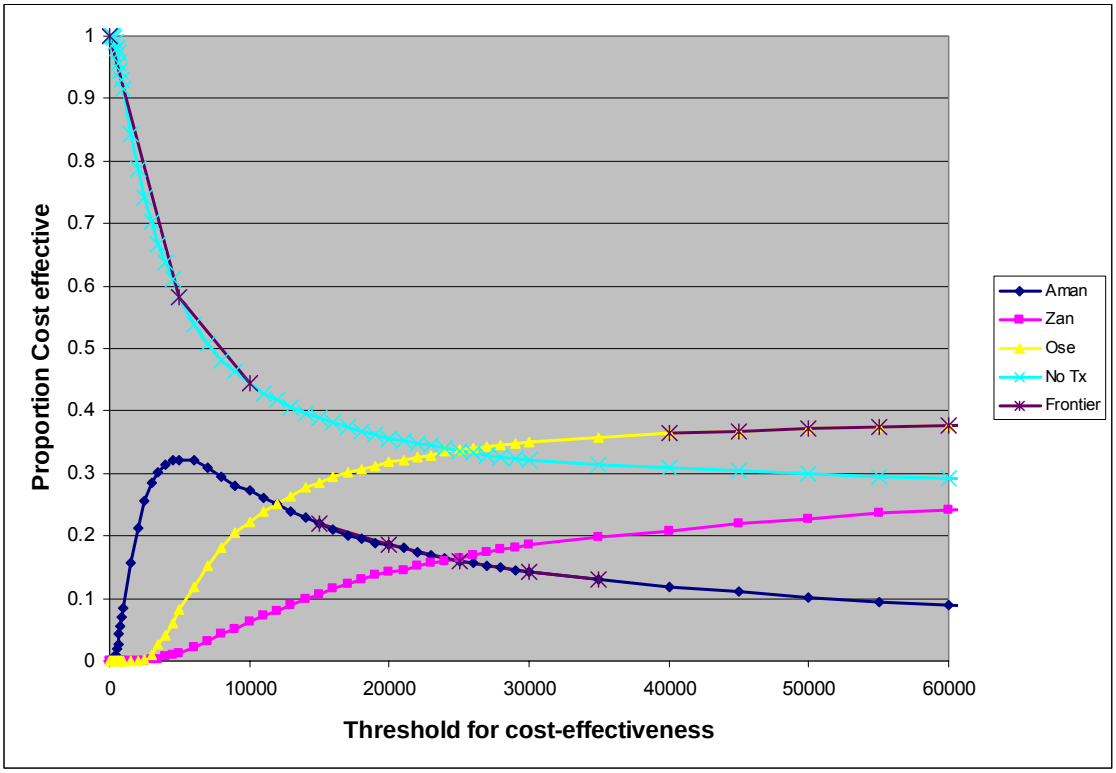


Figure 5.2: Cost effectiveness acceptability curve and frontier – base case analysis

The restructuring of the model in sensitivity analysis 1 does not impact on the central estimates of cost per QALY but reduces the extent to which sampled QALD values for oseltamivir fall below those of no treatment. The resultant reduction in uncertainty for all four strategies translates into the cost-effectiveness acceptability curve shown in figure 3. At a low λ level, no treatment remains the strategy most likely to be cost-effective and amantadine forms part of the cost-effectiveness acceptability frontier between λ values of approximately £14k and £38k at higher levels of probability than in the base case. Across this range, the probability that amantadine is cost-effective exceeds 0.5. At higher λ values, oseltamivir is the strategy most likely to be cost effective with a probability that peaks at approximately 0.8.

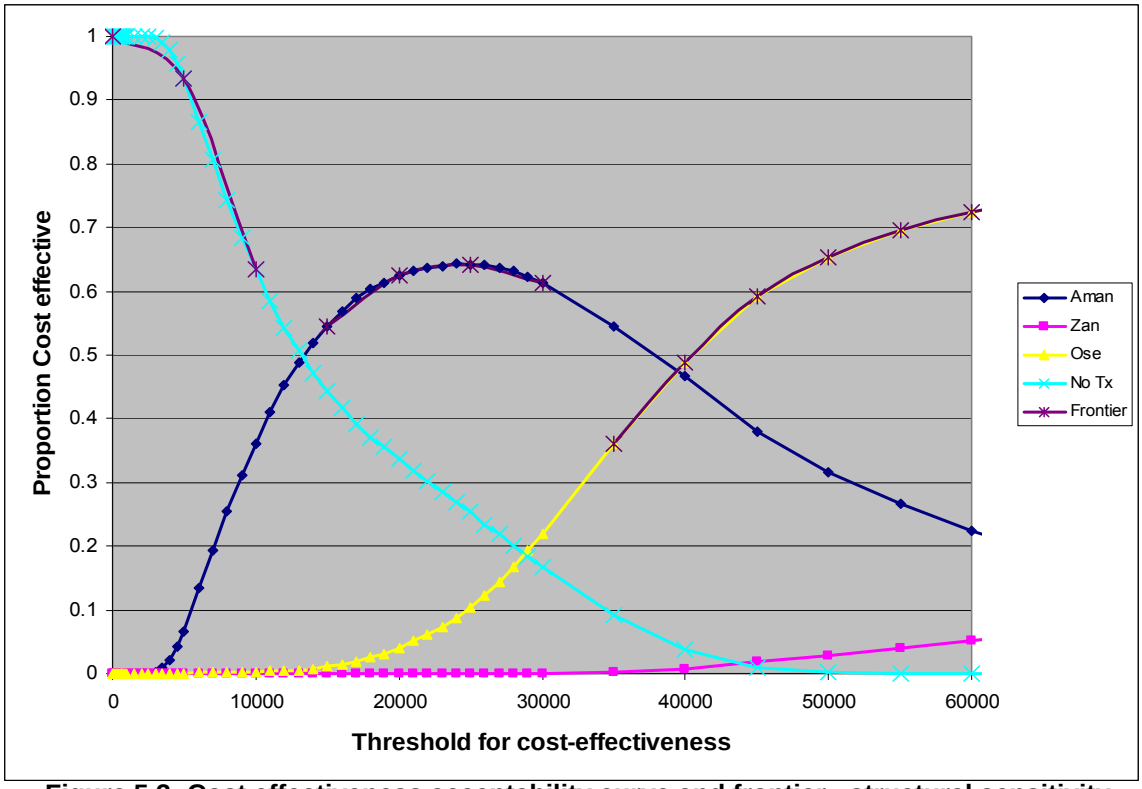


Figure 5.3: Cost effectiveness acceptability curve and frontier - structural sensitivity analysis 1.

The cost effectiveness results associated with sensitivity analysis two are shown in table 5.4 and demonstrate the sensitivity of the results to changes in the probability of consulting the GP. For oseltamivir, the ICER is in excess of £70,000 in this scenario.

Table 5.4: Sensitivity analysis 2 - Cost and benefits

Strategy	Mean Cost	Mean QALD	Cost per QALY- compared to no treatment	Cost per QALY (ICER)
No Treatment	£10.71	6.917	-	
Amantadine	£10.81	6.920	£12,259	£12,259
Zanamivir	£12.71	6.926	£79,946	dominated
Oseltamivir	£12.29	6.928	£54,303	£71,112

5.3.2 Research recommendations

EVPI for base case analysis

Figure 5.4 plots the global EVPI per person as a function of λ between £0 - £100,000 for the base case analysis for 5yr, 10yr and 15yr lifespans of the technologies. EVPI is equal to £38m, £66m and £88m respectively where λ is £30,000 and these values continue to rise as λ increases. Table 5.5 summarises all EVPI results.

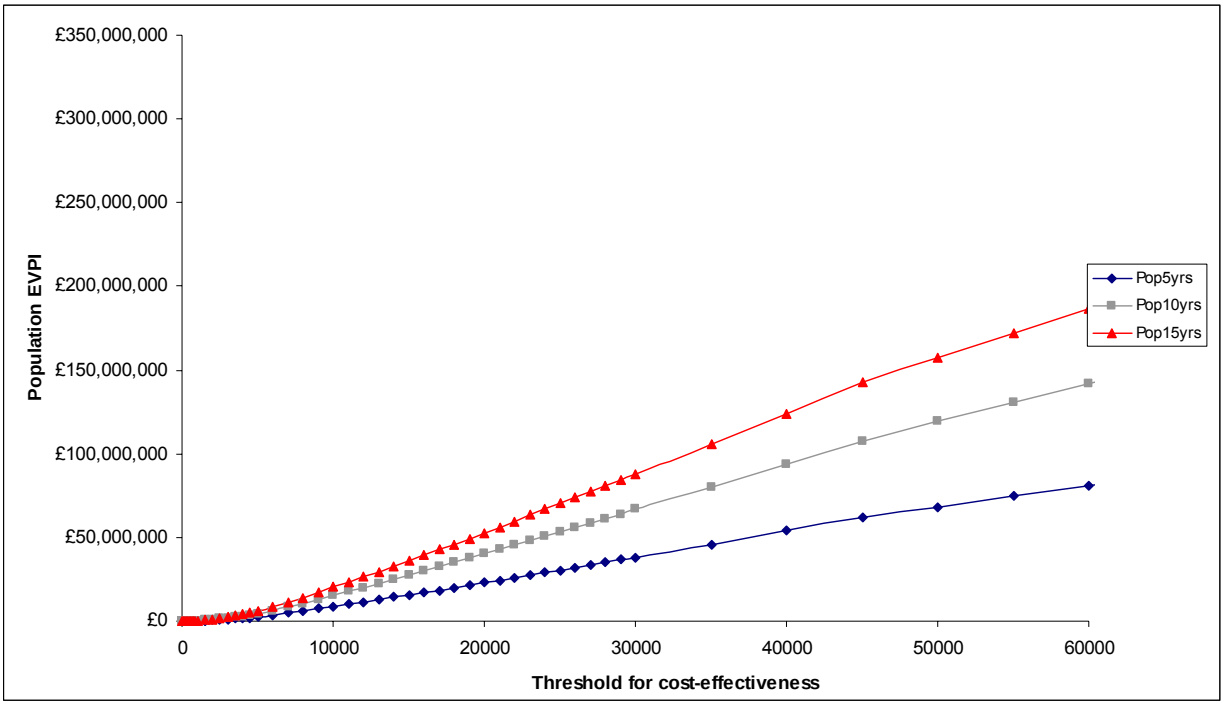


Figure 5.4: Global EVPI – Base case analysis.

EVPI for individual parameters are shown in figure 5.5. The analysis was run on all parameters but only those that generate substantial value at some threshold value (between £0 and £100k) are displayed. The EVPI values for two parameters in particular dwarf all others. The uncertainty surrounding the quality of life scores for both oseltamivir and no treatment, which in turn are used to calculate scores for amantadine and zanamivir, continue to rise as λ increases. At a willingness to pay of £30k per QALY, the EVPI per person for the quality of life scores for oseltamivir and no treatment are £0.94 and £0.92 respectively, which equates to a population value of approximately £30m assuming a 10 yr lifetime

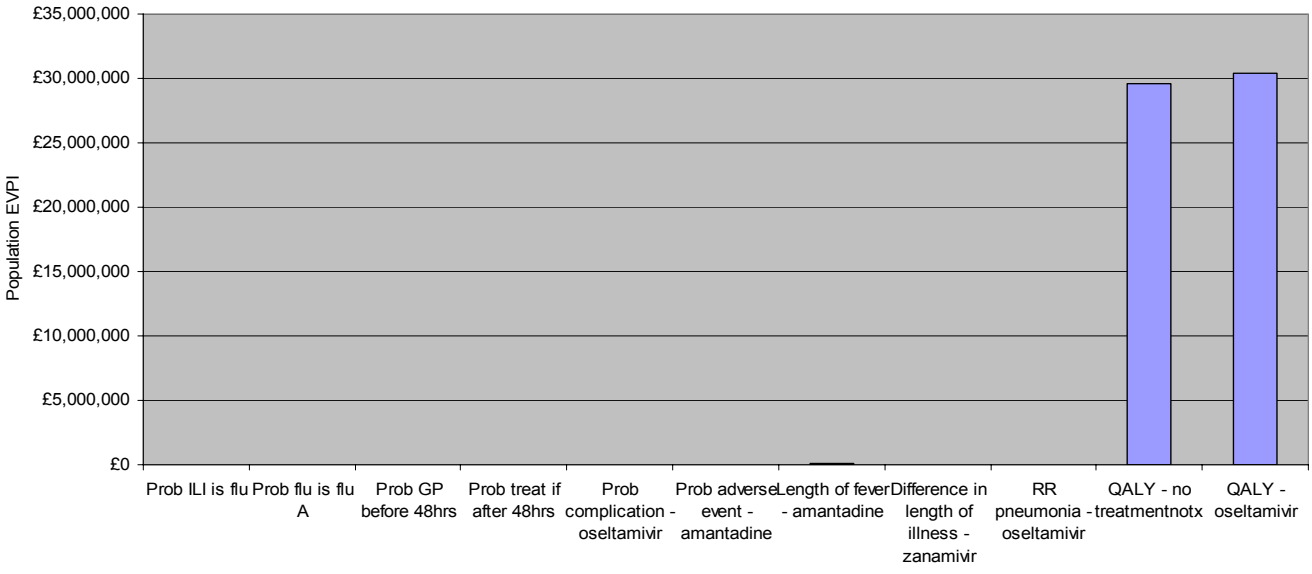


Figure 5.5: Partial EVPI for individual parameters (Lifespan 10yrs, Threshold £30,000)

Table 5.5: Population EVPI (Threshold £30,000)

	Base Case			Sensitivity 1			Sensitivity 2		
	5yrs	10yr	15yrs	5yrs	10yr	15yrs	5yrs	10yr	15yrs
Global	38,173,610	66,699,153	88,015,097	1,737,868	3,036,504	4,006,922	39,198,817	68,490,454	90,378,869
Partial Parameters									
Prob ILI is influenza	0	0	0				0	0	0
Prob influenza A	1,099	1,919	2,533				0	0	0
Prob consult GP <48hrs	0	0	0				0	0	0
Prob drug if after 48hrs	0	0	0				0	0	0
Prob complication – no treatment	0	0	0				0	0	0
Prob complication - amantadine	0	0	0				0	0	0
RR comp - oseltamivir	0	0	0				0	0	0
RR ad event – amantadine	0	0	0				572	999	1,316
Prob pneumonia – no treatment	0	0	0				0	0	0
Prob pneumonia - amantadine	0	0	0				0	0	0
RR pneumonia – oseltamivir	0	0	0				0	0	0
Reduction in fever - amantadine	82,742	144,575	190,776	464,708	811,963	1,071,454			
Reduction in length of illness - zanamivir	0	0	0	0	0	0			
Reduction in length of illness - oseltamivir	242,574	423,840	559,291	963,479	1,683,446	2,221,447			
QALDs - no treatment	16,932,882	29,586,115	39,041,346	0	0	0			
QALDs - oseltamivir	17,420,460	30,438,042	40,165,536						
Prob consult GP (NIs available)	NA	NA	NA	NA	NA	NA	2,720,431	4,753,295	6,272,368
Groups of parameters									
GP surveillance	0	0	0						
QALY study	25,328,373	44,255,206	58,398,435						
Oselatmivir RCT	245,168	428,372	565,273						
Zanamivir RCT	0	0	0						
Amantadine RCT	132,399	231,334	278,247						

A number of combinations of groups of parameters, corresponding to different potential study designs, were also examined (reported in table 5.5). Any study including utility measurement has the greatest impact but it also interesting to note that the expected value to be gained from additional RCTs of any of the three drugs is likely to be negligible unless quality of life measurement is included as part of the trial.

5.3.3 EVPI for structural sensitivity analysis 1.

The analysis was repeated for structural sensitivity analysis 1, described above. The correlation between quality of life and symptom days used in this specification reduces the degree of uncertainty in the model, as shown in figure 5.3, and inevitably this results in a lower global EVPI. Figure 5.6 shows that this figure falls substantially from £66m in the base case analysis to £3m assuming a lifetime of 10yrs and a threshold of £30,000. EVPI is not simply an increasing function of λ in this scenario since the dominant effect of uncertainty in the quality of life values is not present.

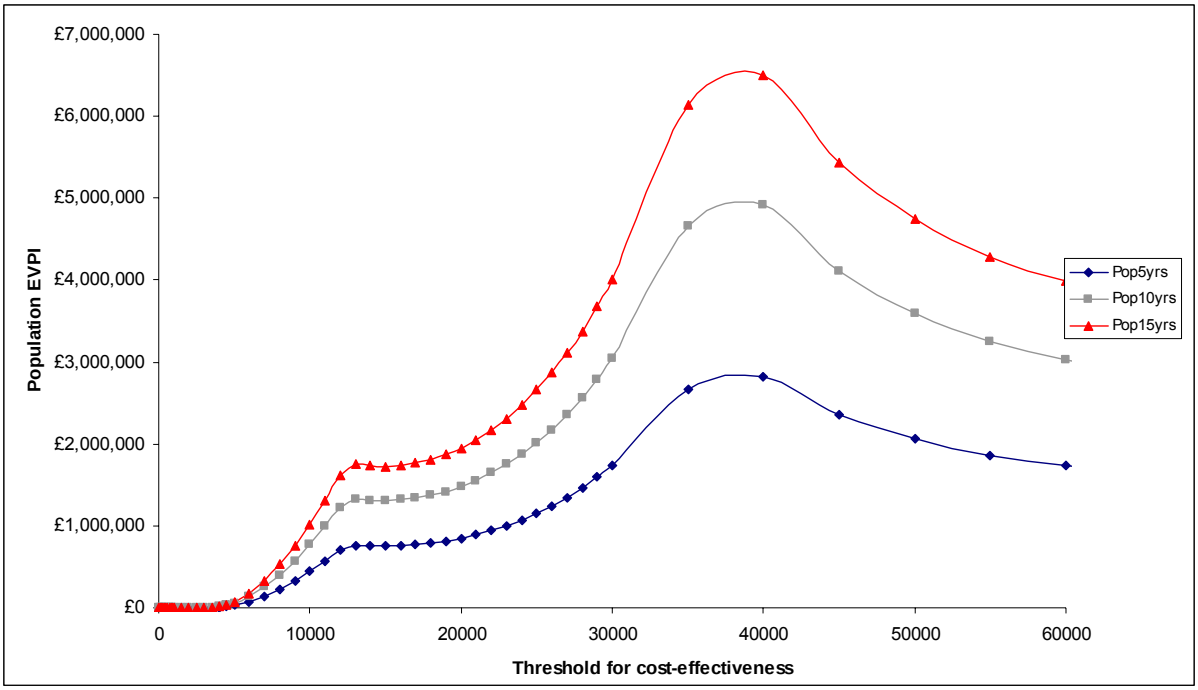


Figure 5 6: Global EVPI per person – structural sensitivity analysis 1.

The EVPI for four separate parameters, which differ in their impact on this specification of the model compared to the base case, was calculated and results are shown in figure 5.7.

5.3.4 EVPI for structural sensitivity analysis 2.

Additional value of information analyses were run using the specification of the model described as sensitivity analysis 2, above. Since QALY scores entered in a similar fashion to the base case, these effects continue to dominate and are not shown again separately. Table 5.5 does highlight the value of information associated with the key parameter that differs in this specification from the base case – the increase in the probability that individuals consult the GP when they have ILI if neuraminidase inhibitors are available. Where the lifetime of the technology is assumed to be 10yrs this value is approximately £5m assuming a threshold value of £30,000.

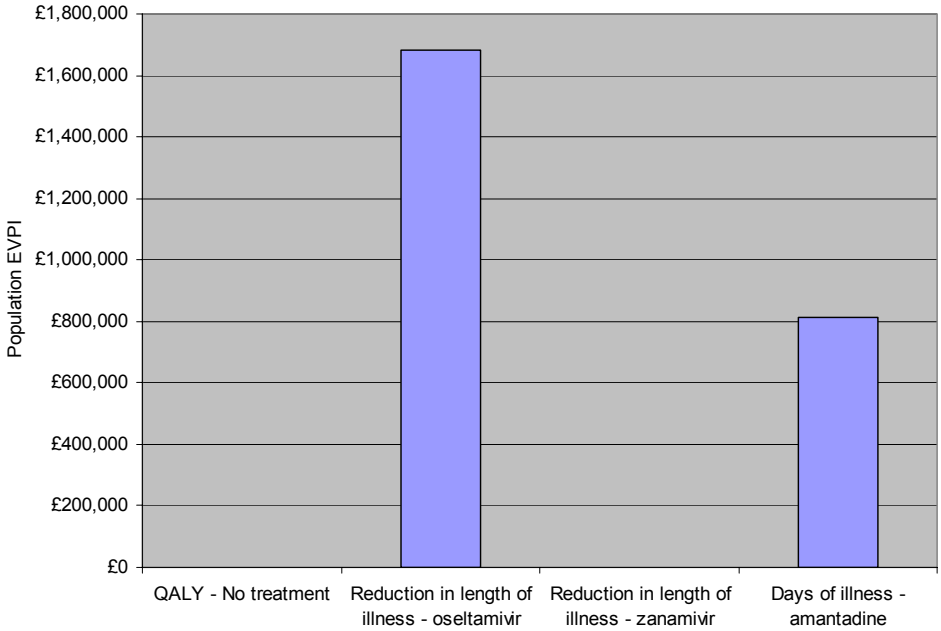


Figure 5.7: Partial EVPI per person – structural sensitivity analysis 1. (Lifetime 10yrs, threshold £30,000)

5.4 Discussion

The base case model is associated with substantial uncertainty which is driven predominantly by the values associated with the QoL for influenza. EVI analysis supports the research recommendations made alongside NICE guidance, that quality of life measurement of people with influenza should be a research priority. Furthermore, the analysis indicates the importance of identifying the quality of life impacts of drugs which potentially make relatively small differences over a period of a few weeks. Days of illness with influenza are not homogenous either on or off treatment.

The analysis also indicates that additional research on parameters that influence the cost effectiveness of amantadine (although this also determines the ICER for other interventions) may be valuable, specifically length of illness, the rate of adverse events and the probability that influenza is influenza A. Other parameters of potential importance are the risk of complications requiring a repeat GP consultation and the risk of pneumonia for oseltamivir, although these are in tun highly dependent on the threshold value of a QALY. At £30,000 the EVPI for these parameters is negligible. Maximum values are achieved around the mean ICER for oseltamivir of £38,000.

The specification of the model is itself subject to uncertainty that is not reflected in analysis of the base case. Uncertainty reflected in both CEAcc and EVI analysis is limited to that which is defined by parameter probability distributions. Several alternative specifications of the model are presented as sensitivity analyses but the evidence base on which these specifications are based and/or the assumptions they entail are not considered sufficiently robust for incorporation in the base case.

Firstly, it is feasible that influenza treatments reduce the likelihood of severe complications leading to hospitalisations or mortality and the inclusion of such effects, based on the proxy measure of pneumonia, substantially improves the cost-effectiveness ratios for both neuraminidase inhibitors (not reported here).

Secondly, the uncertainty in the base case model is driven substantially by quality of life data for oseltamivir and no treatment. An alternative specification that reduces the level of uncertainty is presented which generates quality of life scores for each of the drug interventions as a function of length of illness. However, it must be recognised that this apparent reduction in uncertainty is achieved only at the expense of over-simplification of the determinants of quality of life in this scenario, that is, only the impact of duration of a narrow set of symptoms are relevant.

Thirdly, the potential impact of increasing GP consultations is explored both in terms of cost-effectiveness and EVI. This analysis in particular is included to illustrate the sensitivity of both central estimates of cost-effectiveness and uncertainty analysis rather than directly inform decision makers, since the parameter values are somewhat arbitrary. The findings are important however in highlighting the potential value of additional research, particularly since such research could be relatively low cost to undertake.

Finally, in this analysis it has been assumed that additional information on particular parameters generates value only through the reduction in decision uncertainty relating to the treatment of influenza in otherwise healthy adults. These drugs have prophylactic uses, both seasonal and post-exposure, and are also available for other patient groups (at-risk adults and, in the case of oseltamivir, children). Assessment of cost-effectiveness in these scenarios draws on several parameters reported here [v] and therefore the EVPI figures potentially underestimate the true value of eliminating uncertainty. One of the future challenges for value of information methods is to incorporate these additional considerations.

References:

1. Nicholson KG. Impact of influenza and respiratory syncytial virus on mortality in England and Wales from January 1975 to December 1990. *Epidemiology and Infection* 1996;116:51-63.
2. Health Protection Agency. Frequently asked questions on flu. http://www.hpa.org.uk/infections/topics_az/influenza/flufaq.htm#vacc. [14th October 2003]
3. Burls A, Clark W, Stewart T, Preston C, Bryan S, Jefferson T, et al. Zanamivir for the treatment of influenza in adults: a systematic review and economic evaluation. Health Technology Assessment 2002;6(9).
4. Flu - zanamivir (Relenza) (N0 15) http://www.nice.org.uk/cat.asp?c=426 [19th February 04]
5. NICE. Guidance on the use of zanamivir, oseltamivir and amantadine for the treatment of influenza. February 2003.
6. Turner D, Wailoo A, Nicholson K, Cooper N, Sutton A, Abrams K. Systematic review and economic decision modelling for the prevention and treatment of influenza A and B. Health Technology Assessment 2003;7(35).
7. Burls A, Clark W, Stewart T, Preston C, Bryan S, Jefferson T et al. "Zanamivir for the treatment of influenza in adults". NICE 2000. (Check what is in their HTA report and cite this if appropriate)
8. O'Brien B, Goeree R, Blackhouse G, Smieja M, Loeb M. Oseltamivir for treatment of influenza in healthy adults: pooled trial evidence and cost-effectiveness model for Canada. Value in Health 2003;6:116-125.
9. Jefferson TO , Demicheli V, Deeks JJ, Rivetti D. Amantadine and rimantadine for preventing and treating influenza A in adults. Cochrane Database of Systematic Reviews 2000;2.
10. Nicholson KG, Aoki FY, Osterhaus ADME, Trottier S, Carewicz O, Mercier CH et al. Efficacy and safety of oseltamivir in treatment of acute influenza: a randomised controlled trial. Lancet 2000;355:1845-50.
11. Treanor JJ, Hayden FG, Vrooman PS, Barbarash R, Bettis R, Riff D et al. Efficacy and safety of the oral neuraminidase inhibitor oseltamivir in treating acute influenza: a randomized controlled trial. US Oral Neuraminidase Study Group. *JAMA* 2000;283:1016-24.
12. www.fda.gov/cder/approval/index.htm
13. Unpublished trial
14. Murray CJ, Lopez AD, (eds). The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disabilities from Diseases, Injuries and Risk Factors in 1990 and Projected to 2020. Cambridge MA.
15. MVH Group. The Measurement and valuation of health, First Report on the Main Survey. York: Centre for Health Economics, University of York, 1994.
16. Netten, J. and Curtis, J. Unit Costs of Health and Social Care . 2001. Canterbury, PSSRU, University of Kent.
17. British National Formulary (39). London: British Medical Association and Royal Pharmaceutical Society of Great Britain, 2000.
18. NHS Executive. The New NHS - Reference Costs. Accessed at www.doh.gov.uk/nhsexec/refcosts.htm. 2001.
19. Personal communication, Prescriptions Pricing authority.
20. RCGP. Dr Douglas Fleming personal communication.
21. Public Health Laboratory Service. 2001. Now available from the Health Protection Agency http://www.hpa.org.uk/infections/topics_az/influenza/tables1.htm [accessed 23rd Feb 04].
22. Fleming.D.M. The contribution of influenza to combined acute respiratroy infections, hospital admissions, and deaths in winter. Communicable Disease and Public Health 2000;3:32-8.
23. Meier CR, Napalkov PN, Wegmuller Y, Jefferson T, Jick H. Population-based study on incidence, risk factors, clincial complications and drug utilisation associated with influenza in the United Kingdom. Eur J Clin Microbiol Infect Dis 2000;19:834-42.

24. Office for National Statistics. Population estimates for 1996 for England and Wales. 2002.
25. Ross AM, Kai J, Salter R, Ross J, Fleming DM. Presentation with influenza-like illness in general practice: implications for use of neuramindase inhibitors. Comm Dis Publ Hlth 2000;3 :256-66.
26. Weekly Returns Service. Annual Report. Birmingham; Birmingham Research Unit, Royal College of General Practitioners, PHLS;1999.
27. Weekly Returns Service. Annual Report. Birmingham; Birmingham Research Unit, Royal College of General Practitioners, PHLS;2000.
28. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Economics 2001;10:779-787.
29. Cooper NJ, Sutton AJ, Abrams K, Wailoo AJ, Turner DA, Nicholson KG. Effectiveness of neuraminidase inhibitors in the treatment and prevention of influenza A and B: systematic review and meta-analysis of randomised controlled trials. BMJ 2003;326:1235-9.
30. Office for National Statistics. Population estimates for 1996 for England and Wales. 2002

6. Liquid-based cytology for cervical screening

6.1 Background

6.1.1 Background to disease/condition

Cervical cancer is the 3rd most common form of cancer amongst women, with an incidence rate of around 9.3 per 100,000, and a mortality rate of around 3.7 per 100,000 per annum. The NHS introduced its Cancer Screening Programme in 1988, since which time the incidence of and mortality resulting from cervical cancer have decreased by 40%. The coverage of the screening programme has increased steadily since its introduction, and in 2002/3, 3.7 million women were screened; the majority of these women attended after an invitation from the screening programme.[1]

The main risk factors for cervical cancer include age and the presence of the human papillomavirus (HPV), a sexually-transmitted disease. Cervical cancer is uncommon in women under 25 and over 64, and hence women outside these age bounds are only screened in exceptional circumstances. Under current guidelines, women in England and Wales begin screening at the age of 25, and are screened either every three or five years, up to the age of 64. Screening interval policies vary between health authorities, though it is now becoming increasingly common for women up to the age of 50 to be screened every three years, and every five years thereafter, up to the age of 64. Beyond this age, women are discharged from the screening programme unless their most recent test result shows an abnormality that warrants follow-up. The progression of cervical cancer is particularly slow, taking up to ten years to develop through several pre-invasive stages to full invasive cancer.

Until recently, the most widely used technology for carrying out cervical screening has been the Pap smear, which is a quick method of gathering cells from the cervix for further examination. This method has a tendency to produce smears which are “inadequate”, that is there are insufficient cervical cells to facilitate a full examination, or the sample is contaminated with blood or mucus and so cannot be analysed adequately. The most recent data suggests that around 9% of all smears taken are found to be inadequate.[2] This has cost implications for the screening programme because it means a large number of women have to be re-screened until an adequate smear is obtained.

Liquid-based cytology (LBC) is a novel method of providing cervical cell samples for examination, which has recently been piloted and is intended to become the primary method of cervical screening. Its major advantage over the conventional Pap smear is that it reduces the proportion of tests which are inadequate from 9% to around 2%, by improving the ability to examine poorly taken samples. Samples are collected using a brush-like device rather than a spatula (which is used in Pap screening), the cells from which are then suspended in preservative fluid. Once in the cytology laboratory, any blood or other debris is removed from the suspension before a layer of the cells are deposited onto a slide for examination. LBC has also been found to improve the sensitivity and specificity of cervical screening, and the fact that the cells are held in a preservative fluid has the added advantage that further tests can be carried out at a later date, for example, testing for the presence of HPV.

6.1.2 Technology Assessment Review

In 1999 the NCCHTA commissioned the School of Health and Related Research (ScHARR) at the University of Sheffield to carry out a rapid and systematic review of the effectiveness and cost effectiveness of LBC for cervical screening compared with conventional smear testing.[3] The review and modelling considered three treatment strategies: no screening, conventional Pap screening and LBC screening. For the latter two policies, a number of scenarios were modelled to assess the impact of different screening intervals on the results. Full probabilistic sensitivity analysis was undertaken to examine the uncertainty in model parameters.

Following the submission of the original assessment, NICE did not recommend the use of LBC due to uncertainty surrounding its cost-effectiveness. The assessment was updated in

2003[4] to incorporate data from a series of LBC pilot studies and from literature published during the interim period. This review included updated analyses of the effectiveness and cost-effectiveness of LBC compared to Pap testing.

6.1.3 Current NICE guidance

NICE published guidance on the use of liquid-based cytology for cervical screening in 2003,[1] in which two key recommendations were made:

1. LBC should be used as the primary means of processing samples in the cervical screening programme in England and Wales
2. The NHS Cancer Screening Programme and Cervical Screening Wales may consider evaluating further the different LBC products available

It is anticipated that the roll-out of LBC across England and Wales will take some time due to substantial implications in terms of staff training. The second recommendation relates to the decision as to which LBC product is preferable, which is yet to be made. The guidance is due to be reviewed again in August 2006.

6.1.4 Research recommendations

A number of recommendations have been made for further research into the use of LBC in cervical screening, from the NICE Guidance[1] and the recent update of the HTA monograph.[4] The NICE guidance highlighted four main areas for further research: -

1. Commissioning of high-quality studies to examine differences in performance between existing LBC products;
2. Validation of the number of cells required per sample which are required to establish the adequacy of smears;
3. Further reviews of LBC, clinical data relating to the sensitivity, specificity and rate of inadequate smears should be provided for EasyPrep, Cytoscreen and any future devices;
4. Further evaluation of automated technologies for the analysis of cervical samples.

The analysis conducted in the 2003 updated assessment[4] provided a greater degree of certainty concerning the potential cost-effectiveness of LBC when compared to conventional Pap screening. It proposed that, although the modelling results had provided strong evidence that LBC is an economically attractive screening method, a full cost-effectiveness study of LBC based on a trial of its introduction in a low-prevalence population would provide more definitive answers. However, undertaking such a study would not be justifiable, given the large expenditure involved. Further research was recommended in the area of utility assessment, with a particular focus on the short-term impact of false-positive screening results.

This EVPI assessment considers both the results of the original NICE assessment[3] and the updated NICE assessment[4] following the completion of the LBC pilot.[5] The analysis presented here also incorporates the corresponding results obtained through an update of the previous models, to assess the impact of the use of the Reference Case[6] discount rates and measurement of health benefits.

6.2. Methods

The LBC model was produced with the aim of assessing the impact of introducing LBC compared to conventional Pap screening, with outcomes focused on the incidence of cervical cancer, the associated mortality, and the cost-effectiveness. A cohort of 100,000 women aged 18 was used, whose progression up to the age of 95 was simulated. The model measures health outcomes, resource utilisation and costs for the whole cohort. Life tables are used within the model the represent the risk of age-specific all-cause mortality.

The model uses the state transition methodology to simulate the natural history of cervical cancer, upon which the impact of screening is superimposed. The model uses a six-monthly cycle length using progression data from Sherlaw-Johnson et al.[7] A constant cancer incidence rate is assumed between ages 18 and 64 years. No further incident cases are assumed to occur beyond the age of 64.

The model uses five health states to reflect the progression of the disease from the initial pre-invasive stages through to invasive cancer. The pre-invasive stage of the disease, which is known as cervical intraepithelial neoplasia (CIN) was classified into three states: CIN 1, CIN 2 and CIN 3. The model assumes that regression from CIN to the disease-free state may only occur from CIN 1, and that the disease progresses through each pre-invasive stage in the absence of any screening / treatment intervention. The incidence of the disease was assumed to be the onset of CIN 1.

The model assessed the impact of three screening policies: LBC, Pap smears, and no screening policy. Screening is assumed to taken up by 85% of the cohort, and women are assumed to either attend screening at regular intervals, or not at all. Costs within the model concern the screening, diagnosis and treatment of patients, using unit costs of each smear test, LBC test, colposcopy (a more detailed examination of the cervix) and treatment of both pre-invasive disease and invasive cancer. Marginal costs of LBC over Pap screening were estimated using the associated increases in the costs of consumables and the capital cost of equipment.

The model produces the following outcomes:

- Health benefits of LBC versus the Pap screening and no screening policies, including number of cancers avoided, life years gained, quality adjusted life years gained;
- Resource use, measured by the number of smears and colposcopies required;
- Health economic outcomes including cost per cancer avoided, cost per life-year gained and cost per quality adjusted life year gained.

The ScHARR model uses around thirty parameters which relate to discount rates, probabilities of transiting between health states, effectiveness of treatment, specificy and sensitivity of the screening tests, utility associated with health states, cost parameters, and test result characteristics.

Uncertainty in the model was characterised through assigning statistical distributions with unique characteristics to each model parameters. For simplicity, triangular distributions were assigned to reflect the uncertainty in each parameter, and multivariate Monte Carlo sensitivity analysis was undertaken. Expected value of perfect information (EVPI) analysis was performed to identify those areas in which future research is expected to yield the greatest value. A global EVPI estimate was derived from the model, together with estimates of partial EVPI for individual parameters within the model.

As previously mentioned, the original model was updated in 2003 to incorporate new data from recent literature and pilot studies of LBC. The earlier model included probabilistic sensitivity analysis and EVPI analysis; however, the 2003 model incorporated only limited sensitivity analysis. For the purposes of the analysis presented here, the 2003 model has been updated to include the capacity for EVPI analysis, thus enabling a direct comparison to be made between the two models in terms of the reduction in uncertainty resulting from the collection of additional information during the interim period. The 2003 model measured health benefits in terms of quality adjusted life-years (QALYs); this has since been updated to incorporate the ability to measure these as either life-years gained (LYG) or QALYs, to allow comparison with the original model and to reflect the reference case criteria.[6]

Given the lack of evidence regarding quality of life associated with invasive cancer, a fixed utility value of 0.6 was used in the 2003 model. Owing to the inherent uncertainty surrounding this estimate, the uncertainty surrounding this parameter was described using a triangular distribution with minimum and maximum values of 0.5 and 0.7 respectively.

For the purposes of this pilot study, three separate (but sequential) EVPI analyses were undertaken using the aforementioned three models. The differences between these 3 models are described in Table 6.1.

Table 6.1 Differences between the 3 LBC models

Scenario	Model used	Discount rates used	Health outcome	New data used?
1	Original (2000)	Costs @ 6%, utilities @ 6%	LYG	No
2	Update (2003)	Costs @ 6%, utilities @ 6%	LYG	Yes
3	Update (2003)	Costs @ 3.5%, utilities @ 3.5%	QALYs	Yes

The comparison between Scenarios 1 and 2 is envisaged as a means of quantifying the reduction in uncertainty due resulting from the additional data collected from the pilot studies.[5] Scenario 3 represents the current decision uncertainty, and is closely based upon NICE Reference Case.[6]

The 1-level EVPI algorithm was used in all cases in carrying out the EVPI analysis, necessitating a strong degree of linearity between the model inputs and outputs. There is naturally some uncertainty in precisely how linear a model needs to be in order for the 1-level algorithm to be both appropriate and accurate. The linearity of the model was assessed using linear regression, with the net benefits (NBs) as the dependent variables and the input parameter values as the dependent variables. The regression analysis yielded an adjusted r2 value of 0.91, suggesting a relatively strong degree linearity. Whilst this was considered sufficiently linear to allow the 1-level EVPI algorithm to be used, it is important to note that the absence of perfect linearity can affect the ability of the 1-level algorithm to derive accurate estimates. Multivariate Monte-Carlo sampling was used to generate 5,000 samples of the base-case results, with the net benefits of each screening intervention being the outputs in each case.

The original model simulated a cohort of 100,000 women aged 18, with screening commencing at the age of 21; in the updated model, a cohort of 360,000 was used to more accurately reflect the current female population in England and Wales of this age and who are therefore eligible for screening.

6.3 Results

6.3.1 Adoption decisions

Tables 6.2 **-** 6.4 present the central estimates of cost-effectiveness for Scenarios 1-3 respectively. Table 6.2 shows the incremental cost-effectiveness results for Scenario 1.

Table 6.2 Cost-effectiveness results for Scenario 1

Screening Policy	Marginal discounted life-days gained compared with no screening	**Incremental life-days gained**	Average discounted lifetime cost per woman	**Incremental discounted lifetime cost**	Incremental cost per life-year gained
No screening	-	-	£2.95	-	-
Screening at 5 years - Pap	49.09	49.09	£164.02	**£161.07**	£1,197
LBC	49.61	0.51	£165.56	£1.53	£1,095
Screening at 3 years - Pap	50.65	1.04	£255.36	£89.80	£31,519
LBC	50.98	0.33	£257.66	£2.30	£2,522

Table 6.2 shows that the use of LBC as opposed to Pap testing increases both costs and health benefits. The incremental cost per LYG for LBC is £1,095 for LBC screening using a 5-year screening interval, and £2,522 using a 3-year screening interval. In order to provide a

direct comparison between the cost-effectiveness results of the original NICE assessment and those following the collrection of further information within the LBC pilot study,[5] Scenario 2 uses a discount rate of 6% for both costs and LYGs. Table 6.3 shows that the incremental cost per LYG for LBC is £12,718 over a 5-year screening interval and £44,280 using a 3-year screening interval. It is evident from these results that the re-estimation of the unit costs of LBC has a substantial impact upon the central estimates of cost-effectiveness for the LBC screening programmes.

Table 6.3 Cost-effectiveness results for Scenario 2

Screening Policy	Marginal discounted life-days gained compared with no screening	**Incremental life-days gained**	Average discounted lifetime cost per woman	**Incremental discounted lifetime cost**	Incremental cost per life-year gained
No screening	-	-	£3.15	-	-
Screening at 5 years - Pap	6.94	6.94	£60.97	**£57.82**	£3,040
- LBC	7.08	0.14	£65.86	£4.89	£12,718
Screening at 3 years - Pap	7.29	0.21	£92.89	£27.03	£47,409
- LBC	7.35	0.03	£100.56	£7.67	£44,280

The Scenario 3 considers the impact of using Reference Case[6] discount rates (3.5% for both costs and health benefits), and measured health benefits in terms of QALYs, as opposed to LYG. Table 6.4 summarises the results of the cost-effectiveness analysis.

Table 6.4 Cost-effectiveness results for Scenario 3

Screening Policy	Marginal discounted quality-adjusted life-days gained compared with no screening	**Incremental quality-adjusted life-days gained**	Average discounted lifetime cost per woman	**Incremental discounted lifetime cost**	Incremental cost per QALY
No screening	-	-	£3.15	-	-
Screening at 5 years - Pap	54.04	54.04	£60.97	£57.82	£390
- LBC	54.92	0.88	£65.86	£4.89	£2,033
Screening at 3 years - Pap	D	D	D	D	D
- LBC	55.09	0.17	£10.06	£34.70	£73,388

D: Dominated

It is clear that the inclusion of the reference case discount rates and the valuation of health outcomes as QALYs has a considerable impact upon the cost-effectiveness of these screening programmes.

Figure 6.1 shows the cost-effectiveness acceptability curve (CEAC) for the analysis carried out in Scenario 3. This analysis discounted both costs and QALYs at 3.5%.

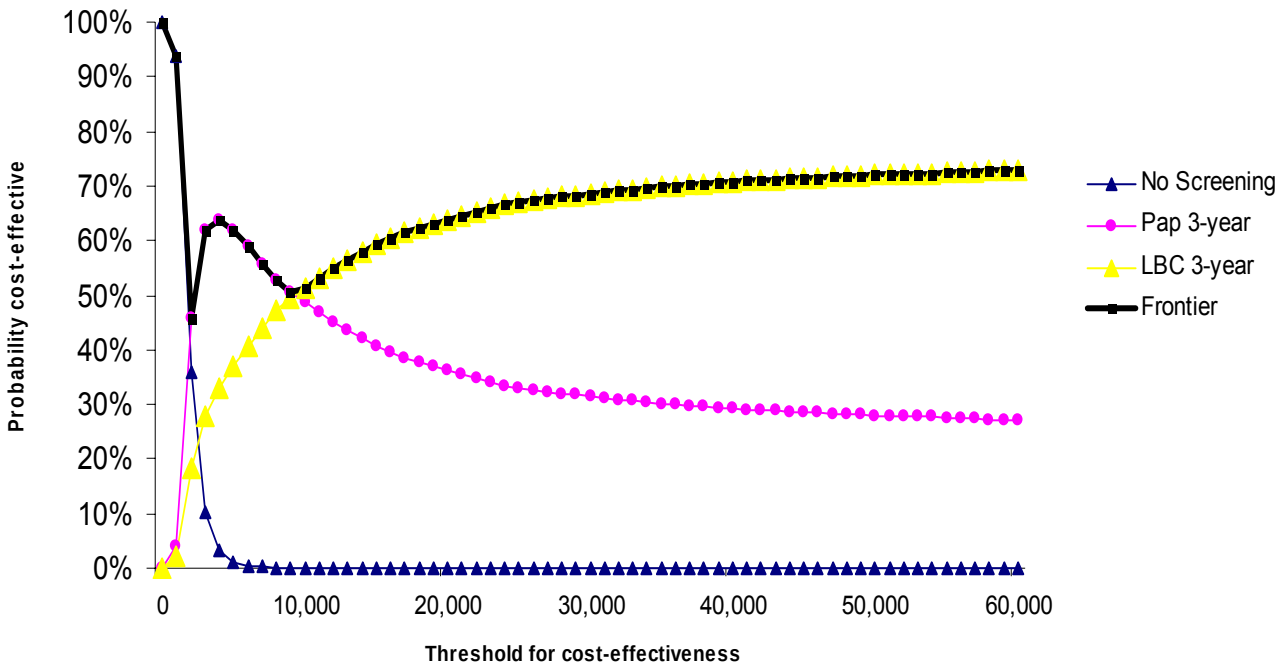


Figure 6.1 Scenario 3 - Cost-effectiveness acceptability curve

Beyond a threshold of around £10,000, LBC is always the optimal screening strategy. The probability of LBC being cost-effective at a threshold of £30,000 is approximately 70%, with Pap testing only 30% likely to be cost-effective. Beyond this threshold, the probabilities of LBC and Pap testing remain relatively stable at these values. The no screening strategy is consistently unfavourable.

The analysis from Scenario 1 used a discount rate of 6% for both costs and LYG respectively. From the EVPI analysis, a global EVPI value of £28.24 per patient was derived (using a cost-effectiveness threshold of £20,000), which equates to a population estimate of £2,824,000 (using a population base of 100,000 women). Sensitivity analyses were carried out to determine the impact on the global EVPI of varying the cost-effectiveness threshold; the EVPI was found to decline approximately linearly either side of the £20,000 threshold, declining to a value of around £1.8 million at a threshold of around £50,000.

Partial EVPI analyses were also carried out to identify those parameters within the model around which further research would be beneficial. These partial EVPI estimates are summarised in the Figure 6.2.

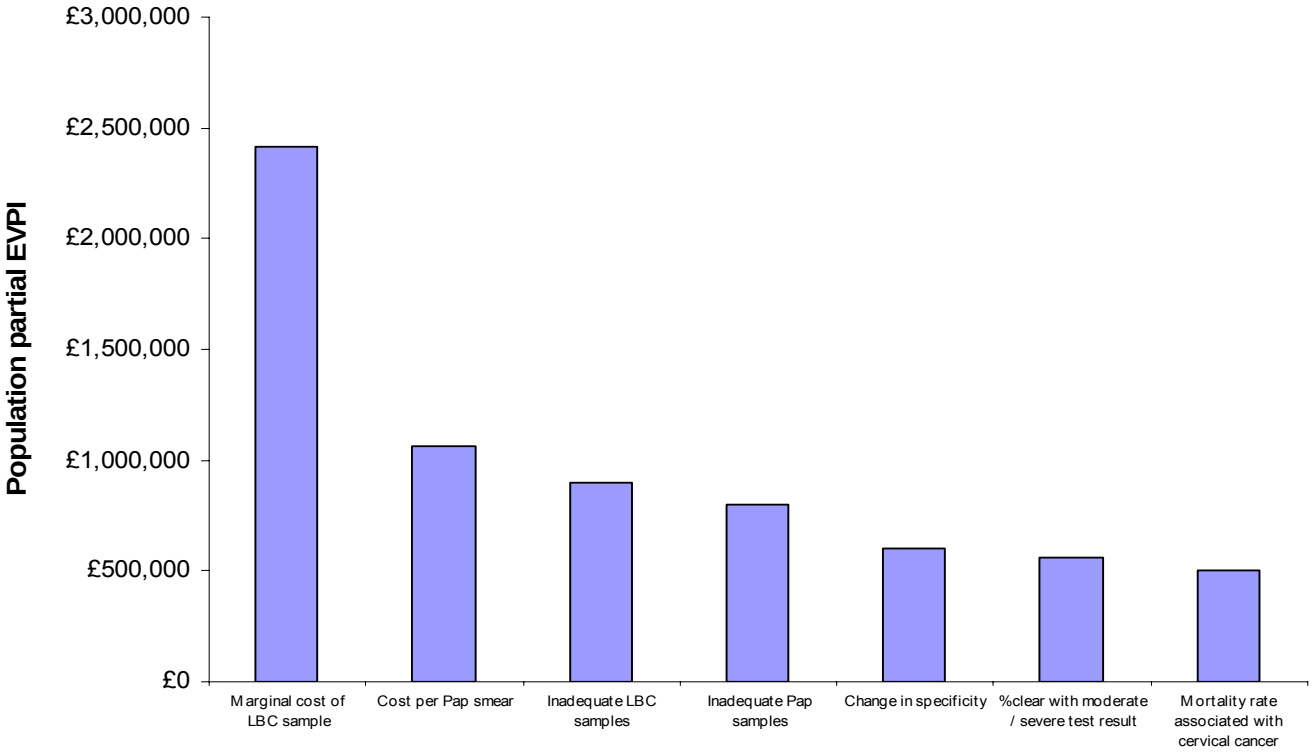


Figure 6.2 Partial EVPI estimates (Scenario 1)

By far the most value can be found in the parameter which measures the marginal cost of taking a sample using LBC, compared to via the Pap method. This reflects the uncertainty in the maximum and minimum parameters used in the triangular distribution representing this cost. Other areas of further research suggested by the analyses include the inadequacy rates of both LBC and Pap testing, the improvement in specificity associated with LBC (compared to Pap) and the mortality rate associated with cervical cancer.

One of the key areas of interest is in how the additional data from the LBC pilot study reduced the uncertainty surrounding the adoption decision. The EVPI analysis from Scenario 2 resulted in a global EVPI of £2.80 per patient, equivalent to a total population EVPI of £280,000 (based on a population of 100,000). Figure 6.3 shows the population EVPI for Scenario 2 across a range of willingness to pay thresholds. The inclusion of this additional information within the Scenario 2 model suggests that there is limited benefit in carrying out further research as LBC is almost always preferable to conventional Pap screening . This is supported by the results of partial EVPI analyses, which suggests that further research on only two parameters (the mortality rate associated with cervical cancer, and the marginal cost of taking a sample using LBC) is likely to yield any further value..

The population EVPI value of £280,000 can be read directly off the graph at a threshold of £20,000. Broadly speaking, the EVPI continues to decrease steadily beyond this threshold. This plot is somewhat different to the plot derived from Scenario 1 in terms of the peak in the EVPI. The Scenario 2 plot shows a peak of approximately £1.5 million at a threshold of £5,000, compared with a peak of £2.8 million at a threshold of £20,000 for Scenario 1. These results demonstrate the value of the additional data obtained from the pilot study. Partial EVPI analyses were also conducted within Scenario 2, the results of which are shown in Figure 6.4.

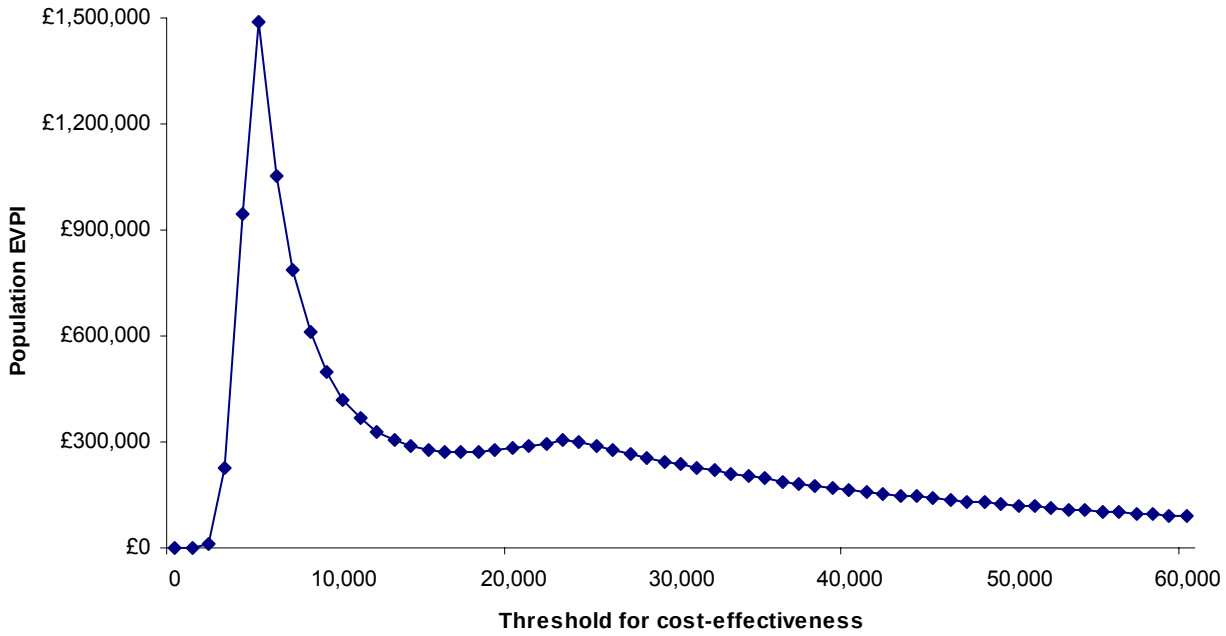


Figure 6.3 Population EVPI (Scenario 2)

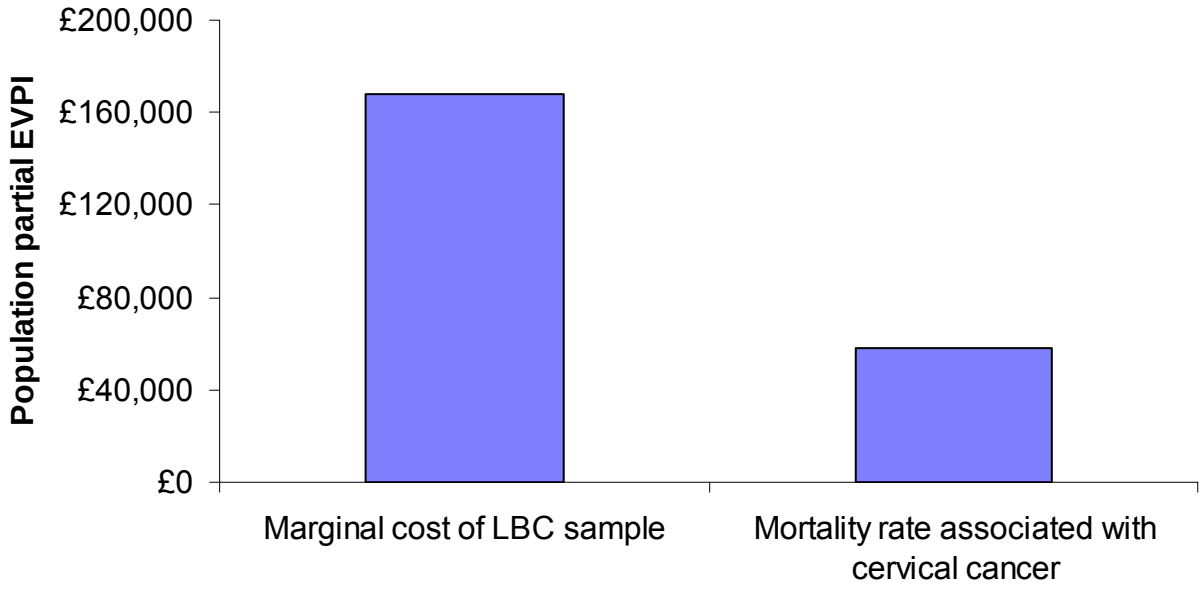


Figure 6.4 Partial EVPI estimates (Scenario 2)

Only two parameters were found to have any value associated with further research. It is evident from Figures 6.3 and 6.4 that further research concerning the cost of obtaining a LBC sample and the mortality rate associated with cervical cancer are likely to yield some value, but this value has decreased in Scenario 2 by an approximate factor of 10.

The analyses for Scenarios 1 and 2 both assumed a relevant population for the decision of 100,000 women; Scenario 3 however assumed a relevant population of 360,000 women to reflect the true female population eligible for screening. The per-patient global EVPI for Scenario 3 was £7.38. Figure 6.5 shows the population EVPI estimates generated under Scenario 3 across a decision lifetime of 5, 10 and 15 years.

Assuming a willingness to pay threshold of £20,000 and decision lifetime of 10 years, the overall decision EVPI is just below £20,000,000 across the entire relevant population. It should be noted here that as the value of λ increases, the so too does the population EVPI; the difference between Scenarios 1-3 is due to the way in which the relevant population is defined. In Scenarios 1 and 2, the population is assumed to be the number of women eligible for screening within a given year (and is thus not discounted), whereas for Scenario 3, a new cohort of screen-eligible women are included for each year and discounted over the selected decision lifetime.

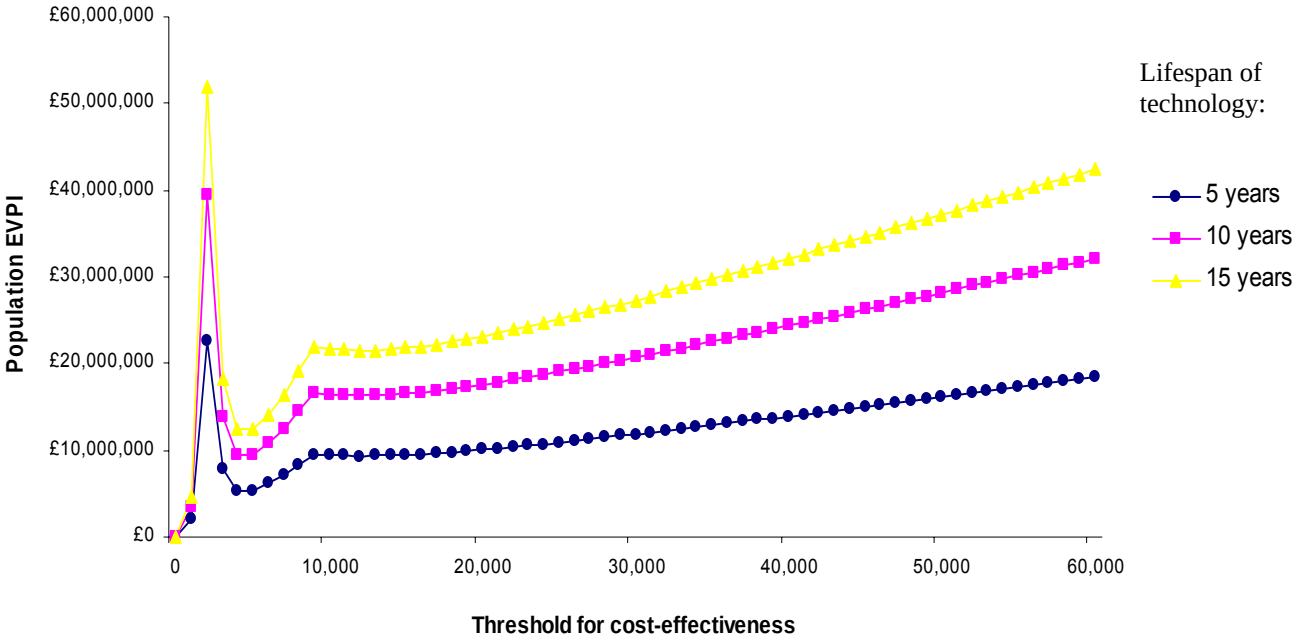


Figure 6.5 Population EVPI across different decision lifetimes (Scenario 3)

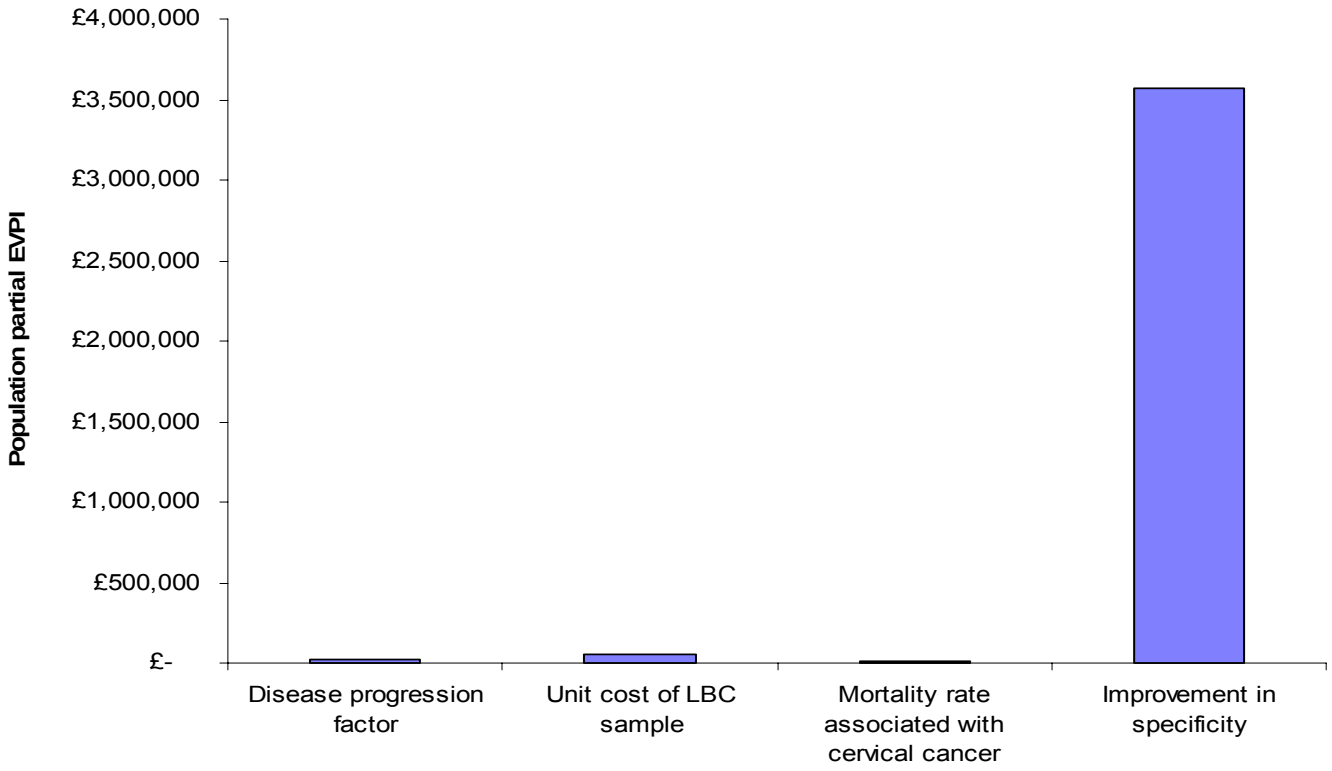


Figure 6.6 Partial EVPIs for Scenario 3

Figure 6.6 shows that the areas in which further research may be warranted in particular in terms of the costs associated with taking a LBC sample and the improvement in specificity of LBC over the conventional Pap smear test. Whilst some value may be associated with obtaining further information on cancer progression and the mortality rate associated with invasive cancer, this value is low and is unlikely to justifiy the cost of obtaining such information. It is reasonable to suggest that the value of further information on the cost of LBC and the improvement in specificity of LBC may outweigh the costs of undertaking such research.

6.4 Discussion

Prior to the pilot study, numerous areas of valuable further research were identified, in particular with regard to the costs of taking a sample using LBC, and the inadequacy rates of LBC and Pap testing. A comparison of Figures 6.2 and 6.4 illustrates the reduction in uncertainty of these particular parameters, and demonstrate the value of the pilot study (the updated model (Scenario 2) incorporated data from the pilot study up until mid-2002, by which time the study was estimated to have cost approximately £1.2 million). Indeed, only two parameters were found to have any value from the analyses in Scenario 2. Although some structural differences exist between the two models used in Scenarios 1 and 2 in the way in which costs are estimated, a significant reduction in the uncertainty in the model was derived through the addition of the pilot study data.

The use of the Reference Case discount rates (Scenario 3) has been shown to impact significantly on the expected value of further research. The analysis in Scenario 2 produced a population EVPI estimate of £2.80 per patient, compared with an EVPI of £7.38 per patient under the analysis in Scenario 3. This difference is attributable to two factors:-

- the measurement of health benefits through QALYs in Scenario 3 rather than life-years gained;
- the alternative discount rates used in Scenario 3.

One of the key uncertainties surrounding this policy decision concerns the identification of the relevant population for the EVPI analysis. For the original NICE assessment (and indeed the comparative analysis presented in Scenario 2), a 'one-off' cohort of 100,000 women was assumed to represent the screen-eligible population. However, owing to unresolved methodological issues concerning the identification of the relevant population, and the implicit assumption of a 1-year decision lifetime, this figure appears arbitrary and may be of limited use in policymaking. A further problem arises however, in that the population affected by the intervention is essentially all asymptomatic women who are eligible for screening. For the sake of consistency across the 6 case studies presented within this report, the current updated EVPI analysis presented in Scenario 3 estimates the population over a 5,10 and 15 year decision lifetime and includes a new cohort of women for each year. It remains unclear however whether this is more appropriate than the population used in the original analysis.

A further methodological issue which is pertinent to this analysis concerns the use of the 1-level EVPI approximation in order to estimate the true value of obtaining perfect information. This method was used because of the computationally expensive nature of using the 2-level sampling algorithm, and whilst there exists a strongly linear relationship between the sampled parameter inputs and the net benefits within the model (r^2 = 0.91), it is unclear whether a higher degree of linearity is required to accurately estimate the true EVPI. It is possible therefore that the EVPI results presented here are distorted owing to this imperfectly linear relationship.

In both the original and updated models, uncertainty in the model parameters was characterised using triangular distributions, which allow a range of values between an upper and lower limit to be sampled. Whilst this approach is open to criticism given the variety of statistical distributions available for characterising uncertainty in specific types of parameters, its justification is that it reflects the inherent uncertainty surrounding the true values of these parameters.

It should be noted that the first 2 scenarios presented within this chapter fail to meet the criteria set out by the new NICE Reference Case[6] on a number of counts. The inclusion of a new revised model which estimates the cost-effectiveness of LBC and Pap smear testing in terms of cost per QALY gained, discounted at a rate of 3.5% means that the updated version of the model closely adheres to the NICE Reference Case, and thus represents the most appropriate analysis for current decision making.

Reference case evaluation: Liquid-based cytology for cervical screening

Element of health technology assessment	Reference case	Criteria met by assessment?	Comments
Defining the decision problem	The scope developed by the Institute	Yes	
Comparator	Alternative therapies routinely used in the NHS	Yes	
Perspective on costs	NHS and PSS	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Synthesis of evidence on outcomes	Based on a systematic review	Yes	
Measure of health benefits	Quality adjusted life years (QALYs)	Yes	
Description of health states for calculation of QALYs	Health states described using a standardised and validated generic instrument	No	There is little robust evidence available , hence assumptions were made regarding the assignment of utilities to health states in the models.
Method of preference elicitation for health state valuation	Choice-based method, for example, time trade-off, standard gamble (not rating scale)	No	
Source of preference data	Representative sample of the public	No	
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes/No	The 3.5% discount rate was agreed following the completion of the original EVPI methods work; the most recent analyses (Scenario 3), however, used these new rates
Equity position	An additional QALY has the same weight, regardless of the other characteristics of the individuals receiving the health benefit	Yes	The use of QALYs in the updated model are discussed in the HTA monograph [4]

References

1. National Institute for Clinical Excellence, Technology Appraisal 69. Guidance on the use of Liquid-Based Cytology for Cervical Screening. National Institute for Clinical Excellence, Technology Appraisal 69 . 2003.
2. Department of Health. Cervical Screening Programme, England 2002-03. *Statistical Bulletin*, 2003/24 . 2003.
3. Payne N, Chilcott J, and McGoogan E. Liquid-based cytology in cervical screening: a rapid and systematic review. *Health Technology Assessment 4(18)*. 2000.
4. Karnon J, Chilcott J, Peters J, and McGoogan E. Liquid-based cytology in cervical screening: an updated and systematic review. *Health Technology Assessment In press*. 2004.
5. Moss SM, Gray A, Legood R, and Henstock E. Evaluation of HPV/LBC Cervical Screening Pilot Studies (First report on evaluation of LBC). 2002.
6. National Institute for Clinical Excellence. Guide to Methods of Technology Appraisal. 2003.
7. Sherlaw-Johnson C, Gallivan S, Jenkins D, and et al. Cytological screening and management of abnormalities in prevention of cervical cancer: an overview with stochastic modelling. Journal of Clinical Pathology 47, 430-435. 1994.

7. Disease-modifying therapies for the management of multiple sclerosis

7.1 Background

7.1.1 Condition and technology

Multiple sclerosis (MS) is a chronic debilitating disease of the central nervous system, which is characterised by progressive disability and intermittent relapse. Evidence suggests that MS results from an autoimmune response, resulting in inflammation, demyelination and axonal loss.[1] MS is one of the most common neurological conditions affecting young adults, and is approximately twice as common in women than men.[1] The prevalence of MS in England and Wales is estimated to be around 110-120 per 100,000, although this varies geographically, with a higher prevalence in the north of England.[2] This translates to between 58,000 and 63,000 cases of MS in England and Wales, although this is conservative in comparison to other published estimates.[2] The annual incidence of MS in England and Wales is estimated to be around 3.8 per 100,000 people (around 2,000 new cases each year).

The management of MS is aimed at improving quality of life through relieving the symptoms of the disease; conventional management typically consists of drug therapy, physiotherapy, psychiatric and social support, and disability aids. Whilst there is no cure for MS, a set of drugs known as 'disease-modifying therapies' (DMTs), namely the interferon-betas and glatiramer acetate, are aimed at slowing disease progression and reducing the number and severity of relapses experienced.

7.1.2 .Technology Assessment Review

The cost-effectiveness of DMTs in the management of MS have been the focus of significant attention for much of the last 10 years owing largely to the substantial annual cost of these therapies, which at the time of the assessment ranged from £6,500 to £12,500 per patient per year, combined with the length of time for which patients might continue to take these drugs. To date, independent and company sponsored evaluations of these therapies have produced a range of cost-effectiveness estimates from in excess of £1 million per QALY gained to cost saving. A review of current economic evidence highlighted significant flaws in the modelling of natural history, efficacy, discontinuation of therapy, mortality and the treatment of uncertainty.[3]

One of the key differences between existing independent and company-sponsored models related to the time horizon used.[3] In general, those models which produced very high cost-effectiveness estimates tended to have shorter time horizons (less than 10 years) or assumed that all benefit ceased when the patient stopped therapy. Conversely, models that assumed long time horizons and sustained benefit after the cessation of therapy produced economically attractive cost-effectiveness estimates.

In 2001, The National Institute for Clinical Excellence (NICE) commissioned a new model from a consortium of universities, using current best available evidence, to explicitly address the limitations of existing models. The ScHARR cost-effectiveness model estimated the cost-effectiveness of four products across six licensed indications, as shown in Table 7.1.

Table 7.1 Interventions evaluated by ScHARR MS model

Product name	Drug	Manufacturer	Dosage (units per week)	Licensed for RRMS?	Licensed for SPMS?
Avonex®	Interferon beta-1a	Biogen	6MIU	Yes	No
Betaferon®	Interferon beta-1b	Schering	8MIU	Yes	Yes
Rebif®	Interferon beta-1a	Serono	a) 22µg b) 44µg	Yes	No
Copaxone®	Glatiramer-acetate	TEVA/Aventis	20mg	Yes	No

7.1.3 Guidance issued

As a result of the scientific and non-scientific evidence made available to the NICE Appraisal Committee, on the basis of their clinical and cost-effectiveness neither beta interferon nor glatiramer acetate were recommended for the treatment of MS in the NHS in England and Wales.[4] The guidance recommended that all NHS patients already receiving these therapies should be given the option to continue treatment until they and their consultant consider it appropriate to stop, in accordance with the Association of British Neurologists Guidelines (ABN).[5]

7.1.4 Further research recommendations

The NICE guidance[4] recommended that NHS trusts and health authorities should be encouraged to collect data on all people with MS who continue on beta interferon or glatiramer acetate. It was recommended that such information should include the preparation used, the patient's relapse frequency and disease progression whilst receiving treatment, the development of adverse effects and neutralising antibodies, compliance with the therapy, the reasons for discontinuing therapy and the subsequent rate of progression of the disease.

Following the dissemination of this guidance,[4] the Department of Health entered into price negotiations with the manufacturers of these therapies. The result of these negotiations was the development of Risk Sharing Scheme (RSS), designed to monitor the cost-effectiveness of these DMTs in the management of MS.[6] The scheme involves the detailed monitoring of a cohort of patients to collect further data on the impact of DMTs on disease progression and the severity and frequency of relapses experienced. The therapies are thus available to all patients with relapsing/remitting multiple sclerosis (RRMS), and those with secondary progressive disease (SPMS) in which relapses are the dominant clinical feature, given their eligibility according to the ABN guidelines.[5] The monitoring process and associated price adjustments are expected to continue for 10 years.

In 2002, the National Co-ordinating Centre for Health Technology Assessment (NCCHTA) commissioned methodological work (HTA Project 02_29_01)[7] to develop methods for undertaking Expected Value of Perfect Information (EVPI) analysis for computationally expensive models. This work involved the development of an outline methodological framework for undertaking EPVI analysis for any health economic model and the direct application of the framework to an updated version of the ScHARR MS cost-effectiveness model.

7.2 Methods

7.2.1 Model structure

The ScHARR model uses the state transition methodology to simulate the natural history of MS over the Expanded Disability Status Scale (EDSS)[8] across RRMS and SPMS. The model evaluates the cost-effectiveness of the six licensed indications of beta interferon and glatiramer acetate using a 20-year time horizon and an annual cycle length. Patients progress through the model according to instantaneous hazard rates derived from a 25-year study undertaken in London, Ontario.[9] The transitions possible during any model cycle are shown in Figure 7.1. Patients who drop off therapy remain in the same EDSS state but transit to the conventional management quadrants of the matrix.

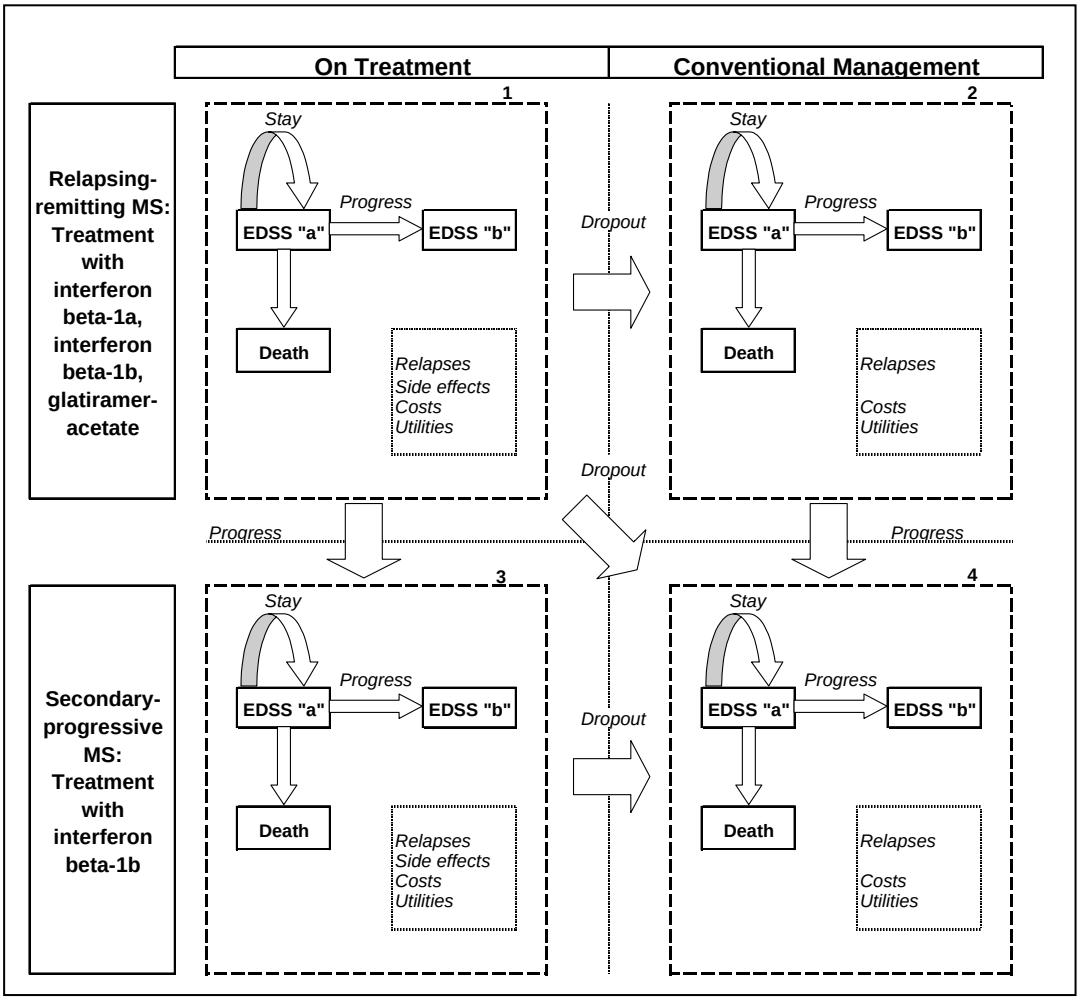


Figure 7.1 Progression diagram fro the ScHARR model [10]

Costs were estimated from the perspective of the NHS, and health outcomes were measured in terms of cost per quality adjusted life year (QALY) gained. Costs and health benefits are discounted at 6% and 1.5% respectively. Costs and utilities are applied directly to the state populations within each of the health states over each model cycle. During any particular model cycle, patients may also experience relapse, whereby a disutility is applied. For those patients receiving DMT therapy, a disutility is also applied to account for the experience of treatment-related side effects.

The effect of beta interferon and glatiramer acetate on disease progression and relapse were modelled using relative risks derived from published studies.[11-15] as well as a re-analysis of commercial-in-confidence trial data made available by two of the manufacturers of these therapies.

A list of parameters values, and their sources is available elsewhere.[10;16] All assumptions made in constructing the model favour the DMTs over conventional management. All transitions are assumed to be progressive only, hence patients cannot regress back to 'better' health states. A sustained effect of treatment on both progression and relapse beyond the trial duration was modelled. Thus, any patient who discontinues therapy subsequently progresses according to natural history rates but retains any benefits received at no additional cost of therapy. Thus, on the EDSS, these patients never 'catch up' with those patients who only receive conventional management. Due to the paucity of evidence concerning the long-term efficacy of any of these therapies, the effects of treatment are assumed to be fixed and did not deteriorate or increase over time. The annual relative risk of 'all-cause' mortality for the

MS cohort is assumed to be the same as a normal healthy population, minus the MS death observed in the natural history cohort. Patients started treatment according to ABN guidelines and are treated until they reach EDSS 7.0 or drop off therapy.

7.2.2 Probabilistic analysis

The model developed for the NICE appraisal was fully probabilistic and all uncertain model parameters were described using appropriate distributional forms. However, the cost and utility data used within the NICE assessment were held as commercial-in-confidence. Furthermore, the version of the model presented to the NICE Appraisal Committee did not recognise correlations between EDSS states for either utilities or costs (i.e. as costs increase, utilities decrease in a systematic pattern reflecting the change in the underlying clinical condition). For the subsequent NCCHTA-funded work,[7] these values were replaced with functions for costs and utilities to ensure that these correlations were reflected in the model appropriately. The specification of these functions drew from data reported in the literature,[17-20] our own experience of analysing cost and quality of life data in MS, as well as our knowledge of methodological issues around cost and quality of life assessment in chronic disabling conditions. The revised version of the model also uses adjusted prices for each of the DMTs, to reflect the current prices assumed within the ongoing Department of Health Risk Sharing Scheme.

7.2.3 Value of information analysis

Although the original assessment model included probabilistic analysis, value of information analysis was not undertaken as part of the original NICE appraisal. Owing to the computational expense of the MS model, this precluded a comprehensive EVPI analysis using the 2-level Monte Carlo algorithm, which would have required around 2,841 years of analysis time. Hence, further methodological work was commissioned by the NCCHTA in order to develop methods for undertaking sensitivity analysis (principally EVPI analysis) within computationally expensive health economic models. As part of this work, a methodological framework for undertaking EVPI analysis was developed and directly applied to the ScHARR MS model. Partial EVPIs were thus generated using 3 separate models:

- the original ScHARR MS cost-effectiveness model (using a 1-level approximation)
- a linear regression metamodel
- a Gaussian Process metamodel

The first two of these models assume a linear relationship exists between sampled parameter inputs and resulting net benefits for each treatment strategy. Conversely, the Gaussian Process methodology is a Bayesian non-parametric regression technique which is particularly effective in approximating highly non-linear models. Despite the considerable potential improvements in predictive accuracy possible using the Gaussian approximation, the method is restricted in terms of the number of model parameters that can be included in the model. Owing to limitations of the current MATLAB® software used to calibrate Gaussian Processes, only 30 model parameters could be specified within the metamodel. As the original ScHARR MS model contains 128 input parameters, Standardised Regression Coefficient analysis, Partial Contribution to Variance analysis, and Partial Rank Correlation Coefficient analysis were used to identify those model parameters that describe the greatest amount of uncertainty in the model.

An additional issue relating to undertaking value of information analysis within the MS model concerns the potential correlations between the efficacies of the six indications of beta interferon and glatiramer acetate. Whilst differences exist between the products being considered, there are also marked similarities; in these circumstances some level of correlation between treatment efficacies must be expected. In order to include efficacy correlation within the model it would be necessary to handle the set of six treatment efficacies as a multivariate normal distribution and to incorporate an uncertain covariance matrix into the model. Sampling using this multivariate distribution would then be facilitated by sequentially sampling a series of standardised normal distributions, and linearly transforming these samples using the Cholesky square root of the covariance matrix. This situation is further

complicated by the necessity to also sample the covariance matrix in order to capture the uncertainty in the correlations between treatments.

On a practical note however, there is a complete absence of quantitative information on the correlations between all treatments; the only option therefore would be to use subjective judgement in defining distributions for the correlation terms. Given these practical difficulties it was decided to take two approaches to the analysis of EVPI for the model:

a) To include all treatment options but assume independence in treatment efficacy. This approach therefore provides an upper estimate to the overall EVPI.
b) To consider a single drug treatment option, i.e. that with the highest incremental net benefit compared to conventional management. The results provided by this analysis will be equivalent to assuming a perfect correlation between treatment efficacies since the rank ordering of expected net benefits will be maintained. This analysis therefore provides a lower estimate for the overall EVPI.

7.3 Results

7.3.1 Adoption decisions

Tables 7.2 and 7.3 show the deterministic results generated using the version of the model used in the NICE appraisal and the revised 'public domain' version of the model used for the EVPI analysis respectively. Owing to the uncertainty surrounding the efficacy of individual therapies, these results consider only the marginal cost-effectiveness of each DMT compared to conventional management. Table 7.2 shows the cost-effectiveness results using public domain efficacy data as presented to the NICE Appraisal Committee. Under the base case assumptions, the cost per QALY ranges between £42,000 and £98,000

Table 7.2 Base-case results for original ScHARR model

Treatment Strategy	Per Patient Results		Marginal Results		Cost per QALY (£)
	Costs	QALYs	Costs	QALYs	
T1 Beta Interferon 1-a (Avonex, Biogen)	£111,954	10.20	£43,500	1.03	£42,041
T2 Beta Interferon 1-a 22mcg (Rebif, Serono)	£112,982	9.89	£44,529	0.73	£60,963
T3 Beta Interferon 1-a 44mcg (Rebif, Serono)	£130,949	10.03	£62,496	0.87	£71,732
T4 Beta Interferon 1-b 8MIU: Treating RR (Schering)	£101,726	9.83	£33,272	0.67	£49,664
T5 Glatiramer Acetate (Copaxone, TEVA)	£101,273	9.50	£32,820	0.34	£97,636
T6 Beta Interferon 1-b 8MIU: Treating RR & SP (Schering)	£107,022	10.03	£38,569	0.87	£44,390
T0 Conventional management	£68,453	9.16			

Table 7.3 shows the central estimates of cost-effectiveness generated using the updated version of the model. As discussed within Section 7.2, the difference in the results shown in Tables 7.2 and 7.3 concern the inclusion of revised cost and utility estimates for each EDSS health state together with the current adjusted prices for the Department of Health Risk Share Scheme. Table 7.3 shows that the revised cost per QALY estimates range between £39,000 and £92,000.

Table 7.3 Base case results using revised EDSS cost and utility estimates alongside RSS drug price

Treatment Strategy	Per Patient Results		Marginal Results		Cost per QALY (£)
	Costs	**QALYs**	Costs	**QALYs**	
T1 Beta Interferon 1-a (Avonex, Biogen)	£128,997	10.68	£39,874	1.02	£39,277
T2 Beta Interferon 1-a 22mcg (Rebif, Serono)	£121,966	10.36	£32,843	0.69	£47,318
T3 Beta Interferon 1-a 44mcg (Rebif, Serono)	£129,363	10.51	£40,240	0.84	£47,828
T4 Beta Interferon 1-b 8MIU: Treating RR (Schering)	£120,788	10.30	£31,665	0.63	£49,973
T5 Glatiramer Acetate (Copaxone, TEVA)	£114,956	9.94	£25,833	0.28	£92,279
T6 Beta Interferon 1-b 8MIU: Treating RR & SP (Schering)	£126,148	10.47	£37,026	0.80	£46,097
T0 Conventional management	£89,123	9.66			

Figure 7.2 presents cost-effectiveness acceptability curves (CEACs) and the cost-effectiveness frontier. The CEACs detail the probability that each strategy is cost-effective over a range of threshold values, and the frontier details the probability that the optimum strategy is cost effective. For values of lambda less than around £40,000, conventional management is the optimal strategy. For values of lambda higher than £40,000, Avonex becomes the optimal strategy. It should be noted here however, that these CEACs do not take account of efficacies between therapies and must therefore be interpreted with caution.

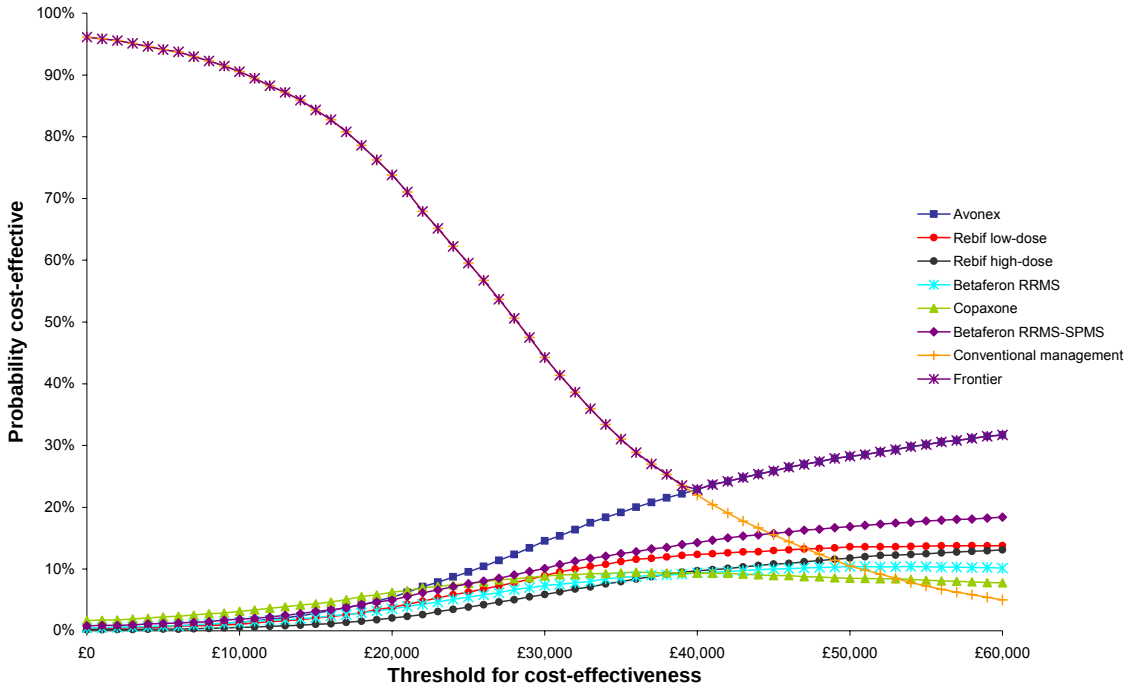


Figure 7.2 CEACs and Frontier for the ScHARR MS model

7.3.2 Research recommendations

Table 7.4 shows the epidemiological data, estimates of treatment eligibility and estimates of uptake from the Department of Health Risk Sharing Scheme; these estimates are used to estimate the relevant population for the EVPI analysis.

Table 7.4 Epidemiological and eligibility parameters used to estimate the relevant population

Variable	Value	Source
Prevalence	5,700	Estimated using data collected from the 64 centres participating in the MS RSS Monitoring Study
Incidence	3.8 per 100,000 population	Richards et al[2]
% RRMS at onset	80%	National Clinical Multiple Sclerosis Guidelines for NICE Management in Primary and Secondary Care
Percentage of patients eligible for DMTs	40-50%	Proportion of newly diagnosed RRMS who progress to DMTs *(Personal communication: Dr Mike Boggild, The Walton Centre for Neurology and Neurosurgery, 13/11/2003)*
Annual incidence of DMT use	800	Calculated from: Incidence x Population x %RRMS x Uptake
Discount rate	3.5%	NICE Guide to Methods of Technology Appraisal – Consultation document[21]

Table 7.5 shows the global 'per patient' and 'population' EVPI estimates for the two scenarios relating to assumptions concerning correlations between treatment efficacies: (1) EVPI estimates for each treatment strategy versus conventional management; and (2) decision EVPI across all 7 treatment strategies.

Table 7.5 Global EVPI results for the ScHARR MS model

	T1 Avonex versus T0 conventional management	T2 Rebif 22mcg versus T0 conventional management	T3 Rebif 44mcg versus T0 conventional management	T4 Betaferon RRMS versus T0 conventional management	T5 Copaxone versus T0 conventional management	T6 Betaferon RRMS/SPMS versus T0 conventional management	Decision EVPI
Per patient EVPI (£)	£4,271	£3,035	£2,827	£2,776	£2,444	£3,514	£8,855
Population EVPI (£)	£41,581,273	£29,545,388	£27,525,355	£27,023,241	£23,794,182	£34,211,482	£86,208,936

These estimates assume a £30,000 willingness to pay threshold and a lifetime for the decision of 10 years. Assuming independent treatment effects the per patient EVPI results shown in Table 7.5 suggests that the value of obtaining perfect information for all uncertain parameters within the case study model is £8,855. This results in a population EVPI of £86,208,936 across the relevant MS population, which represents an upper estimate for the EVPI. The population EVPI estimate for Avonex versus conventional management is £41,583,273; this assumes that treatment efficacies are perfectly correlated and hence represents a lower estimate for the EVPI.

Figure 7.3 shows the population EVPI for the decision over a range of willingness to pay thresholds (λ) over a lifetime for the decision of 5,10 and 15 years.

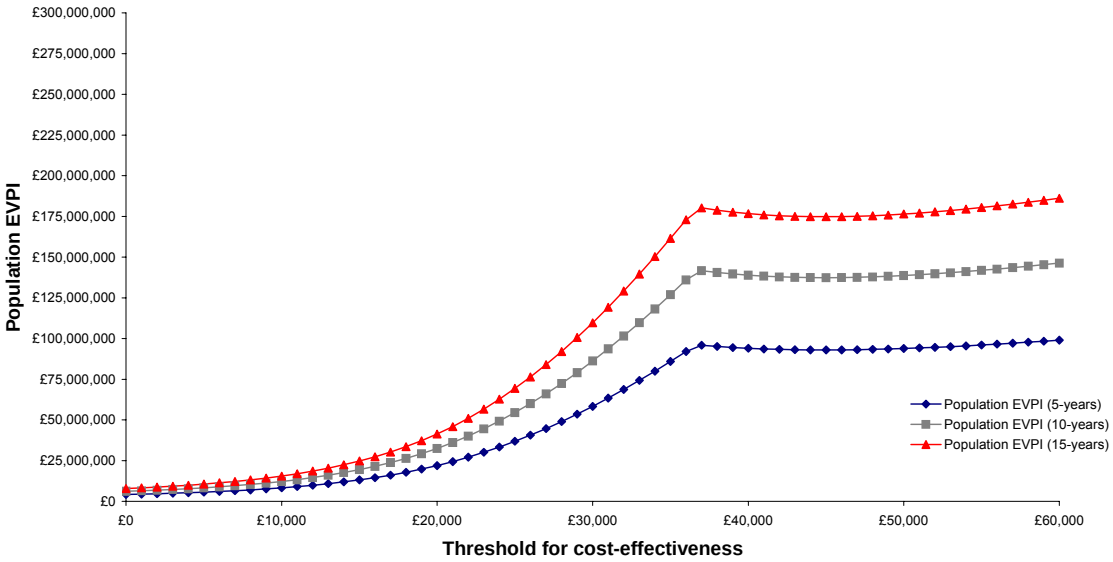


Figure 7.3 Population EVPI for the ScHARR MS model

It is evident from Figure 7.3 that as the value of λ increases, so too does the value of obtaining further research. It should be noted that the anticipated decision lifetime has a considerably greater impact upon the overall population EVPI as this value of λ increases.

7.3.3 EVPI for individual parameters

The key focus of the further methodological work undertaken on behalf of the NCCHTA[7] concerned the estimation of partial EVPIs for parameters. The partial EVPI results generated using all three approaches are reported in full in the monograph.[7] A direct comparison of the 1-level EVPI approximation, the linear regression metamodel and the Gaussian Process metamodel suggested that the Gaussian metamodel enabled a substantially better approximation than both the linear regression metamodel and the 1-level EVPI algorithm, despite the existence of a strong linear relationship between sampled parameter inputs and net benefits. Figure 7.4 shows the 'per patient' partial EVPI for individual parameters within the model as calculated using the Gaussian Process metamodel uplifted to the relevant population for a decision lifetime of 10 years. The partial EVPI analysis clearly suggests that large uncertainties surround many of the model parameters.

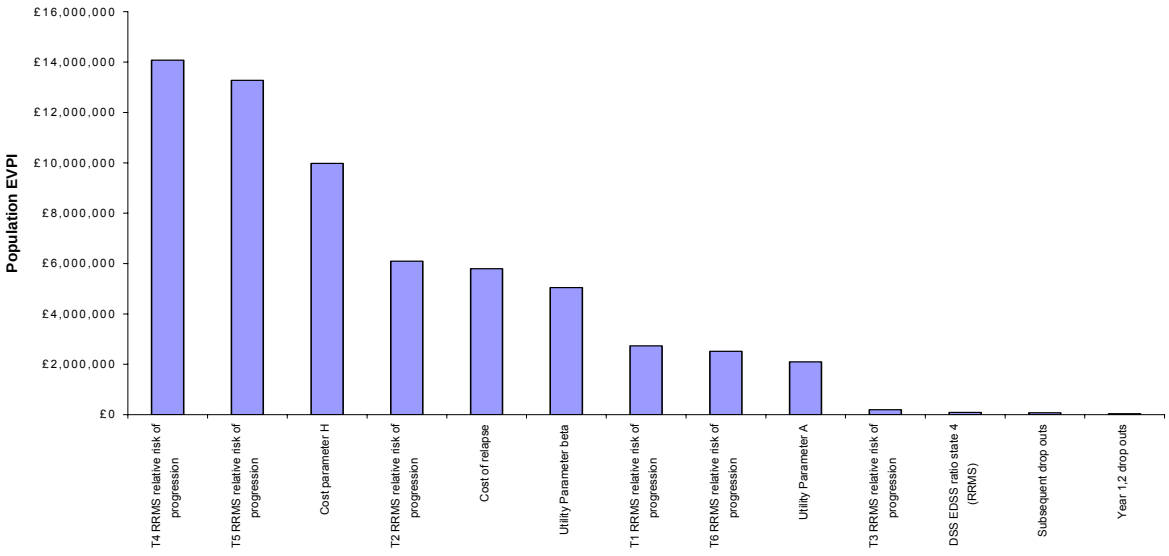


Figure 7.4 Partial EVPI for individual parameters for the ScHARR MS model

Figures 7.4 suggests that further research is merited on the relationship between the EDSS and the cost of care, the relationship between the EDSS and quality of life, the rate at which patients drop off therapy, and in particular, the long-term impact of DMT therapy on MS progression.

7.4 Discussion

The EVPI analysis presented here has identified a number of areas in which further research on the impact of beta interferon and glatiramer acetate is merited. For a £30,000 willingness to pay threshold, the global EVPI is in the range £41,581,273 to £86,208,936. Even at the lower estimate of EVPI, the potential value of obtaining further information is likely to considerably outweigh the costs of funding such research. The partial EVPI analysis suggests that research should focus specifically on the relationship between the EDSS and the cost of care, the relationship between the EDSS and quality of life, the rate at which patients drop off therapy, and in particular, the long-term impact of these therapies on disease progression. Whilst further information on costs associated with particular EDSS states and the rates at which patients drop off therapy may be obtained through non-experimental designs such as observational studies, further useful information on the impact of disease-modifying therapies on disease progression and associated health outcomes would be most reliably obtained through a long-term randomised controlled trial which includes a direct assessment of quality of life.

The large difference in the EVPIs obtained from the completely correlated and independent model assessments indicates that further knowledge on the correlation between treatment efficacies would be highly valuable in commissioning decision making. This is intuitively sensible as learning something about the treatment efficacy of one of the disease modifying therapies is likely to give information about other drugs in this set. Furthermore, the EVPI assessment incorporating all treatment options assumes that a single treatment is identified as optimal and selected for commissioning on the basis of its maximum net benefit. Given the high degree of uncertainty and the likely small differences in treatment efficacy and consequential net benefit an exclusive commissioning recommendation identifying one specific product is unlikely; a broad commissioning decision covering groups of or all products is more probable. Given this, the lower estimate for the expected value of information is likely to give a better estimate for the value of further research than the upper estimate.

Due to the vast computational expense of undertaking comprehensive partial EVPI analysis using the ScHARR MS model, estimates of partial EVPI were generated using three different approaches. The Gaussian Process metamodel suggested the closest approximation to the original simulation model, hence the partial EVPI estimates generated using this model are likely to be the most reliable and robust. However, the Gaussian methodology itself is restricted in terms of the number of parameters that may be specified in the metamodel (typically less than 30). Although the three separate factor screening techniques undertaken to identify the most influential model parameters identified the same 30 parameters to be important, there remains a chance that these importance analysis techniques failed to identify some of the important model parameters and these may have thus been excluded from the metamodel.

Furthermore, there exists an intrinsic problem in that due to the computational time required to perform comprehensive EVPI analysis using the correct 2-level EVPI algorithm within the original ScHARR MS model, there exists no basis for validation of the results of the Gaussian Process metamodel. Whilst the Gaussian Process methodology has previously been validated using a simpler case study,[22] this issue should be addressed in further policy problems.

An additional issue concerns the epidemiological and eligibility parameters used to estimate the population EVPI. The primary focus of this analysis has been on estimating the ‘per patient’ value of information, however the population scaling factors are also highly uncertain. In particular, the eligibility for and uptake of DMTs is particularly uncertain, and is currently based entirely upon subjective judgement. This should be highlighted as an area for further research. In addition, it should be noted that previous value of information studies have

calculated the population EVPI by simply multiplying the 'per patient' EVPI by a fixed number of patients over an assumed lifetime of the decision/technology. However, as the population relevant to a particular decision is itself uncertain, there remains an unresolved methodological issue concerning whether the uncertainty in the epidemiological parameters should also be accounted for within the sensitivity analysis.

It should also be noted that the model deviates from the criteria set out in the NICE reference case[21] on 2 counts (See Appendix ??). Firstly, whilst the model used within the NICE assessment used quality of life data valued using the EQ-5D, these data were held as commercial-in-confidence. The clear need for transparency in the application of EVPI methods to the MS model meant that these data were no longer available for use in the subsequent methodological work. As a result, a functional form describing the relationship between the EDSS and utility was specified from a review of the current literature and our own experience in analysing quality of life data in MS. In addition, the recommendation for discounting both costs and health outcomes at 3.5% was not agreed until after the EVPI methods work was completed; instead the discount rates recommended at the time of the assessment were used throughout.

Reference case evaluation: Disease-modifying therapies in the management of MS

Element of health technology assessment	Reference case	Criteria met by assessment?	Comment
Defining the decision problem	The scope developed by the Institute	Yes	
Comparator	Alternative therapies routinely used in the NHS	Yes	
Perspective on costs	NHS and PSS	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Synthesis of evidence on outcomes	Based on a systematic review	Yes	Systematic review undertaken previously on behalf of NICE[23] by Clegg and colleagues.
Measure of health benefits	Quality adjusted life years (QALYs)	Yes	
Description of health states for calculation of QALYs	Health states described using a standardised and validated generic instrument	No	The original NICE assessment included the results of a utility survey undertaken by the MS Research Trust. These data were held as commercial-in-confidence and were not used in the EVPI analysis. Functional relationships between quality of life and the EDSS were re-specified following a review of current literature and our experience in analysis utility data in MS.
Method of preference elicitation for health state valuation	Choice-based method, for example, time trade-off, standard gamble (not rating scale)	No	
Source of preference data	Representative sample of the public	No	
Discount rate	An annual rate of 3.5% on both costs and health effects	No	The 3.5% discount rate was agreed following the completion of the EVPI methods work; instead the discount rates for costs and health outcomes recommended at the time of the NICE assessment were used.
Equity position	An additional QALY has the same weight, regardless of the other characteristics of the individuals receiving the health benefit	Yes	

References

1. Compston A, Ebers G, Lassman H, McDonald I, Matthews B, Wekerle H. McAlpine's Multiple Sclerosis. London: Churchill Livingstone, 1998.
2. Richards RG, Beard SM, Sampson FC, Tappenden P. A review of the natural history and epidemiology of multiple sclerosis: implications for resource allocation and health economic models. *Health Technology Assessment.* 2002;6:1-73.
3. National Institute for Clinical Excellence. Minutes of the technology appraisals committee meeting 13th December 2000 http://www.nice.org.uk/article.asp?a=16091. 2003.
4. National Institute for Clinical Excellence. Full Guidance on the use of Beta Interferon and Glatiramer for the treatment of Multiple Sclerosis. Technology Appraisal No.32 Available from: http://www.nice.org.uk/Docref.asp?d=27634 [Accessed 08/04/2003]. 2002.
5. Chadwick D, Compston A, Miller D, Zajicek J, Dick J, Martyn C, et al. Guidelines for the use of Beta Interferons and Glatiramer Acetate in Multiple Sclerosis http://www.theabn.org/downloads/msdoc.pdf. Association of British Neurologists . 2001. 14-8-2003.
6. Department of Health. Cost effective provision of disease modifying therapies for people with Multiple Sclerosis. Health Service Circular HSC 2002/004. 4-2-2002. London.
7. Tappenden P, Chilcott JB, Eggington S, Oakley JE, McCabe C. Methods for expected value of information analysis for computationally expensive models. Developments on the health economics of beta interferon and glatiramer acetate for multiple sclerosis. *Health Technology Assessment (In press)* 2004.
8. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale. *Neurology* 1983;33:1444-52.
9. Ebers, G. London Ontario cohort study: Confidential data on file at National Institute for Clinical Excellence. London. 2001.
10. Chilcott JB, McCabe CM, Tappenden P, O'Hagan A, Cooper NJ, Claxton K *et al*. Modelling the cost effectiveness of interferon beta and glatiramer acetate in the management of multiple sclerosis . *British Medical Journal* 2003;326.
11. IFNB Multiple Sclerosis Study Group. Interferon beta-lb in the treatment of multiple sclerosis: final outcome of the randomized controlled trial. *Am Acad Neurol* 1995;45:1277-85.
12. PRISMS Study Group. Randomized double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. *Lancet* 1998;352:1498-504.
13. Jacobs LD, Cookfair DL, Rudick RR, Herndon RM, Richet JR, Salazir AM. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. *Am Neurol Assoc* 1996;39:285-94.
14. Johnson KP, Brooks BR, Cohen JA, Ford CC , Goldstein J, Lisak RP. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multi centre, double-blind, placebo-controlled trial. *Am Acad Neurol* 1995;45:1268-76.
15. European Study Group on interferon beta-1b in secondary progressive MS. Placebo-controlled multicentre randomised trial of interferon beta-1b in treatment of secondary progressive multiple sclerosis. *Lancet* 1998;352:1491-7.
16. Tappenden P, Chilcott JB, O'Hagan, A, McCabe C, Cooper NJ, Abrams KA, and et al. Cost-effectiveness of beta interferon and glatiramer acetate in the management of multiple sclerosis. Report to the National Institute for Clinical Excellence. 2001.
17. Grima DT, Torrance GW, Francis G. Cost and health related quality of life consequences of multiple sclerosis. *Multiple sclerosis* 2000;6:91-8.
18. Murphy N, Confavreux C, Haas J, et al. Economic evaluation of multiple sclerosis in the UK, Germany and France. *Pharmacoeconomics* 1998;13:607-22.
19. Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D. A cost-utility analysis of beta interferon for multiple sclerosis. *Health Technology Assessment.* 1998. 2;4.
20. Patwardhan, M., Matchar, D., Samsa, D., McCrory, D., Li, T., and Williams, R. Cost of multiple sclerosis by level of disability: A review of literature. Poster presented June 22nd 2003 at International Society for Technology Assessment in Health Care, Canmore, Alberta. 2003.
21. National Institute for Clinical Excellence. Guide to Methods of Technology Appraisal. 2003.
22. Oakley, J. E. Value of information for complex cost-effectiveness models. Research Report No. 533/02. 2002. Department of Probability and Statistics, University of Sheffield.

23. Clegg A, Bryant J, Milne R. Disease-modifying drugs for multiple sclerosis: A rapid and systematic review. *Health Technology Assessment* 2000;4.

8. Implications, methodological issues and challenges

8.1 Demonstration of benefits

The framework proved by DA-VOI was successfully implemented for each of the 6 case studies and provides the value of additional information, which may be generated by further research. This is consistent with the objectives of the heath care system (maximise the health gains of the population of England and Wales) and is based on the same resource constraints (the cost-effectiveness threshold which is used to develop guidance on use of the technology).

- For a particular assessment, this allows comparison of the potential benefits of further research (EVPI) with the costs of further investigation. If the potential benefits exceed the additional costs (including opportunity costs to patients) then further investigation maybe required to support guidance on use. The EVPI associated with the groups of parameters indicates the type of evidence that will be most valuable and therefore the type of studies that should be recommended.
- It also allows comparisons to be made across different technology assessments and prioritisation between alternative research recommendations, as well as a comparison between the value of investing resources in research or other activities such as the provision of health service. In this sense it provides a unified and coherent framework for prioritisation of both research and the use of heath care technologies

8.2 Implications for research prioritisation

The value of research differed substantially across the 6 technology appraisals and ranged from £2.8m (LBCs scenario 2) to £865m (ClO scenario 2). The results for selected scenarios for each of the case studies are reported in table 8.1.

Table 8.1: Summary of EVPI results

Case Study	Patient Group	Population EVPI	EVPI for parameters
AMD Screening	Visual acuity 20/40 Visual acuity 20/80	£6.2m £15.3m	Quality of life with and without PDT (£3,370,000 for 20/40)
Glycoprotein IIb/IIIa	Acute treatment following non-ST-elevation acute coronary syndrome (scenario 2)	£171m	Relative risk of death for non acute PCI for GPA as medical management and for Clopidogrel (£85,041,000, and £68,137,000 respectively)
Clopidogrel and dipyridamole for secondary prevention	Stroke Transient Ischaemic Attack Myocardial Infarction Peripheral Arterial Disease (scenario 2)	£865m £250m £710m £240m	Relative risks of vascular and non vascular death (£780m for ASA-MR-dipridamole compared to clopidogrel in the stroke subgroup)
Neurominidase inhibitors	Otherwise healthy adults not at elevated risk of complications	£66.7m	Quality of life with influenza, the effect of oselatimivir and amantadine (£44.3m, £0.43m and £0.23m respectively)
Liquid Based Cytology	Women aged 18 to 64 years (scenario 3)	£20m	Specificity (£3.6m)
Disease modifying therapies for multiple sclerosis	Relapsing remitting and primary progressive multiple sclerosis (scenario 2)	£86.2m	Relative risk of progression for copaxone, Betaferon and rebif (22mg) (£14m, £13.6m and £7m respectively) Also the cost of care, costs of relapse and quality of life (£10m, £7m and £6m respectively)

In some cases the analysis indicated that further research should not be regarded as a priority, e.g., the EVPI surrounding LBC following the evidence from the pilot study was low (£2.8m scenario 2). In other cases it indicated that additional research should be regarded as a priority, e.g., the EVPI surrounding CLO for stroke patients was high (£865). In other cases the analysis re-focuses the original research recommendations, e.g., in the AMD case study, although further research appears to be potentially worthwhile, it is additional evidence about quality of life with and without photo dynamic therapy rather than the performance of self screening itself which is valuable.

The analysis indicated which comparators should be included in future research and also suggested other parameters that could be excluded, For example, the value of information for NIs was significant (£66.7m) but it is further evidence about quality of life with influenza which is most important (£44.3) rather than additional evidence about the effect on symptoms. Although there is some value in further RCTs of the effect of oselativir and amantadine on symptoms (0.43m and 0.23m respectively) there is no value in further trials of Zanamivir. Similarly in the MS case study, although there is value in additional RCT evidence of the effect on progression of the disease, it is the effect of copaxone and betaferon which should be regarded as a priority (£14m and £13.6m respectively). However, in this case evidence about cost of care and relapse, and quality of life are also valuable. In these cases further research will not require experimental design and may be less costly to acquire (£10m and £6m respectively) so maybe regarded as priorities.

Estimates of value of information for the decision problem and for groups of parameters were also presented for relevant patient sub groups, e.g., for example the value of information different across the patient groups considered in the CLO case study (from £856m to £240m). This suggests that further research on the stroke and MI subgroups should be regarded as a priority although research on the TIA and PAD subgroups may also be worth while. Similarly the value of additional evidence for AMD differs by visual acuity (from £6.2m to £15.3m), and suggests that additional research should include those subgroups with lower starting visual acuity.

The analysis also indicates which endpoints should be included in further research. For example, in the GPA case study further research is valuable and should be regarded as a priority. It also indicated that it is RCT evidence of the effect of GPA as medical management and Clopidogrel which is most valuable. However, it also indicates that it is the mortality endpoint for patients with non-acute PCI which should be the primary endpoint in any future trial.

A number of case studies presented scenarios to explore alternative views of the relevant evidence e.g., inclusion of related and “unrelated events” in the assessment of CLO and impact of restricting consideration of evidence at 6 months in GPAs ; different structural assumptions regard mechanism of action, e.g., additive nature of information gains during screening for AMD; as well as the impact on value of information when relevant alternative may have been excluded from the original scope of the appraisal, e.g., the including the potential role of clopidigrel in the GPA case study.

A more detailed discussion of the implications for research prioritisation including the implications for the design of any future research in terms of features such as the relevant patient groups and comparators, and whether experimental design was likely to be required in each of the areas, can be found in chapters 2 to 7.

8.3 The reference case as a sufficient basis for VOI

The results of any analyses are conditional on the use of appropriate model structure, appropriate synthesis of evidence and characterisation of other uncertainties. This is important for estimates of expected cost-effectiveness but even more important for estimates of value of information which are particularly sensitive to these issues. The existing reference case and methodological guidance requires supplementary guidance on the detailed use of methods to ensure that the adequacy of reference case submission can be assessed and that an adequate reference case analysis will provide sufficient basis for value of information

analysis. This process of developing detailed methodological guidance is underway, coordinated by the National Decision Support Unit. It is important that in identifying and recommending methods the full characterisation of decision uncertainty should be a primary concern.

Although each case study was selected to be as close as possible to the new reference case there were a number of departures where the existing TAR model fell short of the new requirements. These are detailed at the end of each case study chapter. However, in general the most common and significant departures surrounded the quality of the evidence on quality of life available in the original TAR analysis. In addition and unsurprisingly the difference in discount rate between the original and new methods guidance was common to al case studies. It should be noted that even when the differences in discounting have a modest effect on estimated of cost-effectiveness then can have a substantial impact on value of information, e.g., in LBC case study changing the discount rate had a significant impact on he value of information.

8.3.1 Feasibility and resource implications

The pilot demonstrates that VOI is feasible within reasonable time lines, even based on pre reference case analysis. The use of VOI as part of the reference case (taking account of the recommendations made above) is for most types of models limited, not by time and resource requirements, but by the capacity to conduct this type of analysis and the dissemination of these methods. Therefore training in VOI methods should be considered as a cost-effective means of easing these capacity constraints. However it should be recognised that the key constraint is the capacity to conduct adequate probabilistic decision analytic modelling.

However, complex and computationally expensive models (particularly patient level simulations) make probabilistic analysis and therefore VOI potential very time and resource intensive. There are therefore 3 issues that should be addressed:

- It should be recognised that using patient level simulation is very costly in the sense that it may prevent reliable estimates of cost-and effect, and decision uncertainty as well as VOI being presented. In these circumstances it should be avoided if possible (by use of alternative structures and programming techniques). More work is required to establish those circumstances were the use of patient level simulation unavoidable.
- Where patient level simulation is required then there are techniques available to solve computationally expensive models, characterise uncertainty and estimate VOI. Indeed these have been used in NICE submissions. Further work is required to pilot their feasibility when patient level simulation is unavoidable and dissemination of appropriate methods.
- Non of the case studies included patient level simulation. However the MS and LBC case studies included computational expensive models and both used different techniques to overcome the computational problems. For example LBC attempted to estimate a linear relationship between model inputs and output, MS case study evaluated a number of approaches including a Gaussian process which does not impose linearity. Linear approximations maybe adequate for estimating costs and effect but my perform less will when estimating the value of information. The use of Gaussian process performs better than linear regression particularly when estimating value of information but the number of parameters, which can be included is limited. Again further works is required to pilot the use of techniques to evaluate computationally expensive models.

8.4 Critical issues and methodological challenges

It should be recognised that the key challenges for this type of analysis are not the VOI methods themselves but structuring decision problems, synthesis of evidence and the characterisation of uncertainty (required for estimating costs and effects as well as VOI). The development of methods in these areas is ongoing and will require continued support from a

variety of sources. Particular issues, many of which have been highlighted in the case studies, include:

8.4.1 Structuring decision problems

- Ensuring a sufficiently wide scope for the assessment to include all the relevant alternative strategies.
  This includes other technologies as well as different clinical policies (start and stop criteria) for a single technology. The exclusion of alternative strategies may not change the overall guidance on use of a technology but in some cases it may have a substantial impact on the value of information and on research recommendations. For example, excluding clopidogrel as an alternative in the GPA case study would have lead to an underestimate of the EVPI. The full range of clinical policies must also be included. For example, the CLO case study was restricted to evaluating 2 year treatment policies. However, if other policies such as life time treatment was evaluated then although the cost-effective strategy may not change the value of information may be higher and may focus on the longer term effect of secondary prevention
- Exploring and reflecting the additional uncertainty surrounding alternative but credible structural assumptions.
  For example, in the AMD case study three scenarios of alternative structural assumptions about how the information from self screening would change the chance of self referral to an ophthalmologist were modelled. Although the alternative assumptions had limited impact on estimates of cost-effectiveness (the overall cost-effectiveness of the strategies were unchanged) they did have a more substantial impact on the value of information (from £6.3m to £30.5m for visual acuity 20/40). This suggests that uncertainty and therefore evidence about the structural relationship may be as valuable as evidence about the value of the model parameters. In this pilot these types of uncertainty have been modelled as scenarios. However it is possible to assign probabilities (priors) to alternative assumptions and generate a value of information this for uncertainty. This would require elicitation of probabilities from experts and decision makers within an iterative process of analysis. Another example of these issues are found in the NI case study, e.g., whilst the base case value of information is driven by the uncertainty in quality of life, when this is modelled solely as a function of length of influenza illness, the value of information is reduced substantially.
- Model complexity and characterising uncertainty.
  The MS and LBC case studies included computational expensive models and both used different techniques to overcome the computational problems. For example, the LBC case study attempted to estimate a linear relationship between model inputs and output, the MS case study evaluated a number of approaches including a Gaussian process which does not impose linearity. Linear approximations maybe adequate for estimating costs and effect but my perform less will when estimating the value of information. The use of Gaussian process performs better than linear regression, particularly when estimating value of information, but the number of parameters which can be included is limited. Again further work is required to pilot the use of techniques to evaluate computationally expensive models.

8.4.2 Evidence synthesis and characterising uncertainty

- The synthesis of both direct and indirect evidence for measures of effect but also for other model parameters.
  This is a key issue for all the case studies. The GPA case study demonstrates that only considering the 6 month trial evidence (scenario 1) would overestimate the uncertainty and value of information. The more appropriate analysis (scenario 2), which included all the trial evidence, required more sophisticated methods of evidence synthesis. The CLO case study employed an indirect comparison of the two main treatments of interest, clopidogrel and modified-release dipyridamole, because no direct trial data were available. Such indirect comparisons are necessary for a full comparison of all treatment options, but are always subject to an increased level of

uncertainty. It is not surprising then that the relative treatment effect of clopidogrel compared to modified-release dipyridamole on mortality was associated with significant value of information.

- Dealing simultaneously with heterogeneity (variability by observed patient characteristics), variability (variability by unobserved characteristic) and uncertainty.
  In the GPA case study, if higher and lower risk patients could have been identified then sub group analysis could have been conducted. However, this was not possible and the analysis was conducted for average risk patients. It should be recognised that the current analysis is for patients with the average risk not the group of patients with variability in risk which make up the average. The latter is more policy relevant but requires both variability and uncertainty to be modelled.
- Reflecting the additional uncertainty due to potential biases in the evidence, which may come from different types of study and/or suffer from publication bias.
  The AMD case study used alternative structural assumptions to explore the substantial uncertainty and potential bias is estimates of self referral rates. This type of scenario analysis indicated that value of information was sensitive to these issues.
- Modelling the exchangeability of the evidence with the parameters required in the model and reflecting any additional uncertainty.
  For example, a meta regression was conducted to establish whether relative risk was related to base line risk in the GPA case study. This enabled the evidence from US trial (where baseline risks differed from the UK) to be used in the UK context. However the additional uncertainty introduced by using the US evidence was not explicitly modelled.
- The inclusion or exclusion of unrelated events from the evidence and the potential role of prior elicitation from “experts”.
  The CLO case study illustrated that the cost-effective treatment strategy and value of information can be significantly affected by the inclusion (exclusion) of unrelated (related) events. Where there is considerable uncertainty about how to include effects on events which may not be connected to the treatments under consideration, as with non-vascular mortality in the antiplatelets, trial. In such instances expert opinion on the validity of excluding the event from the model, or prior beliefs on the magnitude and extent of any potential treatment effects may have an important role in augmenting trial data.
- The potential role of using priors elicited from “experts” within the NICE process and appropriate methods of elicitation of priors.
  In the AMD case study expert judgements for the probabilities of self referral conditional on loses in visual acuity were used. The distributions assigned to these estimates were diffuse to represent the substantial uncertainty in the value of the parameters. However, more formal methods of elicitation for the whole prior distribution are available and would be more appropriate for future reference case analysis.
- Establishing efficient methods of searching for evidence on all model parameters not simply those associated with measures of effect.
  Each case study was based on a TAR which included a systematic review of the evidence on effect. However, the models and the probabilistic analysis rely on estimates of many parameters other than measures of effect. Clearly methods of systematic and efficient searching for all model parameters must be considered if all evidence is to be used to estimate costs and effects and fully characterise uncertainty.

Many of these issues are being addressed though various programmes of research around the UK. However these areas of research require further development and continued support. In additional it should be recognised that these issues require multidisciplinary working with collaboration across many different centres. Infrastructure support to facilitate full collaboration across these areas of work should be sort from a variety of sources. As all these methods evolve, it should be recognised that the detail of what is required within an adequate reference case analysis will also develop over time.

8.4.3 Issues specific to VOI

Although the key challenges are more general and relevant to estimating cost, effect and decision uncertainty there are a number of issues which are specific to VOI which need to be addressed

- Estimating the effective population that may benefit from additional evidence, including estimating time horizons for different technologies and incorporating this uncertainty in the estimates of value of information.
  For example, the CLO case study considered only modified-release dipyridamole, as it is now considered superior to the standard release formulation. Following the technological advance which allowed an extended release formulation, standard release dipyridamole, will likely fall out of use. In contrast, aspirin has been in use for more than the 15 years considered in the current EVPI calculations, and is still recommended for use in the guidance from the CLO appraisal despite the arrival of newer, competing technologies (clopidogrel and dipyridamole). Thus there is uncertainty around the effective future lifetime of these technologies and there is likely to be different effective life times for different technologies relevant to the same decision problem. In the NI case study the size of the population are based on the numbers of people currently presenting to the GP classified as influenza like illness. But the drugs need to be used within 48 hrs of symptom onset and are likely present similarly to a common cold. Therefore true "population" could be much greater than estimated once the technologies are available.
- Estimating the value of information for correlated parameters.
  The methods of evidence synthesis required to make comparisons between technologies generates correlation between estimates of effect. For example, in the CLO case study the relative risks were correlated and it should be recognised that the EVPI for these relative risks individually did not account for the correlation and could either under or over estimate EVPI. Similarly in the MS case study two scenarios were used to explore the impact of regarding each of the treatments as independent or as perfectly correlated (the same drug) and showed the effect on the estimates of value of information. Work is ongoing into accounting for correlations when calculating EVPI for parameters.
- Estimating the overall value of information based on estimates of the value of information for patient subgroups
  The AMD, GPA and CLO case studies included a number of patients sub groups. The EVPI for each patient group is useful in identifying were research will be most valuable. However it must be recognised that further evidence about one sub groups may inform other patient groups. Therefore the subgroup EVPI is likely to be a lower bound on the value of conducting research on that sub group alone. Similarly the summation across subgroups will overestimate the value of research for all patient groups together. What is required is to model the exchangeability between subgroups and indeed to other decision problems.
- Develop methods to assess the stability of estimates and required iterations. The number of iterations required to provide stable estimates of the EVPI and the EVPI for particular parameters depends on the nature of the model (non linearity) the number of parameters and their distributions. No clear general "rule" is available. Further work on measures of stability would be useful
- Presenting the value of information and the value of full implementation of guidance on use with in the same framework of analysis.
  The value of information conducted in this pilot focuses on the value of evidence about what is a cost-effective intervention. There is clearly a separate issue of the value of ensuring that clinical practice is consistent with the current evidence of cost-effectiveness and indeed that if additional research is commissioned that clinical practice will respond to the results.

Again work is currently ongoing on all of these issues but continued support from a number of sources for this methods work as well as support for an infrastructure of collaboration is needed.

8.4.4 Challenges in context

The challenges detailed above are not, with a few exceptions, specific to value of information analysis or indeed to decision modelling generaly. The issues of interpretation of evidence, synthesis, potential bias, exchangeability etc, have always been present in any informal and partial review of evidence. In fact until quite recently these challenging issues could be

conveniently ignored by both policy makers, clinicians and analysts while decision making was opaque and based on implicit criteria and unspecified “weighing” of the evidence. These challenges must be faced as more to explicit and transparent approaches to decision making are being taken. Indeed one of the many advantages of taking more transparent and explicit approach to decision making is that it exposes many important methodological issues which have previously been ignored or avoided by presenting partial analysis which do not directly address the decisions which must be made.

9. Implementing value of information analysis in the NICE research and development process

9.1 Introduction

Value of Information is not specified as part of the Reference case in the latest edition of the Appraisal Methods Guide, despite the fact that Probabilistic Sensitivity Analysis is.[xxxii5] The use of health economics and decisions analysis within the NICE Clinical Guidelines programmes is in its infancy, although the Institute has a strong commitment to addressing this issue in the medium term. Against this background, this chapter considers whether Value of Information Analysis could be incorporated into the standard processes of the Institute, and particularly how it could be used by the Research and Development Committee, (RDC).

The analyses presented in chapters 2 to 7 of this report demonstrate the feasibility of the *routine* incorporation of value of information analysis into analyses undertaken by Technology Assessment teams for the NICE Appraisal Programme. There are two important aspects of these demonstration projects that are worth highlighting at this point.

Firstly, these analyses were undertaken in a very short time scale. The results were available to the Institute approximately 5 weeks after the projects were approved. Thus, when probabilistic sensitivity analyses have been undertaken for the Technology Assessment Report (TAR), the additional burden of producing a Value of Information Analysis is not as large as might be expected, and is certainly small in relation to the value of the information obtained.

Secondly, the same individuals who had produced the TARs, and not a few specialists in VOI analysis, undertook these analyses. This is important, as it indicates that VOI analyses can be produced by any team that is capable of producing the probabilistic sensitivity analysis that is now required in the reference case analysis for NICE appraisals.

Based on the case studies presented above, the starting point for this chapter is that Value of Information Analysis is a pragmatically feasible output cost effectiveness models produced for the NICE Appraisal Programme.

9.1.1 The promise of Value of Information Analysis

VOI analysis is an internally coherent framework for assessing the value of investing in research to reduce decision uncertainty and identifying which research would be most valuable. Assuming a consistent decision criterion across all the decisions the Institute has to make, VOI can be used to identify the most valuable research questions to be addressed across the portfolio of the Institute's activities. VOI analyses can also be used to optimise the design of additional research. Given that resources for research, like all other resources, are limited, the prospect of prioritising research in a coherent and defensible manner is an appealing one.

9.1.2 Function of the NICE Research and Development Committee

The Strategy Document for the NICE RDC is clear that its primary role is not to fund research, but rather to promote research that it identifies as being important, to the established funding bodies such as the Medical Research Council, the Wellcome Trust and the NHS Research and Development Programme.[xxxiii]

The use of Value of Information Analysis to identify the research questions to be promoted to the research funding bodies, would imply that the objective of the RDC is to assist the Institute in its objective of maximising health gain from the expenditure of the NHS budget (subject to certain constraints), as VOI will only be available from analyses which have been

[5] Value of Information analysis is specifically referred to as a possible source of information on the research needs.

undertaken to inform the appraisal or guidelines programmes. The RDC may wish to consider whether it wishes to adopt such a narrow definition of its function.

It has been shown that the value of information is heavily dependent upon the comparators considered in the evaluation. Therefore, the question as it is defined in the Scope for an appraisal will play a crucial role in determining which research questions are identified as having a social high value. Experience of the NICE appraisal process leads us to believe that the scope of an appraisal is determined as many factors *inter alia*, process issues, resources available for the production of the assessment report and the availability of data. There is a risk that the research prioritisation based upon VOI analysis will not be robust to changes in the assumptions. This said, the introduction of the new appraisal process, notably, the stakeholder workshop; *should* lead to scope that are appropriate and therefore, robust value of information analyses.

It is unclear how research in to methods of health technology assessment and appraisal would be considered in a VOI framework. If the RDC were to use these methods to identify research priorities, it is unclear whether this would entail the exclusion of methods research from its remit. Given the Institute's pivotal role in the promotion and adoption of innovative research methods it would seem somewhat perverse for these activities be effectively excluded from the RDC's remit.[xxxiv] However, it is likely that such research would represent a very small proportion of the material that the RDC would be consider, and therefore its exclusion from the standard processes may be a reasonable price to pay for the application of robust analytical methods to the majority of the RDC's considerations.

The RDC has an equal responsibility to promote the research questions identified by all the programmes; not just those from the Appraisal Programme. Currently only the Appraisal Programme would be able to furnish VOI data in support of its research concerns. The work undertaken within the Guidelines, Interventional Procedures and Confidential Enquiries Programmes would not currently be in a position to do so. Either the RDC would need to develop a method for comparing VOI 'apples' with Confidential enquires' 'pears' etc… or the Institute would have to make a substantial investment in the provision of decision science expertise to these programmes. The feasibility of either of these strategies, in the short term requires careful consideration. This said, there is considerable external pressure for the incorporation of high quality decision science expertise in to the clinical guidelines processes. Therefore this potential problem may be resolved, to the benefit of the RDC processes.

9.1.3 Alternatives to EVPI for prioritising research

Given the issues highlighted in the previous sections of this chapter, it would seem sensible to briefly consider the alternatives to EVPI for prioritising research within a decision science framework.

Traditional research prioritisation processes have focussed upon the 'importance' of the question being addressed, the feasibility of undertaking research to address the question, and the quality of the methods of the proposed research study.

In clinical trials funding and health technology assessment more generally, the importance of the question has tended to be addressed in terms of the burden of the disease under consideration and the impact of the intervention upon that burden. Such information tends to be presented to the research funder as a portfolio. In order to make a decision, the funder must in some way attach a value to the burden of the disease, and to the change in that burden of disease. Such decisions are usually made by committees, and each member of the committee will their own assessment of the burden of the disease and the impact of the intervention. In such circumstances, simple voting mechanisms are used to establish a hierarchy of alternative research projects.

In essence, the research funder must undertake an informal and unstructured value of information analysis in their head, in deciding whether the commission a piece of research. The payoff function used by the research funder, is not necessarily have to relate to the decision of whether to reimburse a therapy or not, but it must attach a value, implicitly or

explicitly, the expected change in the state of knowledge that will result from the research. Each member of a research funding committee is likely to have their own pay off function, and the voting mechanism leads to the use of the mode as the estimate of the payoff, rather than the mean.

Thus, traditional research prioritisation mechanisms are merely informal, unstructured and opaque value of information processes. The parallels with previous arguments about Quality Adjusted Life Years and then Cost Effectiveness models in health care resource allocation decisions, are striking. The consideration must be whether research prioritisation processes informed by formal, structured and transparent value of information analyses are likely to lead to better research resource allocation decisions than informal, unstructured and opaque ones.

9.1.4 Acceptability of value of information analysis to research funding agencies

VOI is firmly based in statistical decision theory.[xxxv] The objective of VOI is to establish the value of requiring more information to support a decision. A fundamental precept of the decision theory framework is that traditional (frequentist) statistical inference is arbitrary and irrelevant decision making. By contrast the research funding organisations are *not* primarily concerned with supporting decisions and *are* firmly founded in the frequentist paradigm of statistics. There is clearly an issue about the acceptability to these bodies of conclusions drawn from an analytical framework with which they are at best unfamiliar and may actively reject.

The extent to which the results of VOI will be acceptable will vary from funder to funder. Whilst the NHS HTA programme is relatively familiar with the approach having funded a pilot study to run alongside a recent commissioning exercise,[xxxvi] the MRC and Wellcome Trust are less familiar. As the latter two bodies control by far the largest health care research funding streams, their lack of familiarity with the method represents a substantial challenge to its utility in the context of the RDC promoting research. However, the MRC is aware of the potential of these methods and has recently funded a 5 year programme grant to conduct methodological and applied research using value of information methods.

It seems likely that the Institute generally, and the RDC in particular would need to engage with the research funding bodies to promote understanding of the concepts underlying the framework, the techniques for implementing these concepts and the value of the results from such analysis. Such activities will take time, in the interim, VOI is unlikely to be effective as a means of promoting specific research questions to the research funding bodies. This said, the provision of VOI as supporting information to the Institute's arguments for undertaking specific research projects may act as an effective tool for engaging the funding bodies in discussions about the method. It may be that the provision of a quantitative assessment of the value of research identified by the Institute will challenge the funding bodies demonstrate, in some manner, that the research they propose to support has more value to society. This in itself will represent a valuable contribution to research investment decisions in the UK.

9.1.5 Implementing value of information analysis within NICE

Value of Information Analysis requires the construction of high quality cost effectiveness models, with appropriate probabilistic sensitivity analysis. It is not clear that the cost effectiveness analyses currently provided to the NICE Appraisal programme. In part this is due to a limited supply of the necessary skills and in part it reflects the view of some analysts that probabilistic sensitivity analysis is not necessary to support decision making.

In addition to the logistical problems posed by Assessment teams and collaborating centres that are currently unable or unwilling to provide PSA and VOI analyses to the Institute, there are substantial technical difficulties in the application of PSA which need to be addressed.

All existing PSA analyses, including those reported here, are likely to systematically under estimate the degree of decision uncertainty, and therefore the value of investing in additional research. This is because the data used to parameterise cost effectiveness analyses are approximations to the parameter that the decision maker is actually interested in, rather than

the parameter itself. The most widely known example of this problem is the use of efficacy data from clinical trials when the decision maker is interested in effectiveness. However, what O'Hagan describes as the Data Gap, [xxxvii] applies to all parameters in cost effectiveness analyses. For example, we use uninformed general population health state values when we want informed general population values; [xxxiixxxviii] we use trial based resource use when we are interested in service resource use;(example ref required) we use trial based compliance rates when we are interested in compliance rates seen in the treated population;[xxxiv] and so on… With each parameter we use an approximation, we increase the total uncertainty about how accurate an estimate of the cost effectiveness of the intervention the model provides. An extension of this observation is that if research undertaken in response to VOI analyses gathers data on the approximation to the parameter of interest, rather than the parameter itself, (e.g. if a double blind placebo controlled clinical trial is undertaken to improve the evidence on effectiveness) the value of the information obtained will be less than the standard VOI analysis would suggest (A. O'Hagan, Personal communication May 2004).[6]

Additional technical problems exist in the application of PSA and VOI more generally. The methods of evidence synthesis are in their infancy and there are many types of data which we are unsure how to synthesise; e.g. observational data and trial data. Whilst global and partial VOI calculations are relatively straightforward, the methods for estimating the value of sample information are still under development.

A further challenge to the implementation of VOI is the requirement for a single well defined decision criterion, expressed in terms of a willingness to pay for a unit of health gain. It is quite clear from the most recent Methods Guidance that the Appraisal programme does not have a single fixed decision criterion. [xxxii] With a decision criterion that varies between appraisals, the VOI analyses will not be comparable across appraisals. For decisions on the Value of further research to inform that specific decision, this is may not be important, however, the RDC will be deliberating over the whole range NICE activities and comparability is an important aspect of any analytical tool the RDC adopts. The other programmes within the Institute do not use an explicit cost effectiveness threshold as the decision criterion, and therefore it is not possible to estimate the value of reducing the uncertainty around their decisions.

9.1.6 Summary

In this chapter we have attempted to consider the feasibility of the RDC adopting VOI as the primary source of data to inform its deliberations. Whilst the case studies reported above demonstrate that it is possible to produce Value of Information analyses. Indeed, where the model has already been constructed and probabilistic sensitivity analyses undertaken, the marginal cost of producing global and partial value of information estimates is far from prohibitive.

This said, we have identified a number of challenges to the routine use of VOI by the RDC. Some of these issues relate to process, whilst others are technical.

9.1.7 Process Issues in the adoption of VOI by the RDC

The case studies reported above are all taken from the NICE Appraisal Programme, and were undertaken by TAR teams who have established expertise in PSA and VOI. Whilst the new methods mandate the use of PSA in the reference case analysis for all appraisals, many of the TAR teams are currently ill-equipped to undertake such analyses and it will take a substantial period of time before high quality PSA is routinely provided in all the TARs. Even

[6] It may be that research on the nature of the relationship between the approximation and the parameter of interest will be of more value than research on the true value of the approximation. Analyses which addressed this data gap would require a major change in the way that cost effectiveness models are built, with the explicit incorporation of structures linking the approximation of the parameter of interest, to the real parameter. Current VOI analyses assume that that parameter of interest and the approximation are identical.

when this is the case, VOI will not be routinely provided. If the RDC wishes to receive VOI, it is likely that it will have to find resources to support the additional work.

The RDC has defined is function as the promotion of research priorities to the conventional research funding bodies. It will be necessary to convince these bodies of the validity of the VOI framework before promotion of research priorities based on such analyses will be acceptable. This is likely to take a substantial amount of time and effort.

The RDC has responsibilities to promote the research priorities identified by all programmes within the Institute, not just the research priorities identified through the appraisal programme. None of the other programmes routinely undertaken substantial cost effectiveness analysis, never mind probabilistic sensitivity analysis and VOI. Either these types of analyses need to be implemented across all the Institutes programmes or the RDC will have to find a means of comparing the research prioritisation outputs of the different programmes, when they are identified and expressed using incompatible processes.

9.1.8 Technical issues in the adoption of VOI by the RDC

There are three substantial technical issues in the adoption of VOI by the RDC. The first relates to the technical state of knowledge on probabilistic sensitivity analysis. PSA requires the synthesis of evidence from a variety of sources in order to identify the expected value and the uncertainty around that expectation. For many types of evidence, analysts know how to undertake such synthesis. However, there are some types of evidence for which we do not have adequate techniques. Thus VOI will not always be possible. In part this is using perfection as an argument against the merely good. However, it is important to recognise that PSA and VOI methods are 'works in progress'.

The second technical challenge to the use of VOI is what O'Hagan describes as the 'Data Gap'. The uncertainty expressed in current probabilistic sensitivity analyses is always an under assessment of the true uncertainty, because the data used are from approximations to the parameters of interest to the decision maker, not the parameters of interest. This under assessment of uncertainty will lead to a mis-specification of the value of information and to the degree that further research actually collects information on the approximation rather than the parameter, the value of that information will be less than the analysis indicated. Addressing the problem of the data gap will require a fundamental recasting of cost effectiveness models, such that the model structure includes a function relating the approximation to the parameter of interest to the decision maker. The incorporation of uncertainty around this relationship will allow the analyst to establish the value of further research on the approximation to the parameter compared to research on the relationship between the approximation and the parameter.

VOI requires the specification of a single decision criterion; lambda. This single criterion does not exist even with the appraisal programme, never mind the other programmes of the institute. The adoption of VOI by the RDC would imply substantial changes in the way in which the other programmes within the Institute undertake their analyses, and push the Appraisal Committee towards a much more uniform cost effectiveness decision criterion.

Whilst there are many challenges associated with the use of Value of Information Analysis by the RDC, the alternatives are difficult to defend as they are in essence; informal, unstructured and opaque value of information analyses, which do not have a specified payoff function. Against this background, the RDC has the opportunity to challenge current research resource allocation decisions to improve by incorporating Value of Information Analysis into its processes for the assessment of future research needs.

References:

1. National Institute for Clinical Excellence NICE Appraisal Programme Methods Guide NICE London May 2004
2. NICE National Institute for Clinical Excellence Research and Development Committee: Strategy Document NICE 2003 London
3. Chilcott J, McCabe C, Tappenden P, O'Hagan A, Cooper NJ, Abrams K, Claxton K, Modelling the cost effectiveness of interferon beta and glatiramer acetate in the management of multiple sclerosis BMJ 2003; 326: 522-526
4. Pratt, J., Raiffa, H. and Schlaifer, R. 1995. Statistical decision theory. Cambridge MA: MIT Press.Claxton K Sculpher M et al A pilot study of Valuation of Information Analysis for Commissioning Research NCCHTA 2004 (forthcoming)
6. O'Hagan A. Mind the Gap in Health Technology Assessment Presentation to CHEBS Probabilistic Sensitivity Analysis Course Part 1 April 2004 Sheffield
7. Brazier JE et al Whose Health is it anyway? HEDS DP 04/4

10. Conclusions

The framework proved by DA-VOI was successfully implemented for each of the 6 case studies and provides the value of additional information, which may be generated by further research. This value is consistent with the objectives of the heath care system (maximise the health gains of the population of England and Wales) and is based on the same resource constraints i.e., the cost-effectiveness threshold which is used to develop guidance on use of the technology.

For a particular assessment, this allows comparison of the potential benefits of further research (EVPI) with the costs of further investigation. If the potential benefits exceed the additional costs (including opportunity costs to patients) then further investigation maybe required to support guidance on use. The EVPI associated with the groups of parameters indicates the type of evidence that will be most valuable and therefore the type of studies that should be recommended.

It also allows comparisons to be made across different technology assessments and prioritisation between alternative research recommendations, as well as a comparison between the value of investing resources in research or other activities such as the provision of health service. In this sense it provides a unified and coherent framework for prioritisation of both research and the use of heath care technologies

The results of any analyses are conditional on the use of appropriate model structure, appropriate synthesis of evidence and characterisation of other uncertainties. This is important for estimates of expected cost-effectiveness but even more important for estimates of value of information which are particularly sensitive to these issues. The existing reference case and methodological guidance requires supplementary guidance on the detailed use of methods to ensure that the adequacy of reference case submission can be assessed and that an adequate reference case analysis will provide sufficient basis for value of information analysis.

It should be recognised that the key challenges for this type of analysis are not the VOI methods themselves but structuring decision problems, synthesis of evidence and the characterisation of uncertainty (required for estimating costs and effects as well as VOI). The development of methods in these areas is ongoing and will require continued support from a variety of sources.

The pilot demonstrates that VOI is feasible within reasonable time lines, even based on pre reference case analysis. The use of VOI as part of the reference case is, for most types of models limited, not by time and resource requirements, but by the capacity to conduct this type of analysis and the dissemination of these methods. Therefore training in VOI methods should be considered as a cost-effective means of easing these capacity constraints.

There is a range of possible options to implement DA-VOI within the NICE process, to inform research recommendations. In general there a two levels at which DA-VOI could be implemented. DA-VOI could be implemented at the TAR stage of the process, either selectively or ultimately becoming part of the reference case for the Assessment Report. Alternatively DA-VOI could be implemented in a similar way to the case studies presented here: as a supplementary analysis to and existing TAR once guidance on use and research recommendations has been made. This maybe regarded as the most feasible way of progressing in the short run. But, as capacity and methods develop, a move towards making DA-VOI a routine part of the TAR and the research recommendations in the guidance more firmly grounded on evidence and an explicit analysis may be achievable in the medium term.