The Value of Clinical Information: An Economic Approach to Research Priority Setting

Karl Philip Claxton

D.Phil Submission

University of York
Department of Economics and Related Studies

November 1996

Abstract

This thesis considers two aspects of the value of clinical information: the value of information provided by diagnostic technology; and the value of information generated by clinical research. It is motivated by the methodological problems which are encountered when faced with the economic evaluation of sequential clinical decision problems. A strategy for the evaluation of diagnostic information which could avoid the need for randomised clinical trials is examined. This is generalised to more complex clinical decision problems. It is shown that this strategy will fail in most clinical settings and prospective research will be required. However it is argued that the traditional approach to the design of clincial research is inconsistent with concepts of efficiency even when an economic evaluation is conducted alongside a clinical trial. This poses the problem of how to establish allocative and technical efficiency in clinical research. These issues are addressed by developing decision-analytic and dynamic programming approaches to clinical trial design and research priority- setting. Two hurdles are proposed for clinical research. The first ensures that only potentially cost-effective research is considered. The second ensures that this research will be cost-effective when conducted at the technically efficient scale. The dynamic programming approach enables relevant alternatives which should be compared in a clinical trial to be identified consistently and explicitly. The approach provides a measure of the net benefit of proposed research which can be used to establish allocative efficiency in research and development across clinical decision problems or broader areas of clinical research. Perhaps most importantly, it can be used to establish the optimal allocation of resources between research and development and service provision. Indeed what is clear from this approach is that the value of information and research priorities cannot be separated from the budgetary constraints on service provision.

List of Contents

List of Tables and Figures

Acknowledgement

I would like to acknowledge the support and guidance of my supervisor Dr John Posnett and the members of my thesis advisory group: Professor Peter Smith; and Professor Alan Williams.

Chapter 1 Introduction

Contents: Page

1-1 Introduction

This thesis considers two aspects of the value of clinical information: the value of information provided by diagnostic technology; and the value of information generated by clinical research. It is motivated by the methodological problems which are encountered when faced with the economic evaluation of sequential clinical decisions which include one or more diagnostic processes and a number of treatment strategies. The issues posed by this type of decision problem are: (a) can diagnostic information be valued without the prospective evaluation of all feasible strategies of patient management; (b) if not, is it worth collecting additional information about this decision problem through prospective research; (c) if it is, what is the optimal scale of this research; and (d) which of the many competing strategies of patient management should be included (regarded as relevant alternatives) in the evaluation? The thesis is an attempt to address these practical problems which are the issues of allocative and technical efficiency in research and development.

The thesis presents an approach which can address each of these problems. The methods developed are illustrated in each chapter using simple numerical examples which are introduced in chapters two and three. In chapter two a strategy for the evaluation of diagnostic information which could in principle avoid prospective clinical research with randomised patient selection is examined. In chapter three this approach is generalised to a more complex decision problem. It is shown that this strategy will only provide consistent valuations if critical assumptions, which are unlikely to be met in most clinical settings, hold. This poses a number of problems which include the issues of allocative and technical efficiency in clinical research. These issues are addressed in chapters four, five and six by developing decision-analytic and dynamic programming approaches to clinical trial design and research priority-setting.

Methodological and Policy Issues

In the process of addressing these problems some interesting methodological issues are highlighted. One of the implications from chapter three is that there may be problems when using league tables of cost-effectiveness ratios to set priorities in service provision. Chapter four demonstrates that the traditional approach to clinical trial design is inconsistent with concepts of efficiency even when an economic evaluation is conducted alongside a clinical trial. The decision-analytic approach to the value of information shows that there are circumstances when it will not be efficient to conduct a clinical trial and clinical practice should be based only on prior information. Establishing the expected net benefit of research also means that ethical judgements about proposed research can be based on consistent estimates of the opportunity cost of particular ethical concerns.

The thesis also provides tools which can address some interesting policy questions, in particular methods for research priority-setting. In chapters four, five and six two hurdles are proposed for clinical research. The first ensures that only potentially cost-effective research is considered. The second ensures that this research will be cost-effective when conducted at the technically efficient scale. The expected net benefits of research can be used to establish allocative efficiency in research and development across clinical decision problems or broader areas of clinical research. Perhaps most importantly, it can be used to establish the optimal allocation of resources between research and development and service provision. Indeed what is clear from the analysis is that the value of information and research priorities cannot be separated from the budgetary constraints on service provision.

1.2 The Value of Diagnostic Information

The clinical approach to the evaluation of diagnostic information using measures of accuracy is inadequate because it is not founded on the proposition that information is only valuable insofar as it changes subsequent decision-making. The use of performance measures for diagnostic technology which are independent of the consequences of subsequent treatment decisions will not reflect the most important impact of diagnostic information: the consequences of subsequent changes in patient management.

1.2.1 Clinical Measures of the Value of Diagnostic Information

The simple measures of accuracy of a diagnostic device suffer from a number of problems. Measures of the sensitivity and specificity of a test do not directly address the issue of concern for the clinician; namely the probability of disease for a given test result. Predictive values, although more intuitively appealing, depend on the prior probability of disease (via Bayes) and are context and population specific. Both these types of measures of accuracy are not independent of the positivity criterion (the cut-off used to operate the diagnostic test) and are to some extent arbitrary. The optimal positivity criterion can only be established by considering the net consequences of classifying a test result as positive and the net consequence of classifying a result as negative [137]. This requires information not only on the accuracy of the test but also on the expected costs and health outcomes of subsequent treatment strategies.

More sophisticated measures of accuracy which take account of the possible trade-off between sensitivity and specificity claim to overcome this problem and provide measures of accuracy which are independent of the positivity criterion. The area under the Receiver Operating Characteristic (ROC) curve [2, 24, 67, 97, 130] is a measure of accuracy which is independent of the positivity criterion, because it is a

measure of accuracy across all possible combinations of sensitivity and specificity that the test can provide [92]. The area criterion is really an average of true positive rates over the full range of possible false positive rates [107]. The use of the whole area under the ROC curve implies that the false positive and false negative results are equally valued. However only a small range of false positive rates will be clinically relevant and there will be a large range of false positive rates which will never be considered. The relevant range of false positive rates depends on the patient population with a particular prior probability of disease and also depends on the subsequent treatment strategies which are available. This is a particular problem if the ROC curves of alternative diagnostic test cross [72, 106]. In these circumstances the analyst needs to know where on the ROC curve each test should be operated. But to establish this requires information about the expected costs and health outcomes of the subsequent treatment decisions. Finally it has been argued that the interpretation of the area under the ROC curve is not directly relevant to the decision problem the clinician faces and the energy expended on estimation techniques [68] for ROC curves suggests the imprudent use of a potentially misleading and inconsistent performance measure which lacks relevance to clinical practice [72].

The conclusion that unambiguous and reliable measures require an assessment of the impact of information on the outcome of subsequent treatment choices and not simply diagnostic accuracy seems unavoidable. The economic evaluation of diagnostic information requires information about the accuracy of the test but also the expected costs and health outcomes of the subsequent treatment alternatives [46, 66, 150]. A number of measures have been proposed based on treatment thresholds [103, 104, 105]. In the absence of a diagnostic test a treatment threshold can be estimated based on knowledge of the prior probability of disease and the health outcomes with and without treatment. This threshold indicates the prior probability of disease where the clinician should be indifferent between treatment and no treatment. Thresholds for a new diagnostic test can be established based on an assessment of the accuracy and the existing information about current

practice embodied in the treatment threshold. The testing threshold, the point at which the clinician switches from no treatment to testing, and the test/treatment threshold, the point at which the clinician switches form testing to treatment, defines the range of prior probability of disease where the diagnostic test should be used. Clearly a more accurate diagnostic test will generate a greater range where the test should be used.

There have been a number of approaches to the use of thresholds. Patient orientated performance measures for a diagnostic test have been proposed based on treatment thresholds [64]. These attempt to measure the impact of diagnostic information on patient management: assignment potential [63] is the probability that the diagnostic test result will change clinical practice by moving the post test probability of disease across the treatment threshold; assignment strength [23] measures the distance from the threshold following the results of the test. Other developments have generated stochastic thresholds [69, 70, 98]. However all these approaches share the assumption that current practice is the appropriate baseline against which to compare a new diagnostic technology. It implies that a new diagnostic technology can simply be added to the existing strategies of patient management. These approaches assume that new information has an impact only on whether the patient is assigned to the treatment which was used before the new test was available. This seems an appropriate assumption when considering examples where only one treatment option is available for a given diagnosis, but may not be appropriate when considering more complex sequential clinical decision problems.

1.2.2 A Strategy for the Economic Evaluation of Diagnostic Information

A strategy for the economic evaluation of diagnostic information has been proposed by Phelps and Mushlin [107]. In chapter two this strategy is applied to a simple numerical example. By using information on the costs and outcomes of

current practice prior to the introduction of a new diagnostic technology (assuming it is available) the value of perfect information can be established based only on this prior information. This is the maximum value that any diagnostic technology can provide. By comparing this to an estimate of the cost of the new technology it can be used as a first hurdle that must be overcome. If a proposed diagnostic technology passes the first hurdle then it is potentially cost-effective and prospective research is needed to establish the accuracy of the test. If the new diagnostic technology is non-invasive then random patient selection is unnecessary and a clinical trial of the full diagnostic and treatment process can be avoided

Once the accuracy of the test is established the prior information about current practice is used to estimate the expected value of clinical information. If this exceeds the estimated cost of the test then the new technology is cost-effective and passes the second hurdle. The value of diagnostic information can be established without recourse to a randomised clinical trial of the diagnostic and treatment process, and clinical research can be focused on those diagnostic technologies which are potentially cost-effective.

1.2.3 Consistency in the Economic Evaluation of Diagnostic Information

This strategy assumes that current practice is the appropriate baseline to evaluate a new diagnostic technology. This requires current practice to be cost-effective at the critical cost-effectiveness ratio which will be used to set priorities in service provision when the new technology has been evaluated. In chapter three the Phelps Mushlin strategy is generalised to the more realistic situation where there is more than one treatment option available following diagnosis. Consistency requires that the value of health outcome which is implicit in current practice is the same as the critical ratio which will be used to decide if the new technology will be cost-effective. The analysis in chapter three shows that if this assumption does not hold (because clinicians have a higher implicit value of health outcome or do

not perceive all the costs of the alternative patient management strategies) then current practice may not be the appropriate baseline. The analysis can be subject to two types of error at each of the two hurdles: the value of clinical information may be overestimated and a diagnostic technology which is not cost-effective is evaluated and may be implemented; or the value of clinical information is underestimated and a potentially cost-effective diagnostic technology may be rejected.

The analysis in chapter three shows that the Phelps Mushlin strategy like the threshold approach, assumes that a new diagnostic technology can simply be added to the existing strategies of patient management. However when there is more than one possible treatment for a given diagnosis it is possible that the optimal treatment following diagnosis is not part of existing strategies of patient management. In these circumstances information about current practice will be inadequate to establish the value of diagnostic information. Evaluation will require random patient selection in a trial which includes the diagnostic and treatment processes even when the new test is non-invasive. This poses a number of problems which are addressed in the following chapters of the thesis.

1.3 Allocative and Technical Efficiency in Clinical Research

The analysis of the Phelps Mushlin strategy in chapter three suggests that it is likely to fail when applied to more complex decision problems because: the key assumption of consistency between the value of health outcome implicit in current practice and the critical cost-effectiveness ratio used by an analyst is unlikely to hold; when this assumption is violated the values of both the first and second hurdles will be biased; and it is not necessarily the case that information available about current practice before the test is introduced will be sufficient to establish the value of diagnostic information. The Phelps Mushlin strategy may fail, but if the prospective evaluation of all possible alternatives in a sequential clinical decision problem is not possible then this poses a number of questions: (a) how should information of different quality from different sources be combined consistently and explicitly; (b) which clinical decision problems will be worth evaluating in a clinical trial; (c) if a clinical decision problem is worth evaluating which of the competing alternatives should be compared in a clinical trial; and (d) what is the optimal scale of this prospective research? These are the questions of how to establish both technical efficiency in research design, and how to achieve allocative efficiency in research and development across clinical decision problems and between research and service provision. It is these questions which are addressed in chapters four, five and six.

1.3.1 A Decision Analytic Approach to Clinical Trial Design

The analysis in chapter four uses a decision-analytic approach to the value of information which combines a Bayesian view of probability with a framework for decision making. This approach is used to establish the cost of uncertainty surrounding the clinical decision problem (the expected value of perfect information). The marginal benefit and the marginal cost of acquiring sample information is then explicitly considered. This enables the technically efficient

scale of the research to be identified and the expected net benefit of proposed research to be established. These methods are generalised from the single-stage decision problem considered in chapter four to the two- and four-stage sequential clinical decisions in chapter five. However, although the decision-analytic approach taken in these chapters addresses the issues of which clinical decision problems are worth evaluating and what the technically efficient scale of the research should be, it does not allow the relevant alternatives to be identified.

1.3.2 A Dynamic Programming Approach to Optimal Patient Allocation

In chapter six the fixed and equal allocations rule used in chapters four and five, which assigns equal numbers of trial entrants to each of the alternative arms of the trial, is relaxed. A dynamic programming approach is used to identify the optimal allocation of trial entrants at each stage of the decision problem. By explicitly considering the marginal benefit and marginal cost of assigning trial entrants to the alternative arms of the trial the expected net benefit of the proposed research is higher than with fixed allocation rules. The optimal allocation of trial entrants enables relevant alternatives to be identified because it is possible to assign no sample to an arm of the trial and in this case it can be ruled out as an irrelevant alternative. By the end of chapter six each of the methodological problems which motivate this thesis are addressed and methods are proposed which can provide a practical solution.

1.3.4 Setting Priorities in Clinical Research

The decision-analytic approach which is developed in chapters four, five and six can provide practical policy tools for research priority-setting. The information generated by clinical research is valued in a way which is consistent with concepts of efficiency, and with the methods used to set priorities in service provision. The

simple numerical examples show that these techniques can be used to identify areas of clinical practice where the cost of uncertainty is high, and where the potential benefits of clinical research will also be high.

Two hurdles are constructed which proposed research must overcome before it can be considered cost-effective. The first hurdle can eliminate proposed research which will not be cost-effective before issues of research design must be addressed. Those proposals which pass the first hurdle can be regarded as potentially cost-effective and can be considered at the second hurdle. The second hurdle ensures that the design of potentially cost-effective research is technically efficient, and that it will be cost-effective when conducted at the optimal scale. The value of proposed research to the providers and consumers of health services can be established. This approach provides a means to decide which clinical decision problems are worth evaluating in a clinical trial and what is the technically efficient scale of this research.

These methods can be used to rank competing research proposals so that the maximum health benefits can be gained for limited research and development resources. These tools can be used to establish the optimal level of research and development; the optimal allocation of resources between research and development and service provision; and the optimal allocation between different areas of clinical research. There are also further methodological developments which could be pursed and the application of these methods to research priority-setting suggests a programme of empirical work. It also poses the issue of how to implement this type of approach to research priority-setting. These issues are discussed in chapter seven.

Chapter 2 The Value of Diagnostic Information

Contents: Page

2.1 Introduction

Approaches to the evaluation of diagnostic information which are based on treatment thresholds combine measures of accuracy and the consequences of subsequent treatment decisions [23, 63, 64], but they are essentially qualitative: measuring whether information changes patient management by assigning those with positive test results to treatment and those with negative test results to no treatment. They do not attempt to measure the value to patients of the changes in clinical practice brought about by the diagnostic information. However these measures do require information on the consequences of the treatment options which are assumed to be used following diagnosis. This prior information can be used and combined with information on the costs of treatment and testing to provide estimates of the expected value of diagnostic information.

This is the strategy for the economic evaluation of diagnostic information proposed by Phelps and Mushlin [107]. By using information on the costs and outcomes of current practice prior to the introduction of a new diagnostic technology, the value of diagnostic information can be estimated based only on this prior information and on estimates of the accuracy of new diagnostic technology. The prospective evaluation of subsequent treatment following diagnosis in a full clinical trial is then unnecessary. This strategy is an attempt to combine information from a number of sources and to focus clinical research more sharply by (a) eliminating new technologies which will not be cost-effective, (b) by avoiding randomised experimental design where possible, and (c) by focusing on a clinically relevant range of test and patient characteristics.

This is achieved by constructing two hurdles that a new diagnostic technology must overcome before it can be considered cost-effective. The first hurdle compares the expected value of the test assuming that it provides perfect information (the maximum value that any diagnostic technology can provide) with an estimate of the cost of the new technology. If the Expected Value of Perfect

Information (EVPI) is greater than the estimated cost of the test then it is potentially cost-effective and passes the first hurdle. This first hurdle does not require prospective research and is constructed using only information about current practice which is available before the new technology is introduced.

If a proposed diagnostic technology passes the first hurdle then prospective clinical research is required to establish the accuracy of the test. Once the accuracy of the test is established existing information about current patient management is used to estimate the expected value of this imperfect clinical information. If the Expected Value of Clinical Information (EVCI) is greater than the cost of implementing the new diagnostic technology, the test passes the second hurdle and will be cost-effective over some range of prior probability of disease.

This strategy means that an economic evaluation of a new diagnostic technology can be based on measures of accuracy and information which is available prior to the introduction of the test, combining information which is available from a variety of sources. If the new diagnostic technology is non-invasive then randomised patient selection to establish the accuracy of the test is unnecessary, because double-blind diagnosis with the new test and with a "gold standard" test can be performed on the same patient. A controlled clinical trial of the full diagnostic and treatment process can be avoided: saving research and development resources; avoiding delay in adopting cost-effective technology, and avoiding potential health costs to patients who would otherwise have been enrolled in a clinical trial.

In this chapter I apply the Phelps Mushlin strategy to a simple numerical example of the test/treatment decision and explore the relationship between the value of diagnostic information and the critical cost-effectiveness ratio (the shadow price of the budget constraint on service provision) [108, 140, 141]. This demonstrates that both hurdles are dependent on this decision rule and that the explicit monetary

valuation of health outcome is unavoidable in the valuation of clinical information.

The Phelps Mushlin strategy, in common with test/treatment thresholds, assumes that current practice is the appropriate baseline to evaluate a new diagnostic technology. In the context of economic evaluation this requires current practice to be optimal or the most cost-effective strategy at the critical cost-effectiveness ratio which will be used to set priorities in service provision when the new technology has been evaluated. In the next chapter I generalise the Phelps Mushlin strategy to the more realistic situation where there is more than one treatment option available for a particular diagnosis. This requires consistency between the value of health outcome which is implicit in current practice and the critical cost-effectiveness ratio which will be used to decide if the new technology will be cost-effective. If this assumption is violated (because clinicians have a higher implicit value of health outcome or do not perceive all the costs of alternative patient management strategies) then current practice may not be the appropriate baseline (or the relevant alternative), and the analysis can be subject to two types of errors at each of the two hurdles. Type I errors occur when the value of clinical information is overestimated and a diagnostic technology which is not cost-effective is evaluated and implemented. Type II errors will be made when the value of clinical information is underestimated and a potentially cost-effective diagnostic technology is rejected.

This strategy, like the threshold approach, assumes that a new diagnostic technology can simply be added to the existing strategies of patient management and that diagnostic information changes clinical practice by assigning those with a positive test result to treatment and negative test result to no treatment. This assumption will be appropriate when there is only one possible treatment for a particular diagnosis. However, when there is more than one possible treatment for a given diagnosis, the optimal treatment following the test result may not be the same as the optimal treatment choice without the test. In these circumstances diagnostic information may change patient management in two ways: by changing

the optimal treatment choice for a particular diagnosis; and by changing the probability of assigning a patient to a particular diagnosis.

If the diagnostic information changes the optimal treatment for a given diagnosis it is possible that this treatment may not be part of current practice. This is more likely to be the case if the value of health outcome implicit in existing strategies of patient management is not consistent with the critical cost-effectiveness ratio used by the analyst. It may be reasonable to assume that information on the costs and health outcomes of current practice will be available. However it is very unlikely that information about treatment options which are not part of current practice will be available. In these circumstances existing information will be inadequate to estimate the value of new diagnostic information and the investigator may be forced to consider an experimental design which includes these alternative treatments.

The implications are that when approaches to economic evaluation which are based on existing clinical practice are generalised to more complex sequential decision problems they can only provide reliable information in a limited set of circumstances which impose restrictive assumptions that are unlikely to hold in many clinical settings. However, if it is not possible to draw reliable inferences from existing strategies of patient management then prospective clinical research using randomised controlled clinical trials which consider all feasible strategies of patient management may be required. This poses a number of methodological and practical questions. First is the practical question that it will not be feasible, in terms of resource cost, time, and ethical implications to evaluate every possible alternative strategy in this way. If the prospective evaluation of all possible alternatives is not possible or efficient, but simply relying on existing clinical practice is unreliable, this posses a number of methodological questions including: (a) which clinical decision problems should be subject to prospective clinical research; (b) if a clinical decision problem is worth evaluating in a clinical trial which alternatives should be regarded as relevant and should be compared in the

trial; and (c) what is the optimal or technically efficient scale of this research?

These are the methodological issues of allocative and technical efficiency in clinical research which are dealt with in chapters 4, 5 and 6. The purpose of this chapter and the next is to show that approaches to the economic evaluation of diagnostic information which rely on existing strategies of patient management may be inadequate and to pose the problems which are taken up in subsequent chapters of the thesis.

2.2 A Strategy for the Evaluation of Diagnostic Information

The strategy for the economic evaluation of a new diagnostic technology proposed by Phelps and Mushlin [107] will now be presented and applied to a simple numerical example, and the relationship between the value of information and the value placed on health outcome will be explored. This example uses their approach and notation before it is generalised in chapter 3 to the situation where there is more than one treatment option for a particular diagnosis.

2.2.1 A Numerical Example of the Phelps Mushlin Strategy

Phelps and Mushlin take a simple decision problem to illustrate their approach, and consider a single disease with only two possible health states (i=0,1): disease (i=1); and no disease (i=0). It is assumed that clinicians hold prior beliefs about the likelihood of disease in a particular patient or group of identical patients.

f = prior probability of disease
(1-f) = prior probability of no disease

A simple binary diagnostic test is considered, and following Phelps and Mushlin the standard approach to characterising the performance of the test in terms of sensitivity and specificity is used:

p = probability of a true positive test result (sensitivity)
(1-p) = probability of a false negative test result
q = probability of a false positive test result
(1-q) = probability of a true negative test result (specificity)

The clinician is faced with a simple decision problem with only two possible treatment options (tj; j=0,1): treatment (t_1); and no treatment (t_0), where t_1 follows

a positive test result and t_0 follows a negative result. The decision problem facing the clinician is whether to (a) test and treat according to the test results, or (b) not to test and either treat or not treat. In this simple decision problem there is only one possible treatment for a given diagnosis and it is this restriction which will be relaxed in the next chapter.

The health outcomes of patients depend on health state and the treatment option chosen, and are described as health state utilities, U_{ij} were i indicates health state and j indicates the treatment selected.

U_{11} = utility of a diseased patient treated (t_1)
U_{10} = utility of a diseased patient not treated (t_0)
U_{01} = utility of a healthy patient treated (t_1)
U_{00} = utility of a healthy patient not treated (t_0)

The costs of treatment C_{ij} also depend on health state i and treatment j:

C_{11} = cost of a diseased patient treated (t_1)
C_{10} = cost of a diseased patient not treated (t_0)
C_{01} = cost of a healthy patient treated (t_1)
C_{00} = cost of a healthy patient not treated (t_0)

It is assumed that all patients are identical with the same characteristics and preferences towards health states and costs. It is also assumed that the costs and utilities of the treatment options for a given health state are the same before and after the introduction of the test. This excludes the possibility that: (a) the test itself has any therapeutic value; (b) it is an essential prerequisite to certain treatment; or (c) involves any notion of "process utility" where some value is gained by the patient from the process of testing over and above the increased probability of being treated appropriately following the test. This also excludes the possibility of the test providing prognostic as well as diagnostic information and it

ensures that knowledge of the costs and utilities of current practice before the test is introduced will be sufficient to evaluate the new diagnostic device. In this simple example it is assumed that the test is non-invasive and poses no risks to the patient.

Following the notation of Phelps and Mushlin incremental utilities (ΔU_{ijk}) can also be defined as the utility gained from treating a patient correctly according to their health state, so that ΔU_{ijk} is the utility gained by treating a patient in health state i with treatment j rather than treatment k.

$\Delta U_{110} = (U_{11}-U_{10})$ = incremental utility of treating a diseased patient with t_1
$\Delta U_{001} = (U_{00}-U_{01})$ = incremental utility of not treating a healthy patient

Similarly incremental costs (ΔC_{ijk}) can also be defined as the cost of treating a patient in state i with treatment j rather than treatment k.

$\Delta C_{110} = (C_{11}-C_{10})$ = incremental cost of treating a diseased patient with t_1
$\Delta C_{001} = (C_{00}-C_{01})$ = incremental cost of not treating a healthy patient

The decision problem facing the clinician for an individual patient before the diagnostic device is available is a simple choice between treatment with t_1 and no treatment. When the diagnostic test is available the clinician can also decide to test and treat according to the test results, where t_1 will follow a positive result and t_0 will follow a negative test result.

Using the notation for this simple problem the question of whether the clinician should choose treatment or no treatment without the diagnostic test and whether the clinician should use the diagnostic test when it becomes available can be addressed. The impact on the expected health outcomes and expected costs of implementing this diagnostic technology can also be established. Indeed with this limited information it is possible to estimate how much the clinician should be

willing to pay for perfect information, and also for the imperfect clinical information generated by the new test. These two values of information are the essential elements in the first and second hurdles respectively, and can be compared to an estimate of the cost of the new diagnostic technology to establish whether it will be cost-effective.

Table 2.2

It is assumed that information on the U_{ij} and C_{ij} (which are part of current practice before the new test is introduced) is available. The values of all the variables used in this simple numerical example are reported in table 2.2. The approach proposed by Phelps and Mushlin combines information from a variety of sources including observations of current practice and makes inferences about a new diagnostic technology based solely on information which should be available before the new test is introduced. However the Phelps Mushlin strategy does not provide any explicit method to take into account the variable quality of information from different sources. This issue is addressed in chapters 4, 5 and 6 by using a Bayesian approach where prior distributions are assigned to the key variables.

2.2.2 Selecting the Fallback Strategy

The first step in the evaluation of a new diagnostic device is to establish the appropriate baseline against which it should be compared. The appropriate baseline or the relevant alternative is what the clinician should do in the absence of further diagnostic information. This is called the fallback strategy by Phelps and Mushlin. The value of information is the additional value that it provides and this depends crucially on what the clinician would do if the information was not available. In this simple example before the introduction of the test the decision problem facing the clinician is to treat with t_1 or not to treat (t_0). With no other

diagnostic information this decision must be based on the clinician's prior belief about the likelihood of disease (f). This single-stage decision problem is illustrated in figure 2.2.1 and it represents current practice before the new test is available. The problem is to select the optimal fallback strategy in the absence of diagnostic information. This optimal fallback strategy will be the appropriate baseline (relevant alternative) with which to compare a new diagnostic device.

Figure 2.2.1

Following the decision rules used in the traditional cost-effectiveness approach to the choice between two mutually exclusive alternatives [78, 136, 139, 141], the first step is to establish the expected utility of t_j ($E(U_j) = f.U_{1j} + (1-f).U_{0j}$). The incremental expected utility of choosing t_1 rather than t_0 is the difference between $E(U_1)$ and $E(U_0)$:

$$\begin{aligned} E(U_1) - E(U_0) &= f.(U_{11}-U_{10})-(1-f).(U_{00}-U_{01}) \\ &= f.\Delta U_{110}-(1-f).\Delta U_{001} \end{aligned} \qquad 2.2.1a$$

The expected cost of t_j ($E(C_j) = f.C_{1j}+(1-f).C_{0j}$) must also be established and the incremental cost of choosing t_1 rather than t_0 is the difference between $E(C_1)$ and $E(C_0)$:

$$\begin{aligned} E(C_1) - E(C_0) &= f.(C_{11}-C_{10})-(1-f).(C_{00}-C_{01}) \\ &= f.\Delta C_{110}-(1-f).\Delta C_{001} \end{aligned} \qquad 2.2.1b$$

The traditional cost-effectiveness approach is to establish the incremental cost-effectiveness ratio of choosing t_1 rather than t_0. This represents the cost per unit of health utility gained by moving from t_0 to t_1. To choose between these two options (in the absence of dominance) this ratio must be compared to the critical cost-effectiveness ratio (1/g) which, with a fixed budget or capital constraints, is the cost-effectiveness ratio of the marginal project which will be displaced if t_1 is

implemented and the positive incremental costs of t_1 are incurred [78, 108, 141]. The critical cost-effectiveness ratio is the cost per unit of health utility gained which should be worth paying given the budgetary restrictions on service provision. It is the implicit monetary value placed on health outcome by the existing budget constraint. The decision rule is to choose t_1 if:

$$(f.\Delta C_{110}-(1-f).\Delta C_{001})/(f.\Delta U_{110}-(1-f).\Delta U_{001}) < 1/g \qquad 2.2.2a$$

An equivalent decision rule would be to choose t_1 if the effectiveness-cost ratio is greater than the critical effectiveness-cost ratio (g):

$$(f.\Delta U_{110}-(1-f).\Delta U_{001})/(f.\Delta C_{110}-(1-f).\Delta C_{001}) > g \qquad 2.2.2b$$

The critical effectiveness-cost ratio (g) is the shadow price of the budget constraint on service provision and is the minimum improvement in health outcome per additional unit of cost the clinician should accept. When one alternative does not dominate the other it is not possible to make consistent decisions without reference to a critical ratio or a value of health outcome. However once the a critical ratio has been established [16, 57, 71, 77, 86, 141] then health utilities can be rescaled into monetary terms using 1/g, or equivalently monetary values can be rescaled into health utilities using g. Equation (2.2.2a) can be rearranged and the same decision rule can now be expressed in terms of net benefit measured in health utility. The clinician should choose t_1 if the net benefit of t_1 is greater than the net benefit of t_0, and the incremental net benefit of t_1 is positive:

$$(E(U_1)-g.E(C_1))-(EU_0)-g.E(C_0)) > 0 \qquad 2.2.2c$$

Or alternatively choose t_1 if the net benefit of treating those with disease is greater than the net benefit of not treating those without the disease:

$$f.(\Delta U_{110}-g.\Delta C_{110}) > (1-f).(\Delta U_{001}-g.\Delta C_{001}) \qquad 2.2.2d$$

If t_1 is chosen then the gains from treating those with disease (ΔU_{110}-g.ΔC_{110}) with a probability of f should exceed the benefits forgone (opportunity costs) from being unable to not treat those with no disease (ΔU_{001}-g.ΔC_{001}) with probability (1-f).

An entirely equivalent decision rule would be to rescale health outcome into monetary terms by multiplying (2.2.2c) or (2.2.2d) through by 1/g, and then the net benefits would be measured on a monetary rather than a health utility scale. These decision rules will be equivalent to the net present value decision rule used in a Paretian cost-benefit analysis in the special case where all individuals have identical preferences and their marginal willingness to pay for additional health utility is the same and is equal to 1/g [108].

What is clear from this discussion of decision rules is that when dominance does not exist placing a monetary value on health outcome is absolutely unavoidable if decisions based on cost-effectiveness (or "cost-utility") analysis are to be made. The only issue is whether this decision rule is made explicit by the analyst or whether it is abdicated to social decision-makers, where it may remain implicit and not open to criticism or alternative formulation. In this example, where the fallback strategy must be selected before the value of information can be established, the decision rule must be explicit and the explicit monetary valuation of health outcome by the analyst is unavoidable. This is a general characteristic of all sequential clinical decision problems (which involve contingent decisions and the comparison of a number of mutually exclusive alternative strategies) which poses some interesting methodological issues which are discussed in chapter 3.

Treatment Thresholds

The selection of the fallback strategy can also be expressed in terms of a treatment threshold for f. The information in 2.2.2d can be rearranged to solve for the prior probability of disease where the clinician should be indifferent between t_1 and t_0. The point of indifference can be found by simply setting both sides of 2.2.2d equal to each other and solving for f to give the treatment threshold f_{10}:

$$f_{10} = (\Delta U_{001} - g.\Delta C_{001})/((\Delta U_{110} - g.\Delta C_{110}) + (U_{001} - g.\Delta C_{001})) \qquad 2.2.3$$

The clinician should select t_1 if the prior probability (f) is greater than f_{10} but select t_0 if f is less than f_{10}. When f is equal to f_{10} the clinician is indifferent between t_1 and t_0. At this point the clinician is most uncertain about which fallback treatment to select, and one would expect the value placed on information to reach a maximum at this point of uncertainty. The following parts of this chapter demonstrate that the value of diagnostic information does indeed reach a maximum at these treatment thresholds.

The relationship between f_{10} and the value of health outcome in this numerical example is illustrated in figure 2.2.2 and this demonstrates that the treatment threshold falls as the value placed on health outcome increases.

Figure 2.2.2

Figure 2.2.2 summarises efficient clinical practice for a range of possible values of health outcome. When the value placed on health outcome is low (less than £3,500) $f_{10}=1$, and t_1 is never selected because the additional health benefits of t_1 are not worth the additional costs. Indeed at extreme values when 1/g=0 no value is placed on health benefits and this "accountancy decision rule" is simple cost-minimisation irrespective of health benefits. The least cost option (t_0) will be selected in all circumstances. At the other extreme when g is equal to zero 2.2.3

collapses to the clinical treatment threshold. No value is placed on resource costs and this purely clinical decision rule selects the most effective treatment. An infinite value is placed on health outcome. As 1/g approaches infinity f_{10} tends to its limit ($\lim(f_{10})$) which in this numerical example is at a prior probability of disease of 1/3.

2.2.3 The Test/Treatment Decision

Selecting the efficient fallback strategy is essential because it provides the appropriate baseline or relevant alternative against which the information provided by a diagnostic device can be valued. When the new diagnostic device is available the clinician faces the choice of whether to use the test and treat according to the test results, where t_1 follows a positive test result and t_0 follows and negative result, or to choose not to test and use the efficient fallback strategy. The value of information is the difference between the net benefit when the test is used and the net benefit of the fallback strategy. The clinician should choose to test if the value of information is greater than the costs of the test. This decision problem is illustrated in figure 2.2.3.

Figure 2.2.3

Almost all diagnostic devices are imperfect and the accuracy of this proposed new test is expressed in terms of sensitivity (p) and specificity (1-q). The test is imperfect and it produces false positive results (q), and false negative results (1-p). Initially it is assumed that there is only one way to operate the test and there is no possible trade-off between sensitivity and specificity. However this will be relaxed when optimal test operation and ROC challenge regions are discussed.

The decision problem illustrated in figure 2.2.3 follows the presentation adopted by Phelps and Mushlin [107]. The "no test" arm of the decision tree represents the

fallback strategies which were illustrated in figure 2.2.1. However, the "test" arm of the decision tree does not follow the actual chronology of events because the first event should be the result of the diagnostic test rather than the disease state and the second event should be the predictive values rather than sensitivity and specificity. These should occur after the treatment choice has been made. This presentation was adopted by Phelps and Mushlin (and has been used by others) because sensitivity and specificity enter directly into the tree and this avoids Bayesian probability revision. Although it is not intuitively appealing, in this case it is equivalent to a structure which follows the correct chronology using Bayes. The presentation in this chapter and the next follows Phelps and Mushlin for consistency with their approach, but in later chapters of the thesis decision problems will be structured in the correct chronology[39, 56, 59, 115, 132, 137].

2.2.4 The Expected Value of (Imperfect) Clinical Information

The Expected Value of Clinical Information (EVCI) is the difference between the expected net benefit of using the test and using the fallback strategy, so there is an expression ($EVCI_j$) for each of the possible fallback strategies (j).

The expected net benefit of testing is the difference between the expected utility of testing and the expected cost of testing (which is rescaled to utility values using g). For the diagnostic test in figure 2.2.3 the expected utility of testing is:

$$f.(p.U_{11}+(1-p).U_{10})+(1-f).((1-q).U_{00}+q.U_{01}) \qquad 2.2.4a$$

and the expected cost of testing is:

$$f.(p.C_{11}+(1-p).C_{10})+(1-f).((1-q).C_{00}+q.C_{01}) \qquad 2.2.4b$$

The EVCI when treatment is the fallback

The incremental utility of testing when treatment t_1 is the fallback strategy is simply the difference between the expected utility of testing (2.2.4a) and the expected utility of the fallback t_1 ($f.U_{11}+(1-f).U_{01}$):

$$-f.(1-p).\Delta U_{110}+(1-f).(1-q).\Delta U_{001} \qquad 2.2.5a$$

Similarly the incremental cost of testing (rescaled to utility using g) when t_1 is the fallback strategy is the difference between 2.2.4b and the expected cost of treatment ($f.C_{11}+(1-f).C_{01}$):

$$-f.(1-p).g.\Delta C_{110}+(1-f).(1-q).g.\Delta C_{001} \qquad 2.2.5b$$

The Expected Value of Clinical Information when t_1 is the fallback strategy ($EVCI_1$) is the incremental net benefit of testing, or the difference between the incremental utility (2.2.5a) and the rescaled incremental cost of testing (2.2.5b). The clinician should use the diagnostic test if the $EVCI_1$ is greater than the average variable cost of the test (C'_{te}) rescaled to health utility:

$$\begin{aligned} EVCI_1 = \quad & -f.(1-p).(\Delta U_{110}-g.\Delta C_{110}) \\ & +(1-f).(1-q).(\Delta U_{001}-g.\Delta C_{001}) > g.C'_{te} \end{aligned} \qquad 2.2.5c$$

$$\begin{aligned} \partial EVCI_1/\partial f = \quad & -(1-p).(\Delta U_{110}-g.\Delta C_{110}) \\ & -(1-q).(\Delta U_{001}-g.\Delta C_{001}) < 0 \end{aligned} \qquad 2.2.5d$$

This derivation shows that when the fallback is to treat then the key issue in deciding whether to use the test is the net losses arising from testing and not treating false negatives ($\Delta U_{110}-g.\Delta C_{110}$) compared with the gains from testing and not treating true negatives ($\Delta U_{001}-g.\Delta C_{001}$). The value of information will fall as f increases (2.2.5d) because when f=1 the test will generate some negative results

all of which will be false and some patients with the disease will not be treated, but treatment without testing will not incur any costs of unnecessarily treating healthy patients. So at this extreme the value of imperfect information will be negative (with a fixed combination of p, and q).

The EVCI when no treatment is the fallback

The incremental utility of testing when no treatment (t_0) is the fallback strategy is simply the difference between the expected utility of testing (2.2.4a) and the expected utility of no treatment ($f.U_{10}+(1-f).U_{00}$):

$$f.p.\Delta U_{110}+(1-f).q.\Delta U_{001} \qquad 2.2.6a$$

Similarly the rescaled incremental cost of testing when t_0 is the fallback strategy is the difference between 2.2.4b and the expected cost of no treatment ($f.C_{10}+(1-f).C_{00}$):-

$$f.p.g.\Delta C_{110}+(1-f).q.g.\Delta C_{001} \qquad 2.2.6b$$

The Expected Value of Clinical Information when t_0 is the fallback strategy ($EVCI_0$) is the difference between the incremental utility (2.2.6a) and the rescaled incremental cost of testing (2.2.6b). The clinician should use the diagnostic test if $EVCI_0 > g.C'_{te}$

$$EVCI_0 = f.p.(\Delta U_{110}-g.\Delta C_{110}) -(1-f).q.(\Delta U_{001}-g.\Delta C_{001}) > g.c_{te} \qquad 2.2.6c$$

$$\partial EVCI_0/\partial f = p.(\Delta U_{110}-g.\Delta C_{110}) +q.(\Delta U_{001}-g.\Delta C_{001}) > 0 \qquad 2.2.6d$$

When the fallback is not to treat then the decision can be based on the loss from

testing and treating false positives (ΔU_{001}-g.ΔC_{001}) compared with the gains from testing and treating true positives (ΔU_{110}-g.ΔC_{110}). The $EVCI_0$ will fall as f is reduced (2.2.6d) because when f=0 the test will generate some positive results all of which will be false and some healthy patients will be treated, but no treatment without testing will not incur any costs of not treating patients with the disease. At this extreme the value of imperfect information will also be negative because of false positive results.

2.3 A Strategy for Focusing Clinical Research

2.3.1 Hurdle I: The Expected Value of Perfect Information

The Expected Value of Perfect Information (EVPI) is the maximum value that any proposed diagnostic test could provide. It is an upper bound on the value of information and is used to create the first hurdle in the strategy proposed by Phelps and Mushlin. If the costs of the diagnostic test exceed the maximum EVPI then the test can never be cost-effective and further evaluation is unnecessary. However if EVPI>$g.C'_{te}$ then the diagnostic device is potentially cost-effective over some range of prior probability of disease and the test passes the first hurdle.

The EVPI can be derived from the expressions for $EVCI_j$ by simply setting the sensitivity (p) and specificity (1-q) of the test equal to one in (2.2.5c) and (2.2.6c). The expected value of perfect information when t_1 is the fallback strategy ($EVPI_1$) is as follows:

$$EVPI_1 = (1-f).(\Delta U_{001}-g.\Delta C_{001}) \qquad 2.3.1a$$

$$\partial EVPI_1/\partial f = -(U_{001}-g.C_{001}) < 0 \qquad 2.3.1b$$

The $EVPI_1$ falls with f, but now when f=1 the value of information will be zero because a perfect test does not incur the cost of not treating patients with false negative results.

The $EVPI_0$ (when t_0 is the fallback) is as follows:

$$EVPI_0 = f.(\Delta U_{110}-g.\Delta C_{110}) \qquad 2.3.1c$$

$$\partial EVPI_0/\partial f = (U_{110}-g.C_{110}) > 0 \qquad 2.3.1d$$

The $EVPI_0$ increases with f but now when f=0 the value of information is zero because there are no costs of treating false positive results.

The $EVPI_j$ can be calculated based solely on the utilities and costs of t_1 and t_0 and the prior probability of disease. Since t_1 and t_0 are part of current practice and the utilities, costs, and probabilities required are implicit in current clinical decision-making, it may be reasonable to assume that this information is available prior to the introduction of the test. Indeed this first hurdle can be constructed without any knowledge of the characteristics of a proposed diagnostic device and it relies only on the decision-problem currently facing clinicians.

This first hurdle is constructed for the numerical example by calculating the EVPI for the full range of the prior probability of disease. The EVPI for this numerical example is illustrated in figure 2.3.1a and is rescaled to monetary values using a value of health outcome of £4,000 per unit of health utility gained.

Figure 2.3.1a

From equation (2.2.3) and figure 2.2.2 when 1/g=£4,000 the treatment threshold f_{10} is equal to 0.8. When $f<f_{10}$ no treatment is the fallback strategy and the $EVPI_0$ rises with f, but when $f>f_{10}$ treatment is the fallback and the $EVPI_1$ falls with f, and when $f=f_{10}$ the clinician is indifferent between t_1 and t_0. At this point the EVPI reaches a maximum of £3,200 in figure 2.3.1a where the clinician is most uncertain about which fallback strategy to adopt.

The first hurdle compares the EVPI to an estimate of the variable cost of the new diagnostic device. Figure 2.3.1a illustrates two proposed diagnostic technologies. The first has an estimated variable cost of C'_{te1}=£3,500, which is greater than the maximum EVPI. It will never be cost-effective and should be rejected at this first hurdle. The second test has a cost of C'_{te2}=£2,500 and it is potentially cost-effective for patients with a prior probability of disease between f_0 and f_1. This diagnostic test passes the first hurdle and is potentially cost-effective, but clinical research to establish the characteristics of the test is required before the cost-effectiveness of this test can be established and the new technology implemented.

The EVPI in figure 2.3.1a represents the value of perfect information for an individual patient with a particular prior probability of disease. However a new diagnostic technology may have large fixed costs ($\overline{C}_{te}$) associated with its implementation, and the prevalence of disease in patient population which enters the decision problem in figure 2.2.3 may have a continuous distribution of $\phi(f)$. The first hurdle can be amended so that the population EVPI is compared to the total cost of implementing the proposed technology. The proposed test will pass the first hurdle if:

$$\int_{f_0}^{f_{10}} EVPI_0(f)\phi(f)df + \int_{f_{10}}^{f_1} EVPI_1(f)\phi(f)df > g.C'_{te}\int_{f_0}^{f_1}\phi(f)df + g.\overline{C}_{te} \quad 2.3.1e$$

The relationship between the expected value of perfect information and the value placed on health outcome is illustrated in figure 2.3.1b and 2.3.1c. As the value of health outcome increases the value of information also increases and the first hurdle is dependent on the selection of 1/g.

The value of health outcome determines two aspects of the EVPI: the point at which information is most valuable; and the value placed on that information. In figure 2.3.1b the treatment threshold f_{10} where the EVPI reaches a maximum falls with 1/g from 0.8 when 1/g=£4,000, to 0.414 when 1/g =£20,000. This relationship between 1/g and f_{10} was illustrated in figure 2.2.2 and indicates that as the value placed on health outcome increases the more effective but more costly alternative becomes optimal at lower prior probabilities. The value of information also increases and when 1/g=£4,000 the maximum EVPI is £3,200, but when 1/g is increased to £20,000 the maximum EVPI rises to £28,128.

Figure 2.3.1b

Figure 2.3.1c

These two aspects of the relationship between the EVPI and 1/g are illustrated in figure 2.3.1c, but this also illustrates that when 1/g is low (less than £3,100) the

EVPI is negative. This is because treatment following a true positive test result is not worth the additional costs, and this is also demonstrated in figure 2.2.2 where $f_{10}=1$ when $1/g<£3,100$, and it is never efficient to choose t_1 even when the probability of disease is very high.

2.3.2 Hurdle II: The Expected Value of (Imperfect) Clinical Information

If a proposed diagnostic technology passes the first hurdle it is potentially cost-effective and the next step is to establish the characteristics of the test through prospective clinical research. Once the sensitivity (p) and specificity (1-q) of the test are established the EVCI can be estimated and compared the cost of the test. If the new test is non-invasive then randomised patient selection will be unnecessary because diagnosis with the new test and with a "gold standard test" can be conducted on the same patients.

The EVCI is calculated for this numerical example and is illustrated in figure 2.3.2a. The EVCI reaches a maximum at $f_{10}=0.8$, but the maximum EVCI is lower than the maximum EVPI due to false positive and false negative results. The $EVCI_1$ falls with f when $f>f_{10}$ and the $EVCI_0$ rises with f when $f<f_{10}$, illustrating equations (2.2.5d), and (2.2.6d) above.

Figure 2.3.2a

The EVCI becomes negative at extreme values of f: when f=1 the imperfect test will produce some negative results all of which will be false and some patients with the disease will not be treated, in this case it is better to use the fallback strategy of treating everybody; when f=0 the test will produce some positive results all of which will be false and some patients will receive unnecessary treatment, in this case it will be better not to treat without testing.

At the second hurdle the EVCI of a potentially cost-effective test is compared to an estimate of the variable cost of the new test. Three estimates of the cost of the new device are illustrated in figure 2.3.2a. If the estimated cost of the new test is high (C'_{te1}=£3,500) the test fails at the first hurdle: it is not potentially cost-effective and no further research is required. When the estimated costs are lower (C'_{te2}=£2,500) the test passes the first hurdle and is potentially cost-effective, but when the accuracy of the test is established it fails at the second hurdle. However if the estimated cost was lower still (C'_{te3}=£1,500) the test will pass both the first and second hurdles and will be cost-effective for patients with a prior probability of disease between f_0 and f_1.

The second hurdle can be amended in the same way as the first so that the population EVCI can be compared to the total cost of implementing the proposed technology including any fixed element. The proposed test will pass the second hurdle and should be implemented if:

$$\int_{f_0}^{f_{10}} EVCI_0(f)\phi(f)df + \int_{f_{10}}^{f_1} EVCI_1(f)\phi(f)df > g.C'_{te}\int_{f_0}^{f_1}\phi(f)df + g.\overline{C}_{te} \quad 2.3.2$$

The relationship between the EVCI and 1/g is illustrated in figure 2.3.2b and 2.3.2c, and the value of imperfect clinical information also increases with the value placed on health outcome. Both the first and second hurdles are dependent on the value selected for 1/g.

Figure 2.3.2b

Figure 2.3.2c

The value of health outcome determines the point at which information is most valuable, as well as the value placed on clinical information. Just as in figure 2.3.2b the treatment threshold (f_{10}) where the EVCI reaches a maximum falls with 1/g. The maximum EVCI also increases with 1/g: from £2,240 when 1/g=£4,000

to £19,416 when 1/g = £20,000. However figure 2.3.2b and 2.3.2c also illustrate that the value placed on false negative and false positive results also increases with 1/g and the EVCI at extreme values of f where EVCI<0 falls as 1/g increases.

2.3.3 The Optimal Test Operation

The construction of the second hurdle requires information on the accuracy of the test. So far it has been assumed the test is binary and there is only one combination of p and q which is available. However most diagnostic tests allow a trade-off between the sensitivity of the test and the specificity. The possible combinations of p and q, and the trade-off available can be described by an ROC curve. The information used to construct the first hurdle can also be used to identify the optimal combinations of p and q where the test should be operated.

The approach taken by Phelps and Mushlin takes the full differential with respect to p and q of an expression for the net benefit of testing (NB_{te}). The optimal trade-off between p and q (dp/dq) or slope of the ROC curve can then be found. The net benefit of testing is given by:

$$NB_{te} = f.(p.U_{11}+(1-p).U_{10})+(1-f).((1-q).U_{10}+q.U_{01}) \\ -g.(f.(p.C_{11}+(1-p).C_{10})+(1-f).((1-q).C_{10}+q.C_{01})) \qquad 2.3.3a$$

The full differential of the net benefit of testing:

$$dNB_{te} = f.dp.((U_{11}-U_{10})-g.(C_{11}-C_{10}))-(1-f).((U_{00}-U_{01})-g.(C_{00}-C_{01})).dq \\ = f.dp/dq.(\Delta U_{110}-g.\Delta C_{110})-(1-f).(\Delta U_{001}-g.\Delta C_{001}) \qquad 2.3.3b$$

by setting dNB_{te} =0 and solving for dp/dq:

$$dp/dq = (1-f).(\Delta U_{001}-g.\Delta C_{001}) / f.(\Delta U_{110}-g.\Delta C_{110})$$
$$= EVPI_1/EVPI_0 \qquad 2.3.3c$$

The optimal dp/dq or slope of the ROC curve is the ratio of the expected value of perfect information when t_1 is the fallback, to the expected value of perfect information when t_0 is the fallback. This is equivalent to setting the consequences of designating a result as positive equal to the consequences of designating a result as negative and rearranging so the likelihood ratio or the slope of the ROC curve is the product of the ratio of prior probability of no disease and disease and the ratio of the net consequences of no disease and disease [72, 107, 137].

The optimal dp/dq is simply the ratio of the expected costs of false positive and false negative results and optimal test operation is similar to establishing technical efficiency in production by ensuring that the ratio of factor input prices is equal to the marginal rate of technical substitution between these inputs in production. Production will be technically efficient at a point of tangency between the isoquant (ROC curve) and the budget constraint (dp/dq). If the relative factor prices (expected cost of false positive and false negative results) change then the slope of the budget constraint (dp/dq) will change. The point of tangency with the isoquant (point on the ROC curve) will shift and the technically efficient factor input (optimal combination of sensitivity and specificity) also changes.

Figure 2.3.3

Equation (2.3.3c) shows that the optimal test operation is dependent on both the prior probability of disease and the value placed on health outcome. Figure 2.3.3 illustrates how the range of prior probabilities from the first hurdle in figure 2.3.1a can be used to define the economically relevant portion of the ROC curve. The optimal points dp_1/dq_1 and dp_0/dq_0 are based on the range of prior probabilities of

disease where the test is potentially cost-effective (f_1 and f_0 respectively). dp_1/dq_1 and dp_0/dq_0 defines the economically relevant portion of the ROC curve. When f is low (at f_0=0.38) the expected cost of false positive results is relatively high and the optimal point on the ROC curve will substitute reduced false positives for increased false negatives (dp_0/dq_0=6.53). When f is higher (f_1=0.905) the expected cost of false negatives is relatively high and optimal test operation will substitute reduced false negatives for increased false positives (dp_1/dq_1=0.42).

This strategy makes it clear that it is unnecessary to evaluate the whole ROC curve and the first hurdle can be used to focus prospective research on the economically relevant range of test characteristics. Points on the ROC curve with a slope greater than dp_0/dq_0 or less than dp_1/dq_1 are irrelevant because the diagnostic test will never be cost-effective when operated beyond these points. Phelps and Mushlin go on to define ROC challenge regions: the minimum combinations of p and q which a new device must achieve to be cost-effective. The first hurdle then asks if there is any prospect of a new technology achieving a point within the challenge region, if not the test fails at the first hurdle. This approach also makes it clear that the optimal operation of diagnostic technology is dependent on specifying a value of health outcome. If economic criteria are used to evaluate a new technology then the same criteria must be used to establish how it should be operated before an economic evaluation can take place. The selection of 1/g prior to the economic evaluation is unavoidable.

2.3.4 Focusing Clinical Research

The Phelps Mushlin strategy and their proposed hurdles can be constructed based solely on prior information about health outcomes, costs, prior probabilities of disease and a measure of the accuracy of the test. It can combine information which may already be available from a variety of sources. Health outcomes can be based on literature review and observation of current practice and health state

utilities can be elicited using established methods. Similarly, information on prior probabilities and the distribution of probabilities of disease can be based on published epidemiological studies. The costs of treatment can be estimated by observing current practice or may be readily available from routinely collected information. However the approach proposed by Phelps and Mushlin does not include any way to take into account the variable quality of information from different sources, and this issue will be addressed in chapters 4, 5, and 6 where prior distributions are assigned to the health outcomes and path probabilities.

This strategy not only focuses clinical research by eliminating those proposed new technologies which will not be cost-effective, it also focuses prospective clinical research on potentially cost-effective devices, and on those variables where prior information is not available. The only information that will require prospective research is the accuracy of the new test, and if the test is non-invasive this can be done without patient selection and randomised design can be avoided. Furthermore if the diagnostic test allows a trade-off between sensitivity and specificity the first hurdle can define the economically relevant portion of the ROC curve and the clinical evaluation of the whole ROC curve will be unnecessary. It provides an approach to the economic evaluation of clinical information which can avoid randomised trial designs which include both the diagnostic and treatment process. This could substantially reduce the cost of research and development in terms of resources, the opportunity cost of delaying the implementation of cost-effective technology (or providing evaluative evidence before an unproven diagnostic technology is widely implemented), and the opportunity costs to individuals enrolled in less effective arms of a clinical trial.

Appendix A Tables and Figures for chapter 2

Contents:

Table 2.2 A Numerical Example

Utilities (U_{ij})		Incremental Utilities (ΔU_{ijk})		Costs (C_{ij})		Incremental Costs (ΔC_{ijk})	
U_{11}	6	ΔU_{110}	4	C_{11}	£12,000	ΔC_{110}	£12,000
U_{01}	8	ΔU_{001}	2	C_{01}	£8,000	ΔC_{001}	-£8,000
U_{10}	2			C_{10}	£0		
U_{00}	10			C_{00}	£0		

The characteristics of the proposed diagnostic test are: p=0.9, and q=0.2

Figure 2.2.1 Decision Tree for the Fallback Decision

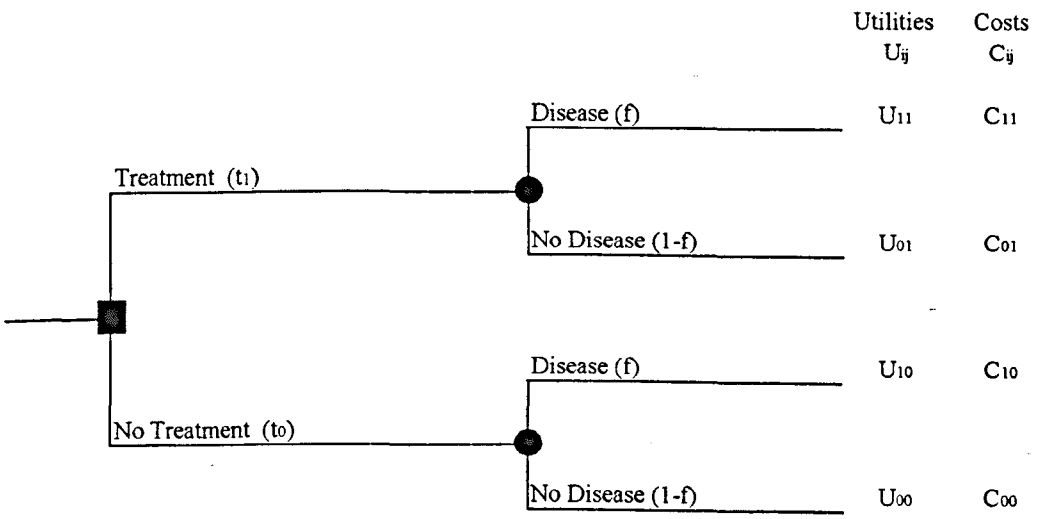


Figure 2.2.2 Treatment Threshold (f_{10})

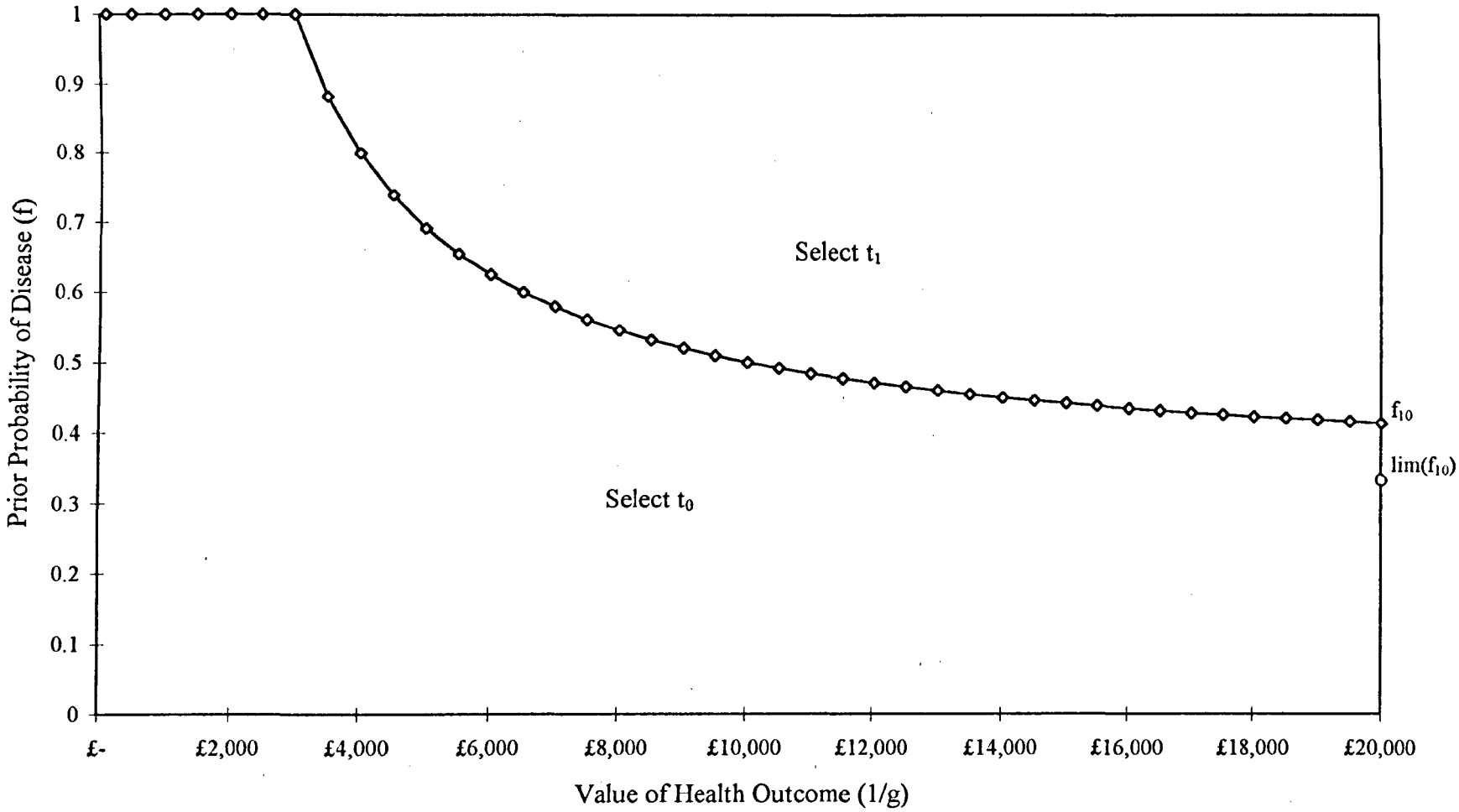


Figure 2.2.3 Decision Tree for the Test Treatment Decision

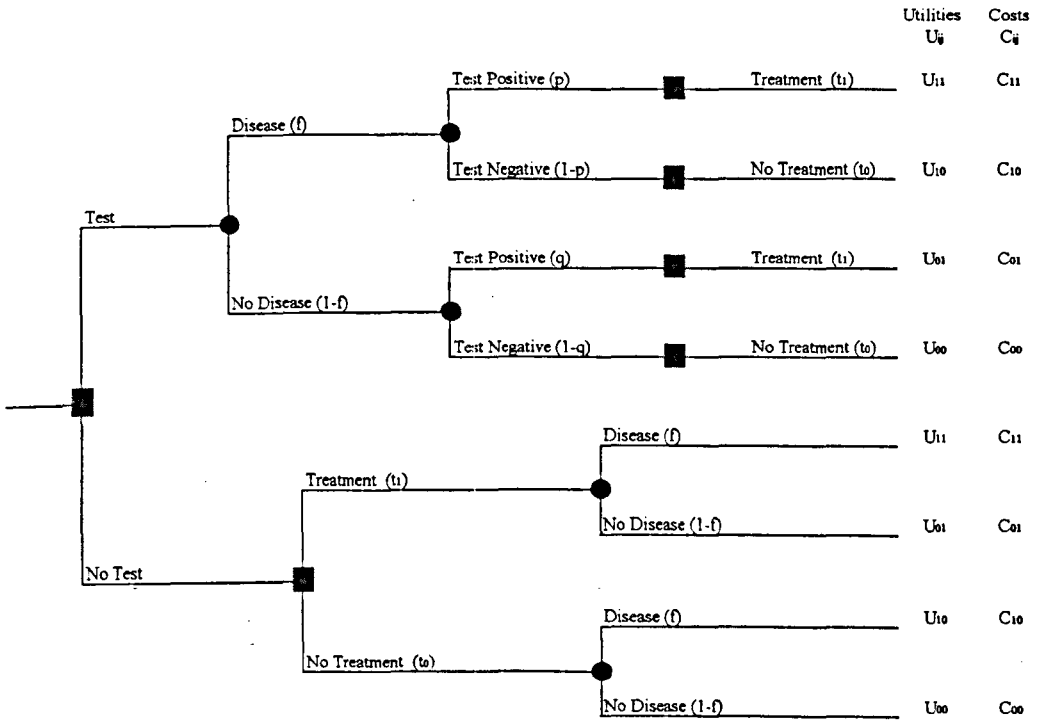


Figure 2.3.1a Expected Value of Perfect Information (EVPI) (1/g=£4,000)

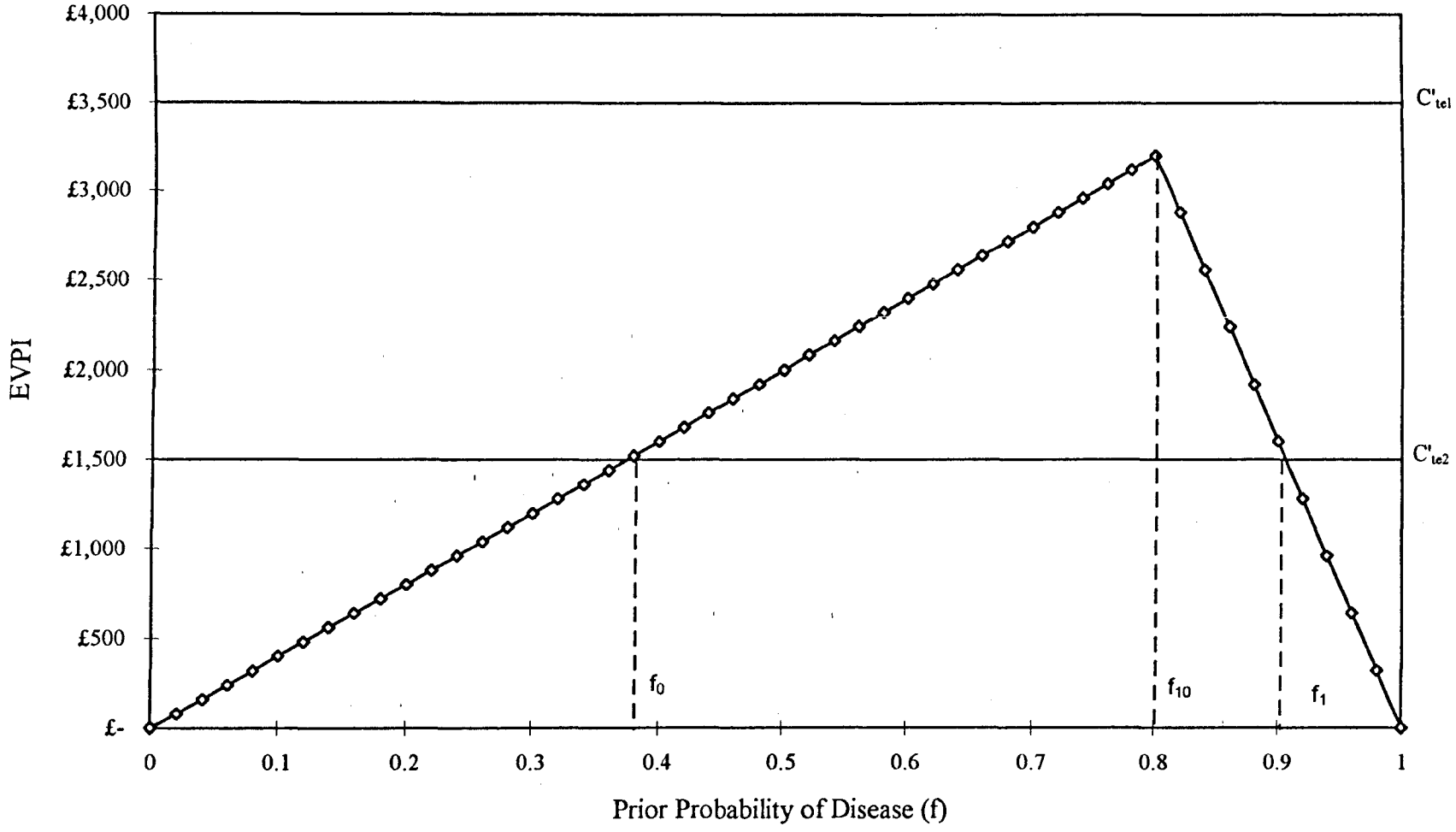


Figure 2.3.1b Expected Value of Perfect Information (EVPI)

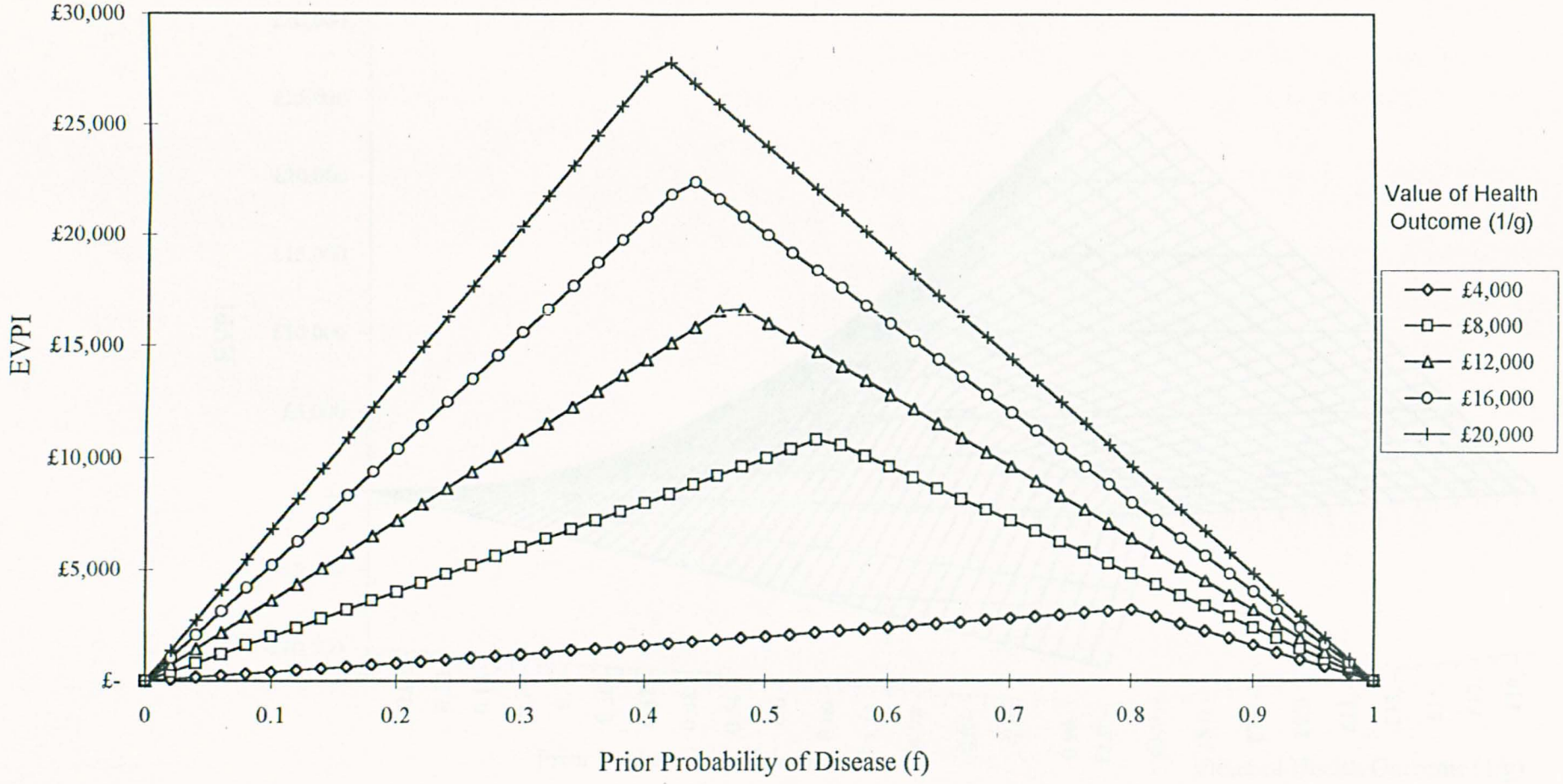


Figure 2.3.1c The Expected Value of Perfect Information (EVPI)

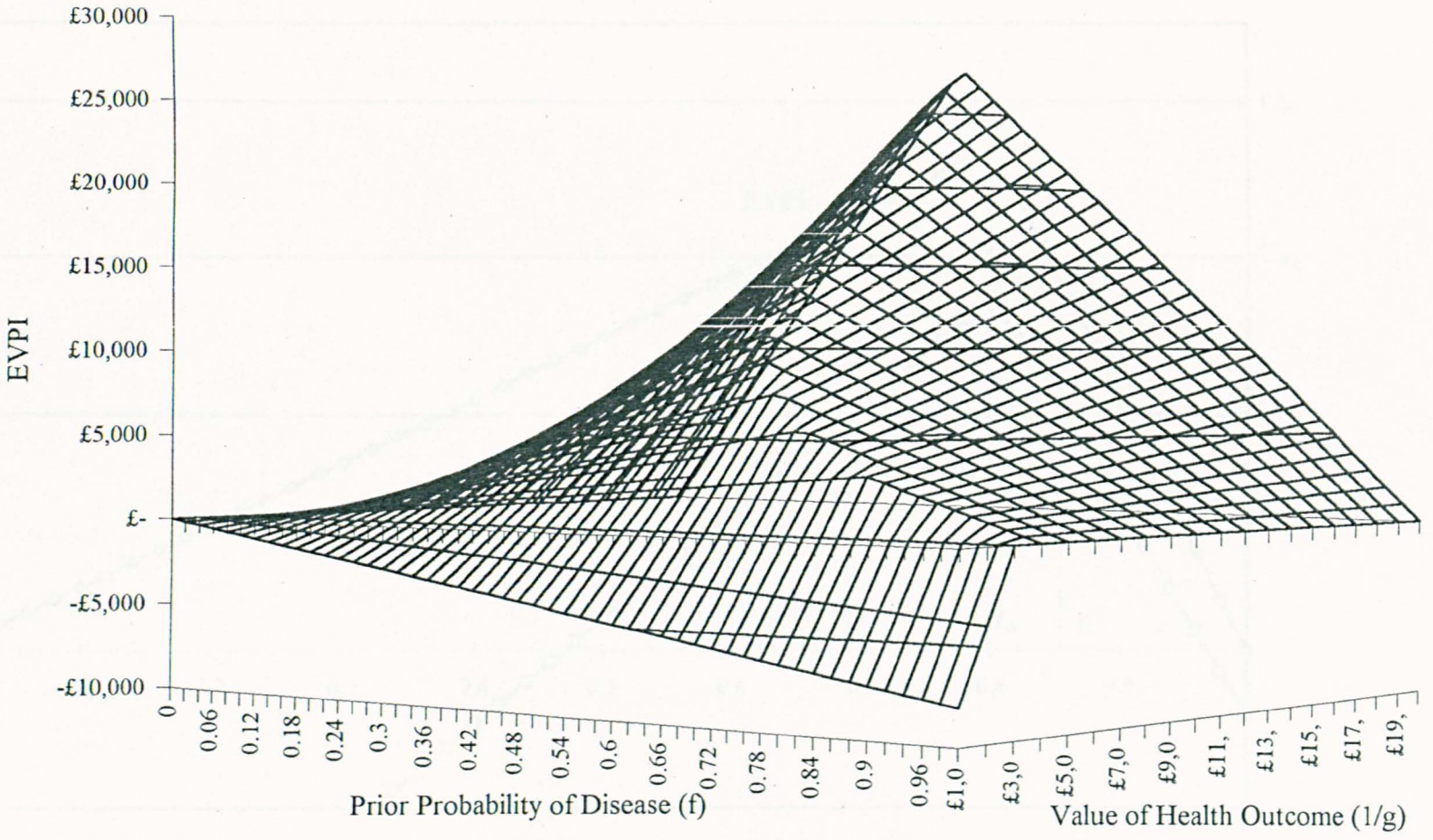


Figure 2.3.2a The Expected Value of Clinical Information (1/g=£4,000)

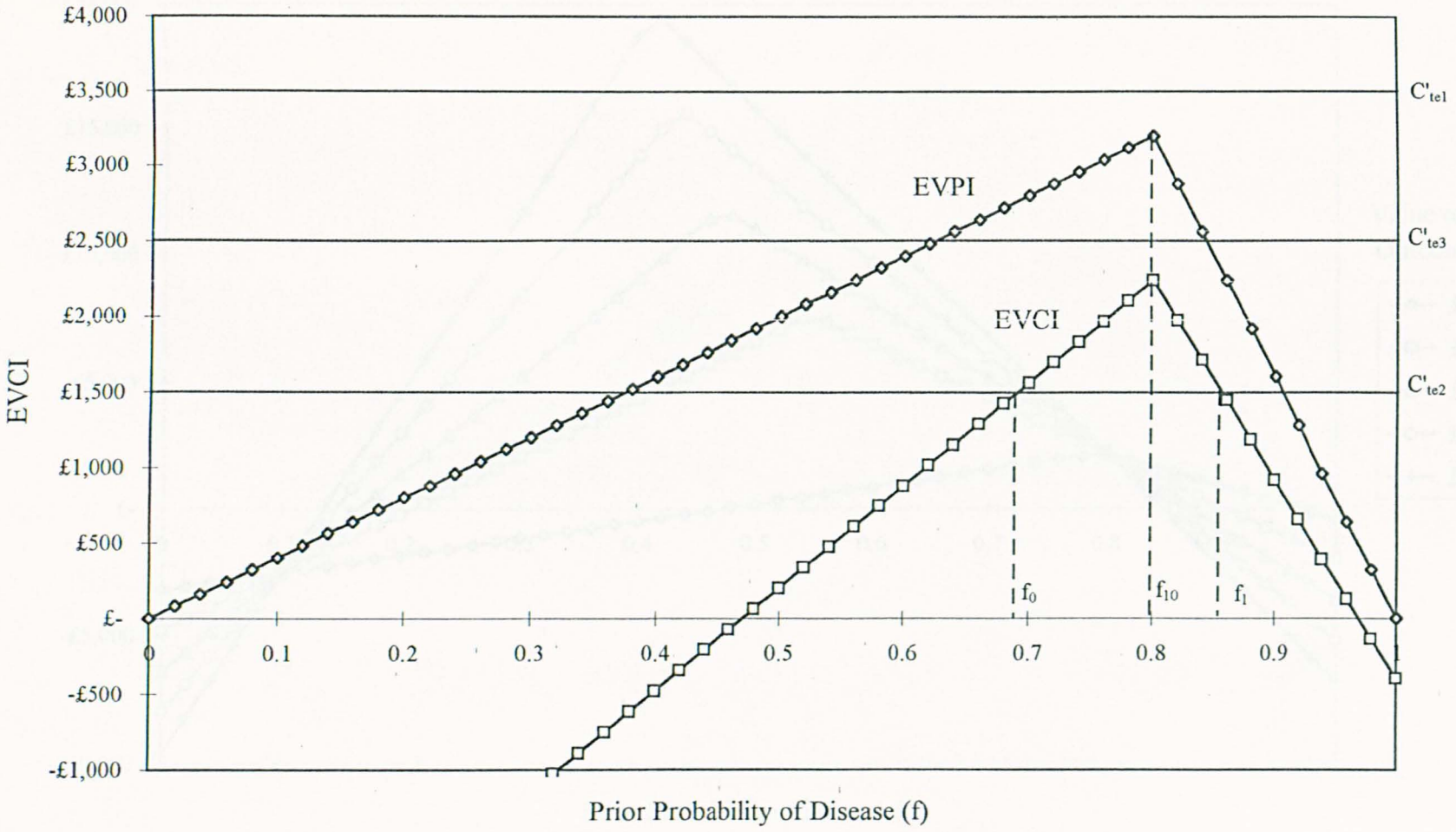


Figure 2.3.2b The Expected Value of Clinical Information (EVCI)

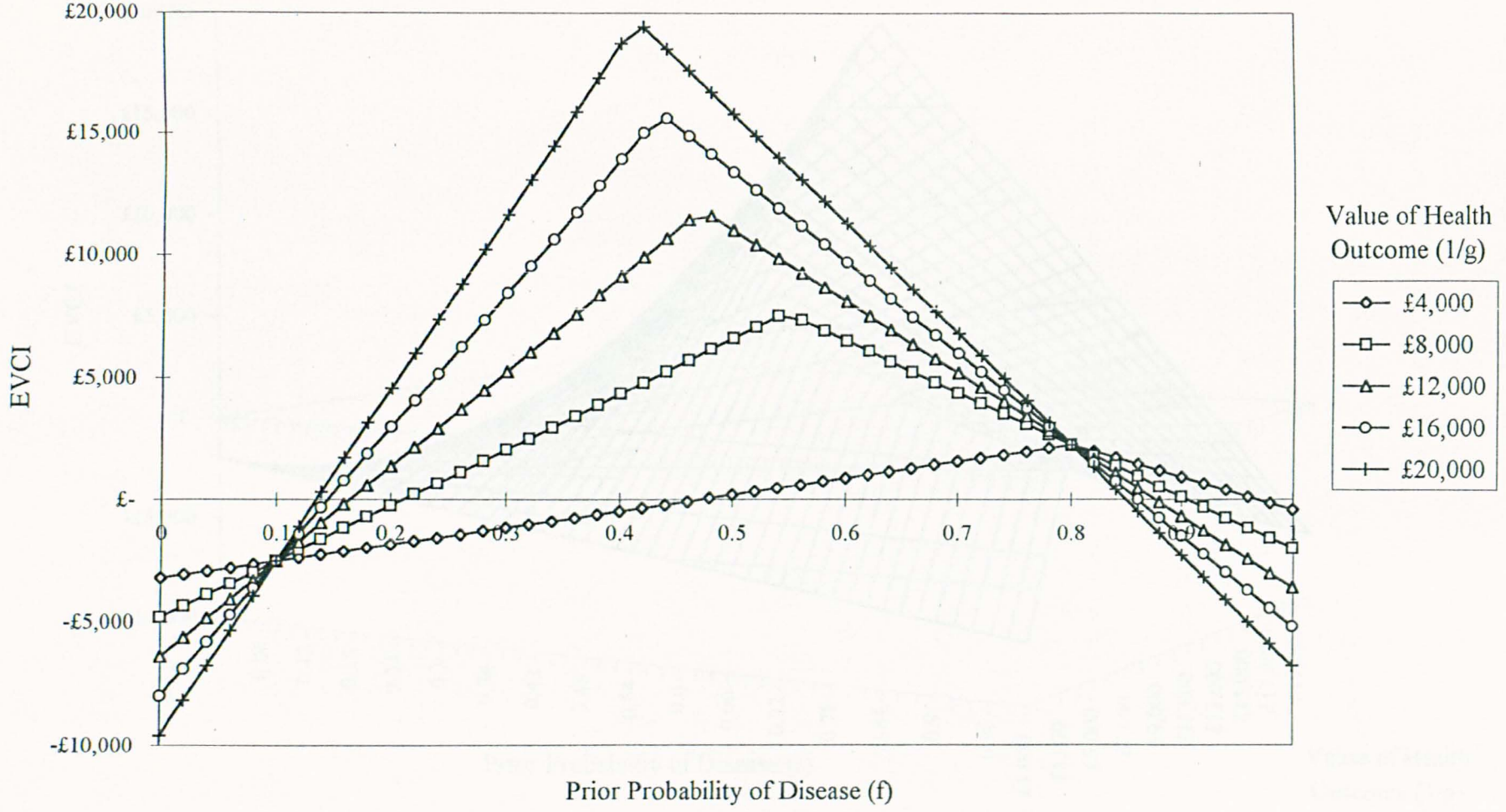


Figure 2.3.2c Expected Value of Clinical Information (EVCI)

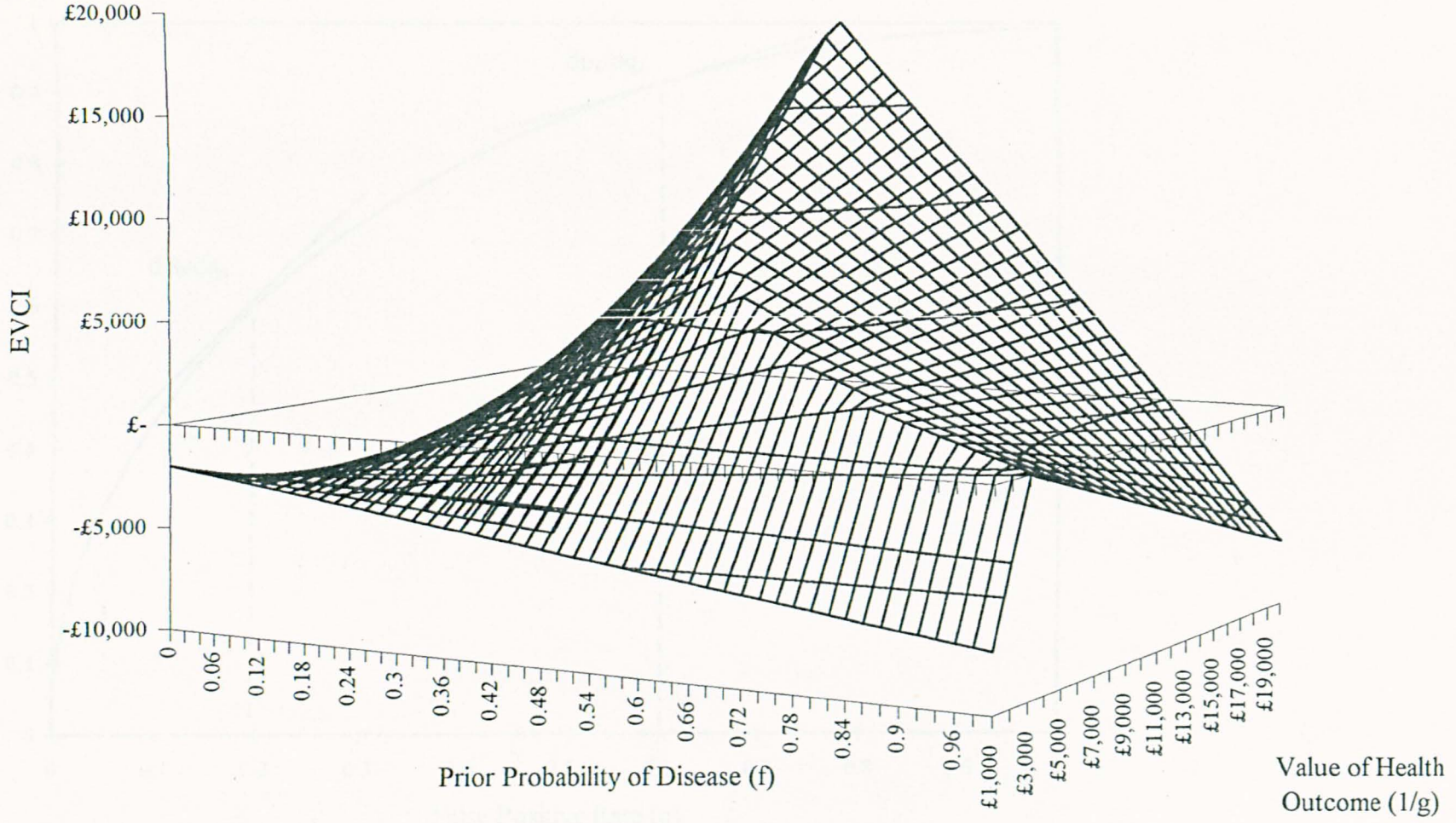
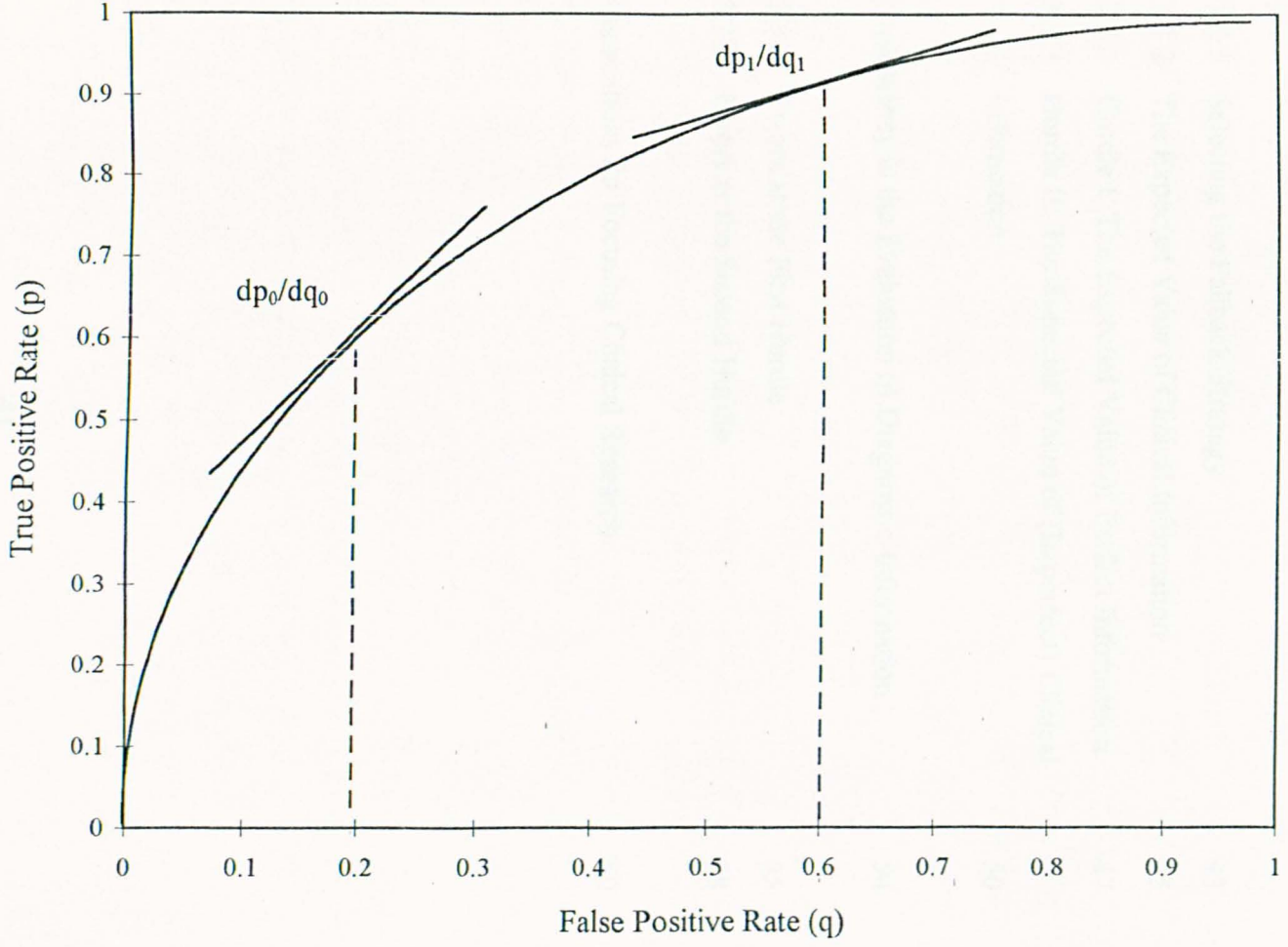


Figure 2.3.3 Reciever Operating Characteristic (ROC) Curve



Chapter 3 Consistency in the Evaluation of Diagnostic Information

Contents: Page

3.1 Introduction

The approach to the economic evaluation of diagnostic information which was outlined in the previous chapter has a number of advantages and attractions: it is consistent with economic decision rules; it enables research to be focused on new diagnostic technologies which are potentially cost-effective; and it can combine information that is available from a number of different sources. However the presentation of the approach by Phelps and Mushlin and its application to a simple numerical example assumed that there is only one possible treatment for a given diagnosis. This is clear in figure 2.2.3 where the decision problem facing the clinician is simply whether to test and treat according to test results: where t_1 follows a positive test and t_0 follows a negative test. Treatment following diagnosis is determined only by the test result and this implies that no other treatment alternatives are possible. The approach has simply added a diagnostic device to existing patient management strategies and clinical practice is changed only to the extent that the test changes the probability of assigning a patient to a particular diagnosis.

However in most clinical decision problems there is a range of treatment alternatives (and other diagnostic processes) which are at least possible following the results of the test, even if these alternatives are currently not used as part of existing patient management. In these circumstances where there is more than one treatment alternative for a given diagnosis, diagnostic information may change patient management by changing the probability of assigning a patient to a particular diagnosis and by changing the optimal treatment choice. To establish the circumstances in which this strategy will be appropriate to these less restrictive decision problems the first step is to generalise the Phelps Mushlin strategy to accommodate more than one treatment for a given diagnosis and identify the assumptions that are required. The second step is to examine the consequences of violating these assumptions.

3.2 Generalising the Phelps Mushlin Strategy

In this section the approach is generalised and applied to the same numerical example but with an additional treatment option (t_2). This treatment is less costly than t_1 but is less effective for those with the disease, although it has fewer side effects than t_1 for patients with no disease. The details of this numerical example are reported in table 3.2.1. The decision problem facing the clinician before the test is introduced is to choose either t_0, t_1, or t_2. Once the test is introduced the clinician can choose to test and treat with either t_1 or t_2 following a positive result, and t_0 following a negative result. This decision problem is illustrated in figure 3.2.1.

Table 3.2.1

Figure 3.2.1

Following chapter 2 it may be reasonable to assume that some information on the costs and health utilities of these three treatment alternatives will be available if they are part of existing patient management before the test is introduced.

3.2.1 Selecting the Fallback Strategy

Following the presentation of the previous section the first step is to establish the appropriate baseline or fallback strategy. The clinician must now decide whether to treat with either t_1 or t_2 or to select no treatment. The same decision rules from (2.2.2a) and (2.2.2b) can be applied to the choice between t_0 and t_1 and t_2. The equivalent decision rule from (2.2.2d) can be expressed in terms of net benefits and t_j (j=0, 1, 2) should be chosen rather than t_k (k=0, 1, 2) if:

$$f.(\Delta U_{1jk}-g.\Delta C_{1jk}) > (1-f).(\Delta U_{0kj}-g.\Delta C_{0kj}) \qquad 3.2.1a$$

This decision rule can also be expressed in terms of a treatment threshold for f and t_j should be chosen rather than t_k if $f>f_{jk}$ where:

$$f_{jk} = (\Delta U_{0kj}-g.\Delta C_{0kj})/((\Delta U_{1jk}-g.\Delta C_{1jk})+(U_{0kj}-g.\Delta C_{0kj})) \qquad 3.2.1b$$

There are now three treatment thresholds: f_{20}, where the clinician is indifferent between t_0 and t_2; f_{10}, where the clinician is indifferent between t_1 and t_0; and f_{12} where the clinician is indifferent between t_1 and t_2. These thresholds are illustrated for this numerical example in figure 3.2.2 and summarise the selection of optimal fallback strategies.

Figure 3.2.2

The treatment threshold f_{10} is the same as in figure 2.2.2, but now t_2 is also available to the clinician and in this numerical example $f_{20}\leq f_{10}\leq f_{12}$. Efficient clinical practice is as follows: when $f<f_{20}$ the clinician should select t_0; when $f_{20}<f<f_{12}$ the clinician should select t_2; and when $f>f_{12}$ the clinician should select t_1. Although there are three treatment thresholds there are only two points where the clinician will be most uncertain about selecting the fallback strategy: when they are indifferent between no treatment and treatment (with either t_1 or t_2); and when they are indifferent between the treatment options. A treatment threshold (f^*) can be defined as the minimum of f_{10} and f_{20} (in this case $f_{20}\leq f_{10}$ and $f^*=f_{20}$) and the clinician will be uncertain about selecting treatment or no treatment at $f^*=f_{20}$ and about selecting which treatment option at $f=f_{12}$. There are now two points where the value of clinical information may reach a maximum.

Figure 3.2.2 also illustrates the relationship between the value of health outcome and efficient clinical practice. When 1/g is low (in this case less than £1,500) $f_{20}=f_{10}=f_{12}=1$, and there is no value of f where either treatment would be efficient. t_0 should be chosen in all circumstances. Figure 3.2.2 shows that when t_2 is available it will be efficient to treat at lower prior probabilities of disease. When

1/g=£4,000 the clinician will move from no treatment to treatment when f reaches 0.533 rather than 0.8 when only t_1 is available. Indeed at this value of 1/g, t_1 is never part of current practice because $f_{12}=1$. t_1 only becomes part of the fallback strategy when $1/g \geq £5,000$ and $f_{12}<1$.

As the value placed on health outcome is increased the treatment thresholds fall and in the limit the thresholds collapse to purely clinical decision rules. In this particular numerical example $f_{20}=f_{10}=f_{12}=1/3$ when g=0 and when a purely clinical decision rule is used t_2 will not be part of the fallback strategy because when $f<1/3$, t_0 should be chosen but when $f>1/3$, t_1 should be chosen. If t_2 is not part of existing patient management then information on the utilities and costs of the treatment will not be available by simply observing current practice. By making minor changes to this numerical example it can be shown that t_2 will not be part of existing patient management even when economic rather than purely clinical decision rules are used. This poses a problem for Phelps Mushlin approach which will be examined in more detail in section 3.3 of this chapter.

3.2.2 The Expected Value of Clinical Information

The Expected Value of Clinical Information is the difference in expected net benefit between the test and the fallback strategies. The $EVCI_{hj}$ can be defined for each combination of the three possible fallback strategies (j=0,1,2) and two possible testing strategies (h=1, 2) where the clinician can treat with t_1 (h=1) or t_2 (h=2) following a positive test and t_0 following a negative test.

The $EVCI_{hj}$ is equivalent to the $EVCI_j$ (2.2.5c) when treatment with either t_1 or t_2 is the fallback and the fallback treatment is the same as the treatment which follows a positive test result (j=h=1,2). 3.2.2a is equal to 2.2.5c when j=h=1.

$$EVCI_{hj} = -f.(1-p).(\Delta U_{1j0}-g.\Delta C_{1j0}) + (1-f).(1-q).(\Delta U_{00j}-g.\Delta C_{00j}) \quad 3.2.2a$$

The $EVCI_{hj}$ falls with f and (3.2.2b) is equal to (2.2.5d) when j=h=1.

$$\partial EVCI_{hj}/\partial f = -(1-p).(\Delta U_{1j0}-g.\Delta C_{1j0}) - (1-q).(\Delta U_{00j}-g.\Delta C_{00j}) < 0 \quad 3.2.2b$$

The $EVCI_{hj}$ is equivalent to the $EVCI_0$ (2.2.6c) when no treatment is the fallback strategy (j=0; and h=1,2) and (3.2.2c) is equal to (2.2.6c) when h=1.

$$EVCI_{h0} = f.p.(\Delta U_{1h0}-g.\Delta C_{1h0}) - (1-f).q.(\Delta U_{00h}-g.\Delta C_{00h}) \quad 3.2.2c$$

The $EVCI_{hj}$ rises with f and (3.2.2d) is equal to (2.2.6d) when h=1:

$$\partial EVCI_{h0}/\partial f = p.(\Delta U_{1h0}-g.\Delta C_{1h0}) + q.(\Delta U_{00h}-g.\Delta C_{00h}) > 0 \quad 3.2.2d$$

Now that two treatments are available it is possible that the fallback strategy is to treat but the fallback treatment is not the same as the treatment which follows a positive test result ($j \neq h=1, 2$). The diagnostic device not only changes the probability of assigning a patient to a particular diagnosis but it can also change the optimal treatment choice for a given diagnosis. In these circumstances the $EVCI_{hj}$ is not equivalent to the $EVCI_j$.

$$EVCI_{hj} = f.(p.(\Delta U_{1h0}-g.\Delta C_{1h0})-(\Delta U_{1j0}-g.\Delta C_{1j0})) - (1-f).(q.(\Delta U_{00h}-g.\Delta C_{00h})-(\Delta U_{00j}-g.\Delta C_{00j}) \quad 3.2.2e$$

The $EVCI_{hj}$ will fall with f if:

$$\partial EVCI_{hj}/\partial f = p.(\Delta U_{1h0}-g.\Delta C_{1h0})+q.(\Delta U_{00h}-g.\Delta C_{00h}) -((\Delta U_{1j0}-g.\Delta C_{1j0})+(\Delta U_{00j}-g.\Delta C_{00j})) < 0 \qquad 3.2.2f$$

If the $\partial EVCI_{12}/\partial f < 0$ then the EVCI will reach a maximum at $f_{20,}$ but if $\partial EVCI_{12}/\partial f > 0$ then the EVCI will reach a maximum at f_{12}. The value of $\partial EVCI_{12}/\partial f$ is determined by the ΔU_{ijk} and ΔC_{ijk} from table 3.2.1, and the value of health outcome: when g=0 the $\partial EVCI_{12}/\partial f > 0$ and the EVCI reaches a maximum at f_{12}, but when 1/g low then $\partial EVCI_{12}/\partial f < 0$ and the EVCI reaches a maximum at f_{20}.

3.2.3 Hurdle I: The Expected Value of Perfect Information.

The $EVPI_{hj}$ is derived from the expressions for the $EVCI_{hj}$ in the same way as in chapter 2 by setting p=1 and q=0. The $EVPI_{hj}$ when treatment is the fallback and the fallback and test/treatment strategies are the same (j=1,2; and h=j) is equivalent to the $EVPI_j$ (3.2.3a is equal to 2.3.1a when h=1), and from 3.2.2a:

$$EVPI_{hj} = (1-f).(\Delta U_{00j}-g.\Delta C_{00j}) \qquad 3.2.3a$$

The $EVPI_{hj}$ falls with f and from 3.2.2b:

$$\partial EVPI_{hj}/\partial f = -(\Delta U_{00j}-g.\Delta C_{00j}) < 0 \qquad 3.2.3b$$

The $EVPI_{hj}$ when no treatment is the fallback (j=0; and h=1,2) is equivalent to the $EVPI_0$ (3.2.3c is equal to 2.3.1c when h=1), and from 3.2.2c:

$$EVPI_{h0} = f.(\Delta U_{1h0}-g.\Delta C_{1h0}) \qquad 3.2.3c$$

The $EVPI_{hj}$ rises with f, and from 3.2.2d:

$$\partial EVPI_{h0}/\partial f = (\Delta U_{1h0} - g.\Delta C_{1h0}) > 0 \qquad 3.2.3d$$

If the fallback is to treat but the fallback treatment is not the same as the treatment following a positive test result ($j \neq h=1,2$) the $EVPI_{hj}$ will not be equivalent to the $EVPI_{j}$.

$$EVPI_{hj} = f.((\Delta U_{1h0} - g.\Delta C_{1h0}) - (\Delta U_{1j0} - g.\Delta C_{1j0})) - (1-f).(q.(\Delta U_{00h} - g.\Delta C_{00h}) - (\Delta U_{00j} - g.\Delta C_{00j})) \qquad 3.2.3e$$

The $EVPI_{hj}$ may rise or fall with f depending on the values of ΔU_{ij}, ΔC_{ij}, and 1/g, and from 3.2.2f:

$$\partial EVPI_{hj}/\partial f = (\Delta U_{1h0} - g.\Delta C_{1h0}) - ((\Delta U_{1j0} - g.\Delta C_{1j0}) + (\Delta U_{00j} - g.\Delta C_{00j})) \qquad 3.2.3f$$

The first hurdle for this numerical example is constructed by calculating the EVPI for the full range of prior probability of disease and this is illustrated in figure 3.2.3 where the EVPI has been rescaled into monetary values using 1/g=£4,000.

Figure 3.2.3a

The optimal treatment following a positive test result is t_2. The $EVPI_{hj}$ reaches a maximum at f_{20} where the clinician is indifferent between t_0 and t_2 so when $f<f_{20}$ the fallback is not to treat, and from (3.2.3d) the $EVPI_{20}$ rises with f. When $f>f_{20}$ the fallback is to treat with t_2 and from (3.2.3b) the $EVPI_{22}$ will fall with f. Now that t_2 is available, t_1 is not part of either the fallback or the test treatment strategy. If the cost of the test is £1,500 then it is potentially cost-effective when $f_0<f<f_2$.

Figures 3.2.3b and 3.2.3c illustrate the relationship between the $EVPI_{hj}$ and the value of health outcome. The optimal treatment following a positive test is t_1 when 1/g is increased to £8,000 in figure 3.2.3b. The optimal fallback strategies

are as follows: when $f<f_{20}$ the fallback is t_0 and the EVPI rises with f; when $f>f_{12}$ the clinician should select t_1 as both the fallback and test treatment strategy and the EVPI falls with f; however when $f_{20}>f>f_{12}$ the clinician should select t_2 as the fallback but t_1 as the optimal treatment following a positive test result. The $EVPI_{12}$ falls with f when $1/g \leq £12,000$ and it reaches a maximum at f_{20}, but when $1/g>£12,000$ the $EVPI_{12}$ rises with f and it reaches a maximum at f_{12}. So once again the value placed on health outcome determines the point at which information is most valuable, and the value placed on the information.

Figure 3.2.3b

Figure 3.2.3c

Table 3.2.2 details the optimal fallback and testing strategies which lie behind these figures, and shows that t_1 does become part of both the fallback and testing strategy when 1/g is increased. The shaded area indicates the circumstances in which the fallback is to treat but the fallback and treatment strategies differ. This is where the test changes the optimal treatment for a given diagnosis as well as changing the probability of being assigned to a particular diagnosis. If a diagnostic test not only changes the probability of being assigned to a particular diagnosis but also the optimal treatment for a given diagnosis then it will not be appropriate to simply add a diagnostic device to existing (fallback) strategies of patient management.

Table 3.2.2

The first hurdle operates in the same way as in the previous chapter: if the cost of the test is greater then the maximum value of the EVPI (at $f=f_{20}$, or $f=f_{12}$) then the device will never be cost-effective; but if the maximum EVPI exceeds the cost then the proposed test is potentially cost-effective over some range of prior probability of disease. The population EVPI can be estimated and compared to the total costs of implementing the new technology. In figure 3.2.3a (1/g=£4,000)

the population EVPI from (2.3.1e) requires only minor amendment. The device is potentially cost-effective if:

$$\int_{f_l}^{f_{ll}} EVPI_{20}(f)\phi(f)df + \int_{f_{ll}}^{f_l} EVPI_{22}(f)\phi(f)df > g.C'_{te}\int_{f_l}^{f_l}\phi(f)df + g.\overline{C}_{te} \quad 3.2.3g$$

In figure 3.2.3b when 1/g=£8,000 and the variable cost is C'_{te2} then the population EVPI requires further amendment. The device will be potentially cost-effective if:

$$\int_{f_l}^{f_{ll}} EVPI_{10}(f)\phi(f)df + \int_{f_{ll}}^{f_l} EVPI_{12}(f)\phi(f)df + \int_{f_{ll}}^{f_l} EVPI_{11}(f)\phi(f)df$$

$$> g.C'_{te2}\int_{f_l}^{f_l}\phi(f)df + g.\overline{C}_{te2} \quad 3.2.3h$$

3.2.4 Hurdle II: The Expected Value of (Imperfect) Clinical Information.

The second hurdle is constructed for this numerical example by calculating the $EVCI_{hj}$ for the full range of prior probability of disease. This second hurdle when 1/g=£4,000 is illustrated in figure 3.2.4a. The optimal treatment following a positive test result is t_2. Treatment t_1 is not part of either the fallback or the test treatment strategies. The $EVCI_{20}$ rises with f and the $EVCI_{22}$ falls with f so that the EVCI reaches a maximum at f_{20}. If the cost of the test is C'_{te1}=£1,500 then it will be cost effective when $f_0>f>f_2$ in figure 3.2.4a.

Figure 3.2.4a

The relationship between the EVCI and the value placed on health outcome is illustrated in figure 3.2.4b and figure 3.2.4c. The optimal treatment following a

positive test result now changes with the prior probability of disease. In figure 3.2.4b when 1/g=£8,000 the optimal strategy is to use t_2 following a positive test when the prior probability of disease is less than 0.34, but to use t_1 when f>0.34. As the value placed on health outcome increases, the probability of disease where the clinician is indifferent between using t_2 or t_1 as the test treatment strategy falls, because t_1 is more costly than but more effective than t_2.

Figure 3.2.4b

Figure 3.2.4c

Table 3.2.3 details the optimal fallback and test treatment strategies which lie behind these figures. Treatment t_1 only becomes both the optimal fallback and testing strategy at higher prior probabilities of disease and at higher values of 1/g. Indeed t_1 becomes the optimal treatment following a positive test while t_2 remains the optimal fallback strategy. In this case the test changes the optimal treatment for a given diagnosis and the circumstances where the diagnostic test can't simply be added to existing patient management are indicated by the shaded areas in table 3.2.3.

Table 3.2.3

There are now four combinations of fallback and test/treatment strategy which make up the second hurdle and it is clear that a new diagnostic device can not simply be added to existing strategies of patient management. The test changes the optimal treatment for a given diagnosis when $f_{20}>f>f_{12}$ and the optimal treatment following a positive test results also changes with the prior probability of disease.

The second hurdle operates in the same way as in chapter 2 and if the cost of the test is greater than the maximum value of the EVCI then there is no range of prior probability of disease where the test will be cost-effective and it should be

rejected. If the maximum value of the EVCI exceeds the cost of the test then there will be a range of prior probability of disease where the test will be cost-effective. In the example illustrated in figure 3.2.4b the $EVCI_{12}$ falls with f when 1/g=£8,000 and the EVCI reaches a maximum at f_{20}, however when 1/g≥£12,000 the $EVCI_{12}$ rises with f and the EVCI reaches a maximum at f_{12}. The population EVCI can be calculated in the same way as the population EVPI (but taking account of the changes in the test treatment strategy) and compared to the total cost of implementing the new technology.

The preceding example illustrates how the Phelps Muslin strategy can be generalised to take account of more complex clinical decision problems where there is more than one treatment for a given diagnosis, and both hurdles can be constructed. However once more than one treatment for a given diagnosis is available a new diagnostic device cannot simply be added to existing patient management. Diagnostic information can now change patient management not only by changing the probability of assigning a patient to a particular diagnosis but also by changing the optimal treatment for a given diagnosis.

Once an alternative treatment is available it is not necessarily the case that the treatment which may be optimal once the diagnostic device is in place will be part of existing patient management. In this example when 1/g<£5,000 t_1 is not part of existing patient management. Similarly when a purely clinical decision rule is used t_2 will not be part of the fallback strategy and with minor changes to this numerical example t_2 will never be part of current practice even when economic criteria are used to select current practice. If a treatment is not part of the fallback strategy then information on the U_{ij} and C_{ij} will not be available from observing current practice . The investigator may be forced to consider an experimental design which includes both the diagnostic test and the subsequent treatment choices. In these circumstances random allocation may be unavoidable. This generalisation has imposed a number of assumptions which are unlikely to hold in many clinical settings. These assumptions are discussed in more detail in the next section of this

chapter and the implications of violating these assumptions are illustrated using the same numerical example.

3.3 Consistency in the Evaluation of Diagnostic Information

Although the strategy of economic evaluation proposed by Phelps and Mushlin can be generalised to more complex decision problems it depends critically on two assumptions:

Firstly, as already noted in chapter 2, it is assumed that the decision problem facing the clinician prior to the introduction of the test must be identical to the decision problem when the test is introduced and the test results are known. The utilities and costs for a particular disease state and treatment alternative are identical before and after the test is introduced. This assumption enables the prior information about the U_{ij} and C_{ij} from current practice to be used to estimate the EVPI and EVCI. This assumption may be violated if: (a) the test also provides prognostic information; (b) the test results are required to direct treatment (for example coronary angiography prior to coronary artery surgery); or (c) if the results of the proposed diagnostic test are not conditionally independent of other diagnostic tests which may be part of current practice.

Secondly any approach to economic evaluation which accepts current practice as an appropriate baseline (or relevant alternative) to evaluate a new diagnostic device implicitly assumes that the existing strategies of patient management are correct. In the context of an economic evaluation this means that existing strategies of patient management must be cost-effective at the critical cost-effectiveness ratio (value of health outcome) selected to evaluate the new device. Current practice will only be the relevant alternative if there is consistency between the value of health outcome which is implicit in the selection of current practice (1/g) and the value of the critical cost-effectiveness ratio (CCER) which will be used to decide if the new technology will be cost-effective.

This assumption is unlikely to hold because the appropriate critical cost-effectiveness ratio (the shadow price of the budget constraint) is uncertain (due to

incomplete information on competing programmes within the budget) and depends crucially on the perspective of the evaluation which determines the budget that is regarded as relevant [16, 108, 141]. There are a number of reasons to believe that the value of health outcome implicit in existing patient management may be greater than a CCER selected by an analyst. For example: clinicians may only consider clinical effectiveness or have a higher (infinite) implicit value of health outcome; they may not perceive all the costs of the alternative patient management strategies; or they may not have full information about the budget constraint they face and the competing programmes within the budget.

If the value of health outcome implicit in the selection of current practice is greater than the CCER then an analysis which uses current practice as a baseline to value a new diagnostic device may overestimate the value of diagnostic information because it will be compared to an inefficient fallback strategy. If both the fallback and the test/treatment strategy are selected using a value of health outcome which is inconsistent with the CCER then the analysis can be subject to two types of errors at each of the two hurdles. The value of clinical information can be overestimated and a diagnostic technology which is not cost-effective may be accepted. In addition a second type of error will be made when the value of clinical information is underestimated and a potentially cost-effective diagnostic technology is rejected. These potential biases are illustrated using the same numerical example and are discussed in detail below.

3.3.1 Errors at the First Hurdle

At the first hurdle a new diagnostic test is potentially cost-effective if the maximum EVPI is greater than the cost of the new test and from section 3.2 it can be seen that the EVPI will reach a maximum at either f_{20} or f_{12}. If the fallback strategy is selected using an implicit value of health outcome which is inconsistent with the CCER selected to evaluate the new diagnostic device then the value of

information will be overestimated and the point at which the EVPI reaches a maximum will also be biased. This is illustrated in figure 3.3.1a where the CCER is £4,000 but the value of health outcome implicit in the selection of the fallback strategy is £20,000.

Figure 3.3.1a

This inconsistency in the decision rules used to select current practice and to evaluate a new diagnostic device leads to an overestimation of the value of perfect information because the optimal testing strategy is compared to an inefficient alternative. When the CCER=1/g=£4,000 the test treatment strategy is t_2 and the optimal fallback is t_0 when $f<0.54$ and t_2 when $f\geq0.54$. However when 1/g=£20,000 current practice differs from the optimal fallback strategy in a number of important respects and the value of perfect information is overestimated. When $0.36<f<0.46$ the $EVPI_{22}$ overestimates the $EVPI_{20}$ and when $f\geq0.46$ the $EVPI_{21}$ overestimates the $EVPI_{22}$. The EVPI reaches a maximum at $f_{12}=0.46$ rather than $f_{20}=0.54$. The EVPI is seriously overestimated because the test is not being compared to the relevant alternative (the optimal fallback at the CCER). The discontinuities in the EVPI are due to the fact that the alternatives have been selected using one decision rule (1/g=£20,000) but then valued using another (CCER=£4,000). There is a danger that a diagnostic test which cannot be cost-effective will pass the first hurdle. A diagnostic test which costs between £9,300 and £3,000 will pass this first hurdle but at a CCER of £4,000 it is not potentially cost-effective and should be rejected. If the cost of the test was less than £3,000 the range of prior probability of disease where it will be regarded as potentially cost-effective will be overestimated, biasing estimates of the population EVPI.

Figure 3.3.1b

The first hurdle is very sensitive to the way in which current practice is selected

and figure 3.2.2b illustrates the errors that will be made if the value of health outcome implicit in the selection of current practice is greater than the CCER=£4,000. Even when 1/g=£5,000 the maximum EVPI will be overestimated. The consequences of these overestimates is that a test may pass the first hurdle when it is not potentially cost-effective and research and development resources will be wasted.

If both current practice and the treatment which will follow a positive test result are selected using an implicit decision rule which is inconsistent with the CCER then two types of errors can be made at the first hurdle. The first type of error will occur when the value of information is overestimated but now a second type of error can also be made where the value of information is underestimated and a potentially cost-effective test may be rejected at the first hurdle.

Figure 3.3.2a

These errors are illustrated in figure 3.2.2a where the CCER=£4,000 but the value of health outcome implicit in the selection of the fallback and the testing strategy is £4,000, £5,000 or £20,000. As before the EVPI may be overestimated if current practice is not the optimal fallback strategy but now the treatment that follows a positive test results will not necessary be optimal at the CCER. In figure 3.3.2a when 1/g=£20,000 and when $0.36<f<0.46$ the $EVPI_{12}$ overestimates the $EVPI_{20}$, and when $f>0.46$ the $EVPI_{11}$ overestimates the $EVPI_{22}$. But now when $f<0.36$ the $EVPI_{10}$ underestimates the $EVPI_{20}$, because although the fallback t_0 is optimal, the testing strategy is inefficient and the value of information will be underestimated.

This second type of error is more clearly illustrated when 1/g=£5,000 and the maximum EVPI is underestimated because although current practice is optimal, the testing strategy selected is inefficient. When $f<0.5$ the $EVPI_{10}$ underestimates the $EVPI_{20}$ and when $0.5 \leq f<0.98$ the $EVPI_{12}$ underestimates the $EVPI_{22}$. The

range of prior probability of disease where the test is potentially cost-effective will be underestimated and estimates of the population EVPI will be biased. In these circumstances it is possible that a potentially cost-effective diagnostic test will be rejected at the first hurdle.

Figure 3.3.2b

When both the testing and fallback strategies are selected using an implicit value of health outcome which is inconsistent with the CCER the first hurdle is very sensitive to differences in these decision rules. The errors which will be made in estimates of the EVPI are illustrated in figure 3.2.2b and in this example one of the two types of error will be always be made if 1/g>CCER.

3.3.2 Errors at the Second Hurdle

Similar errors can occur at the second hurdle but now the consequences are more serious because if the EVCI is overestimated then there is a danger that a diagnostic device which is not cost-effective will pass the second hurdle and will become accepted as part of efficient clinical practice, incurring the opportunity cost of the greater health benefits which could be gained from an alternative use of these resources.

Figure 3.3.3

Figure 3.3.3 illustrates the way that the EVCI will be overestimated when current practice is selected using an implicit value of 1/g=£20,000. The errors follow the same pattern as at the first hurdle because the errors are due to an inefficient fallback strategy being selected rather than differences in the testing strategy. Just as at the first hurdle the $EVCI_{12}$ overestimates the $EVCI_{20}$, and the $EVCI_{21}$ overestimates the $EVCI_{22}$. A diagnostic test with an estimated cost between

£8,400 and £2,000 will pass this second hurdle and become part of what is regarded as efficient clinical practice. However at a CCER of £4,000 it is not cost-effective and should be rejected. A test which costs less than £2,000 is cost-effective but the range of f where the test should be used will be overestimated, biasing estimates of the population EVCI.

Figure 3.3.4a

Figure 3.3.4b

If both the fallback and the testing strategies are selected using an implicit value of health outcome which is inconsistent with the CCER then the EVCI may be under or overestimated and there is now a possibility that either a cost-effective test will be rejected or an inefficient test will be accepted. This is illustrated in figure 3.3.4a. When the implicit value of health outcome is £20,000 then the $EVCI_{10}$ underestimates the $EVCI_{20}$ when $0.12 \leq f \leq 0.36$; the $EVCI_{12}$ underestimates the $EVCI_{20}$ when $0.36 < f < 0.46$; and when $f \geq 0.46$ the $EVCI_{11}$ underestimates the $EVCI_{22}$. Similarly when the value of 1/g implicit in the selection of the alternative strategies is £5,000 the $EVCI_{22}$ will be underestimated by the $EVCI_{12}$ when $0.84 < f < 0.98$.

The second hurdle is sensitive to the decision rule that is used to select the alternatives which are compared in the economic evaluation. If there is an inconsistency in the decision rule implicit in the selection of these alternatives and the decision rule which will be used to decide whether the new test is cost-effective then the estimates of the value of information will be biased. There is a danger that one of the two errors could be made at the second hurdle. These errors in the estimates of the EVCI are also illustrated in figure 3.3.4b for a range of values of 1/g. This shows that for this numerical example there will be some range of f where the value of information will be either under of over estimated if there is any discrepancy in the value of 1/g used to select alternatives and the CCER selected by the analyst to evaluate the new device.

3.4 Implications for Focusing Clinical Research

Section 3.2 demonstrated that the Phelps Mushlin approach can be generalised to more complex decision problems, but when more than one treatment is available for a given diagnosis the optimal treatment following the test result may not be the same as the optimal treatment choice without the test. This demonstrates that simple measures of accuracy, including the area under the ROC curve, which only measure the ability of a test to assign patients to a particular diagnosis, will be inappropriate because they do not measure the impact of information on changing the optimal treatment for a given diagnosis. Similarly other intermediate output measures from a diagnostic process which are based on measures of accuracy or assignment, such as the number of cases found, and assignment strength or assignment potential, will not reflect these changes in patient management and may lead to an underestimate of the value of the diagnostic information.

However the generalisation of the Phelps Mushlin strategy relies on the assumption that current practice (or the fallback strategies) are correct which in this context means that they are the most cost-effective strategies at the CCER which will be used to evaluate the new test. The numerical example in section 3.3 demonstrates that the value of diagnostic information will be overestimated when the existing fallback strategies are not optimal at the CCER, and will be underestimated when the fallback is optimal but the test treatment strategy is inefficient at the CCER. These are examples of the errors generated when the alternatives compared in an economic evaluation are not the relevant or efficient alternatives. Clearly almost any proposal can appear to be cost-effective if it is compared to an alterative which is sufficiently inefficient.

This demonstrates that the selection of relevant alternatives depends crucially on the value of the CCER and that relying on the alternatives currently selected by existing clinical practice will introduce bias unless the value of health outcome implicit in current practice is consistent with the CCER. This consistency requires

that in existing clinical practice all competing projects are already allocated efficiently within the budget. This implies that decision-makers have full information about the budget constraint the costs and benefits of all competing programmes; and make consistent decisions using the shadow price of the budget constraint. These are conditions which are unlikely to hold in most clinical settings.

There are good reasons to believe that the value of health outcome implicit in the selection of current practice may be greater than the CCER. In these circumstances any approach to economic evaluation which accepts current clinical practice as a relevant alternative may introduce bias into the analysis. This is an example of "second best" [129] where by applying first best rules (assuming existing clinical practice is efficient and 1/g=CCER) in a second best world (where clinical practice is not necessarily efficient and 1/g>CCER) will bias the results of any evaluation and lead to a further inefficient allocation of resources when a cost-effective test is rejected and an inefficient test is accepted.

In a second best world existing clinical practice cannot be used to identify which alternatives are relevant in an economic evaluation. The alternatives which should be regarded as relevant depends on the decision rule which will be used to evaluate the new technology. In the simple decision problem considered in this chapter there are three fallback and two testing strategies giving six possible comparisons between the test and no test alternatives. In this numerical example when the prior probability of disease is 0.6 then there are 4 comparisons involving five alternatives which will become relevant at different values of 1/g. In table 3.2.3: when $1/g \leq £3,000$ t_0 is the relevant fallback and t_2 is the relevant testing strategy; when $£4,000 \leq 1/g \leq £5,000$ t_2 is the relevant fallback and t_2 is the relevant testing strategy; when $£6,000 \leq 1/g \leq £10,000$ t_2 is the relevant fallback and t_1 is the relevant testing strategy; and when $1/g \geq £11,000$ t_1 is the relevant fallback and t_1 is the relevant testing strategy. There are four comparisons generating four different cost-effectiveness ratios for this new diagnostic test. The correct ratio which

compares relevant alternatives depends on the value of health outcome which will be used to decide if the test is cost-effective [15, 16, 78]. This poses the problem of which ratio should be placed in a league table of cost-effectiveness ratios of competing programmes [47, 61]. The implications of this issue for the traditional approach to priority setting and decision making using cost-effectiveness and cost utility ratios is discussed in chapter 7.

It has been argued that the value of health outcome implicit in exiting patient management is likely to be greater than a CCER selected by an analyst. A less effective but less costly alternative treatment may exist which would be optimal at the CCER but may not be part of current practice and it can no longer be assumed that treatment alternatives which are optimal following the new test will be part of existing patient management. In this numerical example when a purely clinical decision rule is used to select current practice (g=0) t_2 will never be selected. However if the analyst used a CCER of £4,000 to evaluate the new technology t_2 would be the optimal treatment which should follow a positive diagnostic test result. Not only will the EVPI and the EVCI be overestimated if current practice is used as a baseline, but there will be no information about t_2 by simply observing current practice. It will not be possible to estimate the EVPI or the EVCI based on existing information. The investigator may be forced to consider an experimental design which includes both the diagnostic test and the subsequent treatment choices, to establish the value of U_{ij} and C_{ij}. In these circumstances random patient allocation in a clinical trial may be unavoidable even if the diagnostic test is non-invasive.

The Phelps Mushlin approach to evaluating diagnostic information and focusing clinical research is likely to fail when it is generalised to more complex decision problems because: (a) the key assumption of consistency between the value of health outcome implicit in current practice and the CCER is unlikely to hold; (b) when this assumption is violated the values of both the first and second hurdles will be biased; and (c) it is not necessarily the case that information about current

practice before the test is introduced will be sufficient to construct the first hurdle.

This poses some practical and methodological problems for the economic evaluation of sequential clinical decision problems and for clinical research. If valid inferences cannot be based on observing current clinical practice, but the prospective evaluation of all possible alternatives in a sequential clinical decision problem is not possible, efficient, or ethical then: (a) how should information of different quality from different sources be combined consistently and explicitly; (b) which clinical decision problems will be worth evaluating in a clinical trial; (c) if a clinical decision problem is worth evaluating which of the competing alternatives should be compared in a clinical trial; and (d) what is the optimal scale of this prospective research? These are questions about how to establish technical efficiency in clinical and economic research design, and how to achieve allocative efficiency in clinical research across clinical decision problems and between research and service provision. It is these questions which are addressed in the following chapters of the thesis.

Appendix B Tables and Figures for Chapter 3

Contents:

Table 3.2.1 A Numerical Example

Utilities (U_{ij})		Incremental Utilities (ΔU_{ijk})		Costs (C_{ij})		Incremental Costs (ΔC_{ijk})	
U_{11}	6	ΔU_{110}	4	C_{11}	£12,000	ΔC_{110}	£12,000
U_{01}	8	ΔU_{001}	2	C_{01}	£8,000	ΔC_{001}	-£8,000
U_{10}	2	ΔU_{112}	2	C_{10}	£0	ΔC_{112}	£9,600
U_{00}	10	ΔU_{120}	2	C_{00}	£0	ΔC_{120}	£2,400
U_{12}	4	ΔU_{121}	-2	C_{12}	£2,400	ΔC_{121}	-£9,600
U_{02}	9	ΔU_{002}	1	C_{02}	£2400	ΔC_{002}	-£2,400

The characteristics of the proposed diagnostic test are: p=0.9, and q=0.2

Figure 3.2.1 Decision Tree for the Test Treatment Decision

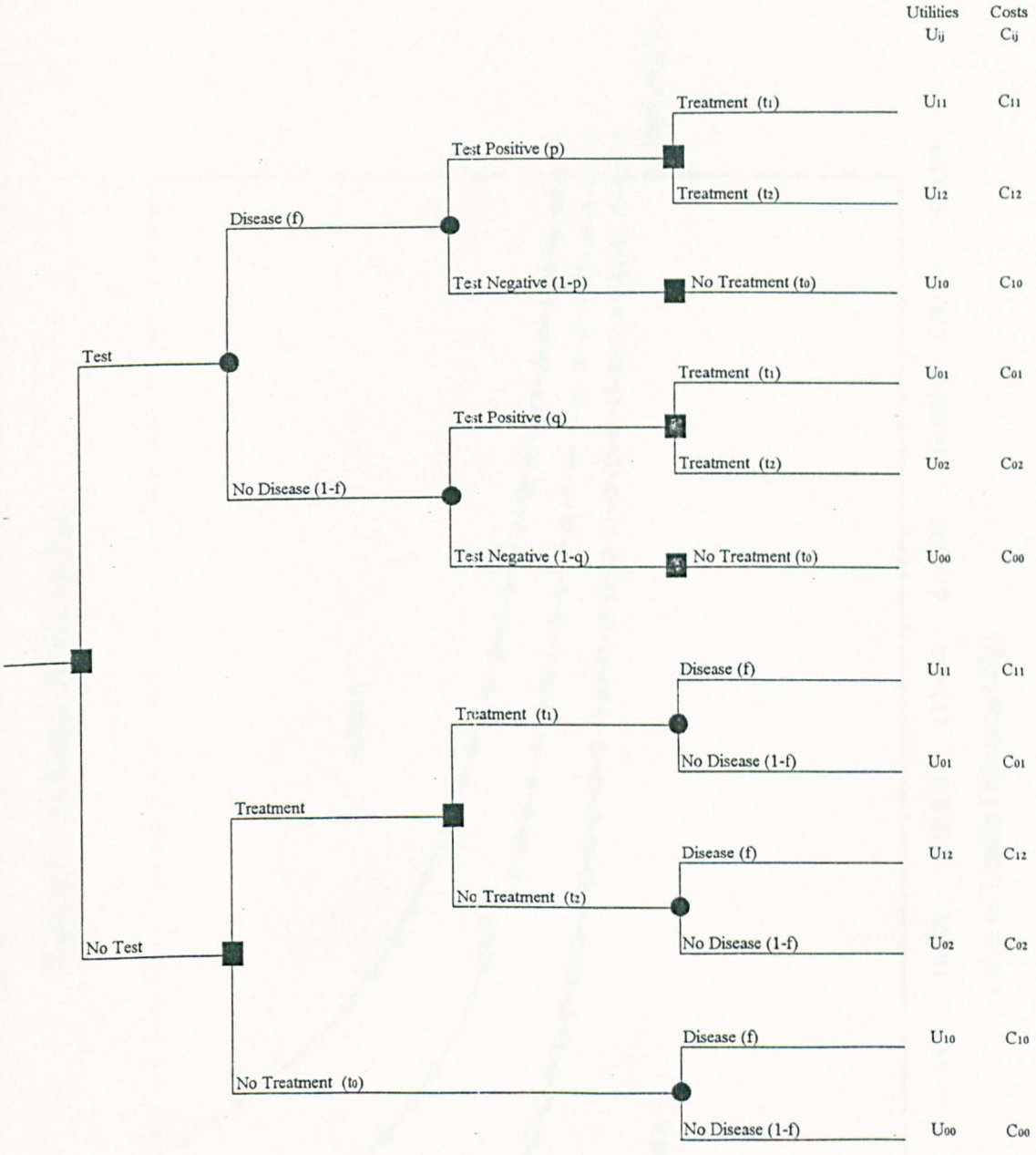


Figure 3.2.2 Treatment Thresholds (f_{jk})

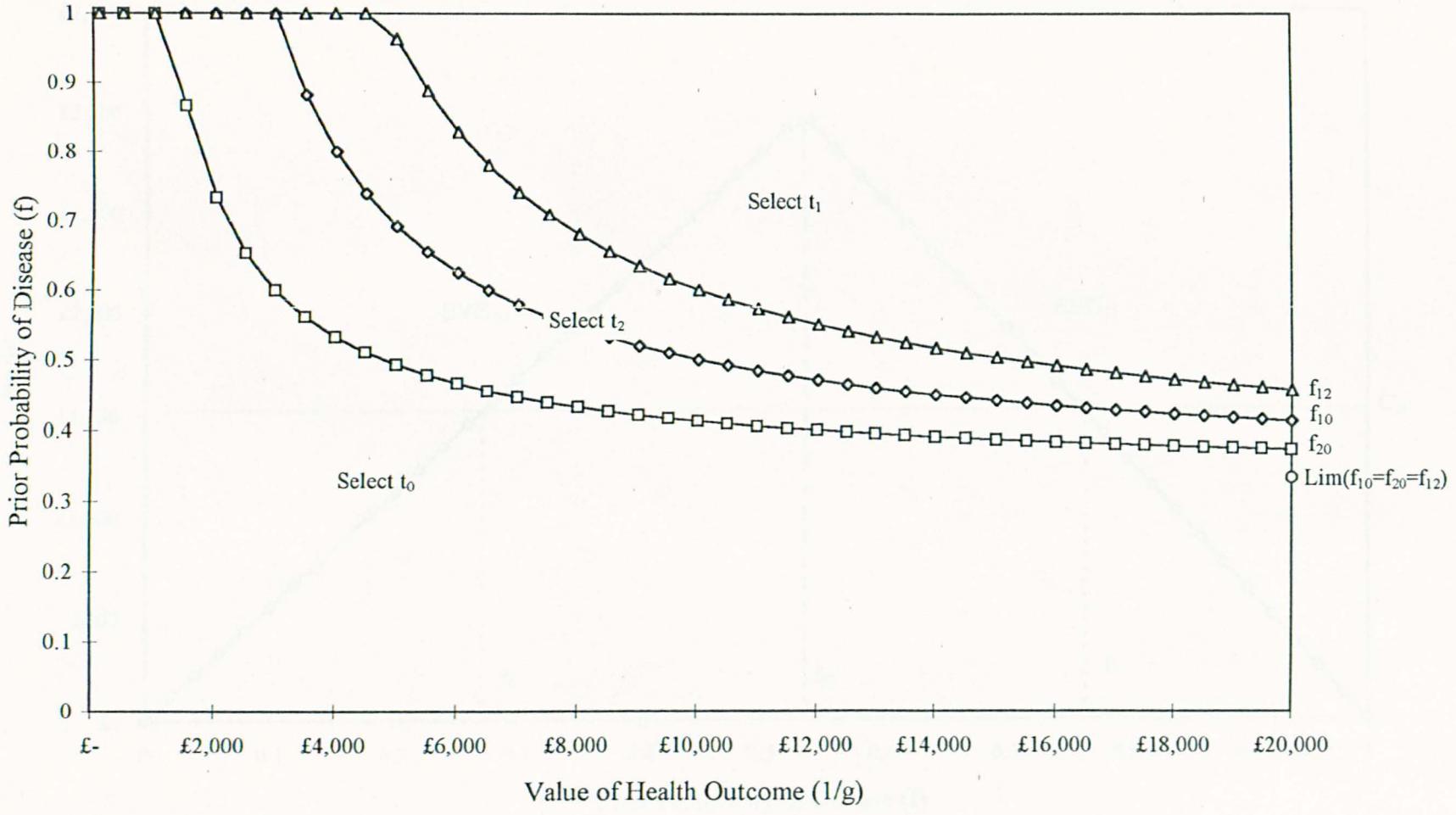


Figure 3.2.3a Expected Value of Perfect Information (EVPI) (1/g=£4,000)

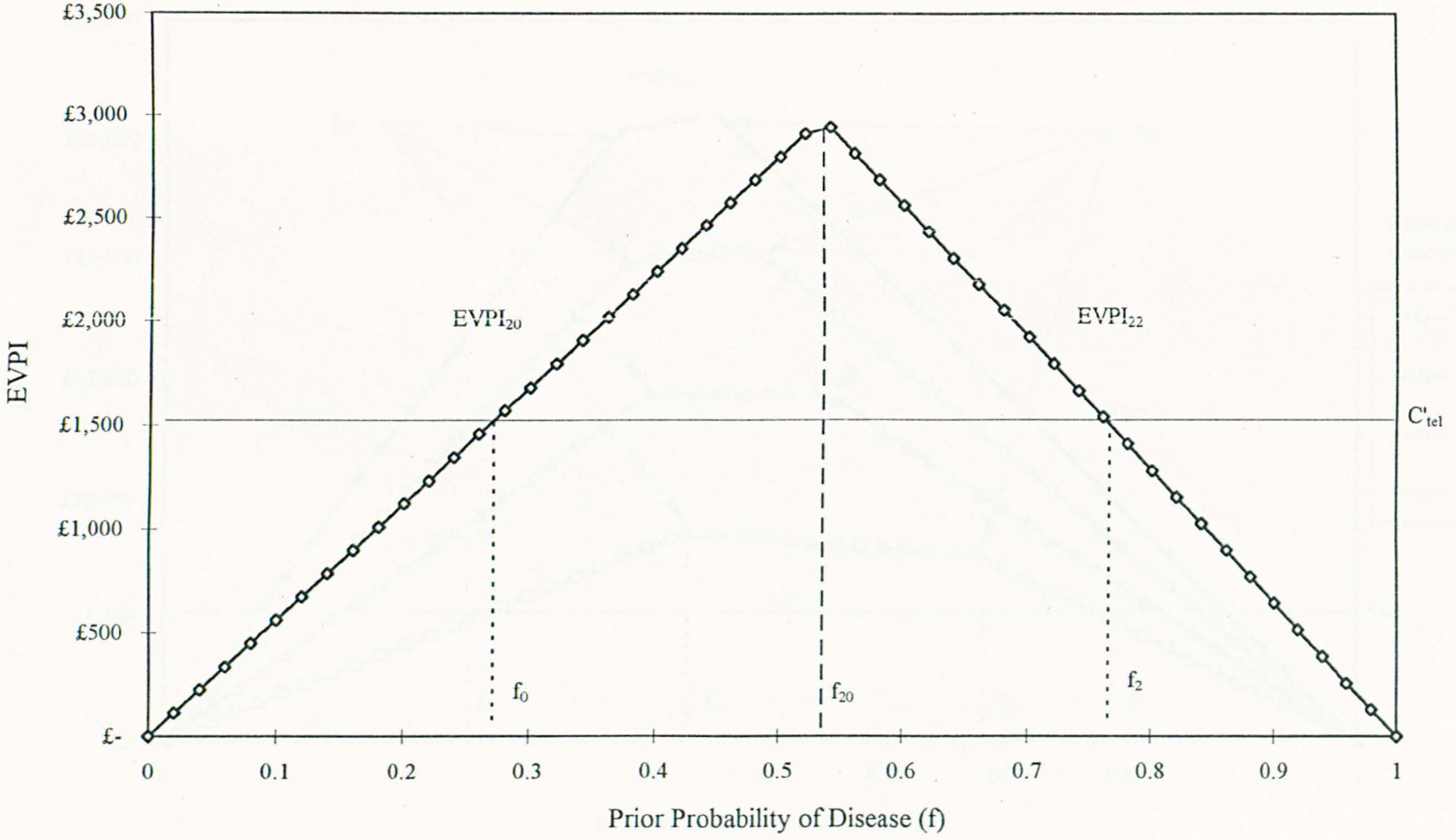
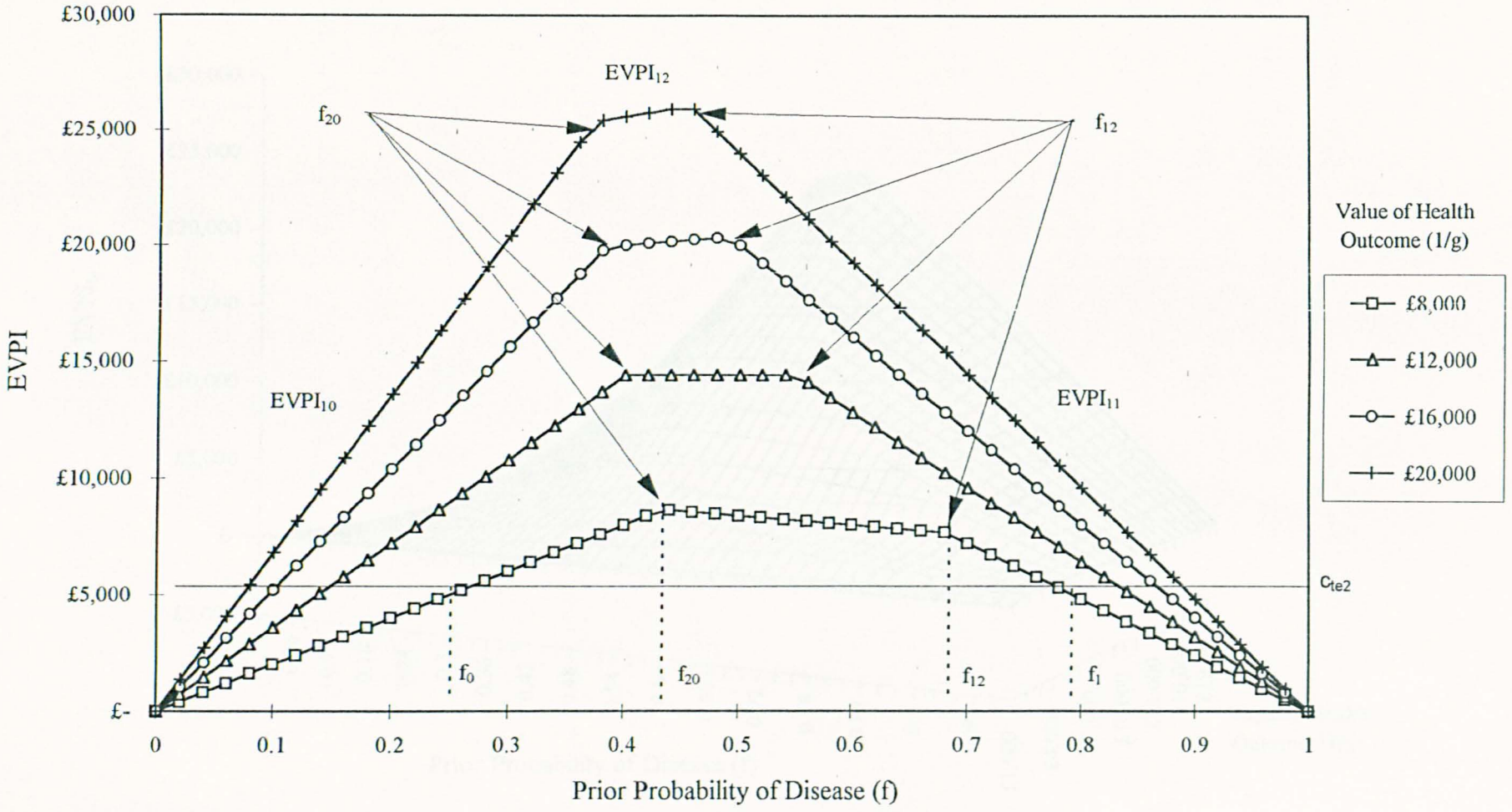


Figure 3.2.3b Expected Value of Perfect Information ($EVPI_{hj}$)



Figue 3.2.3c Expected Value of Perfect Information ($EVPI_{hj}$)

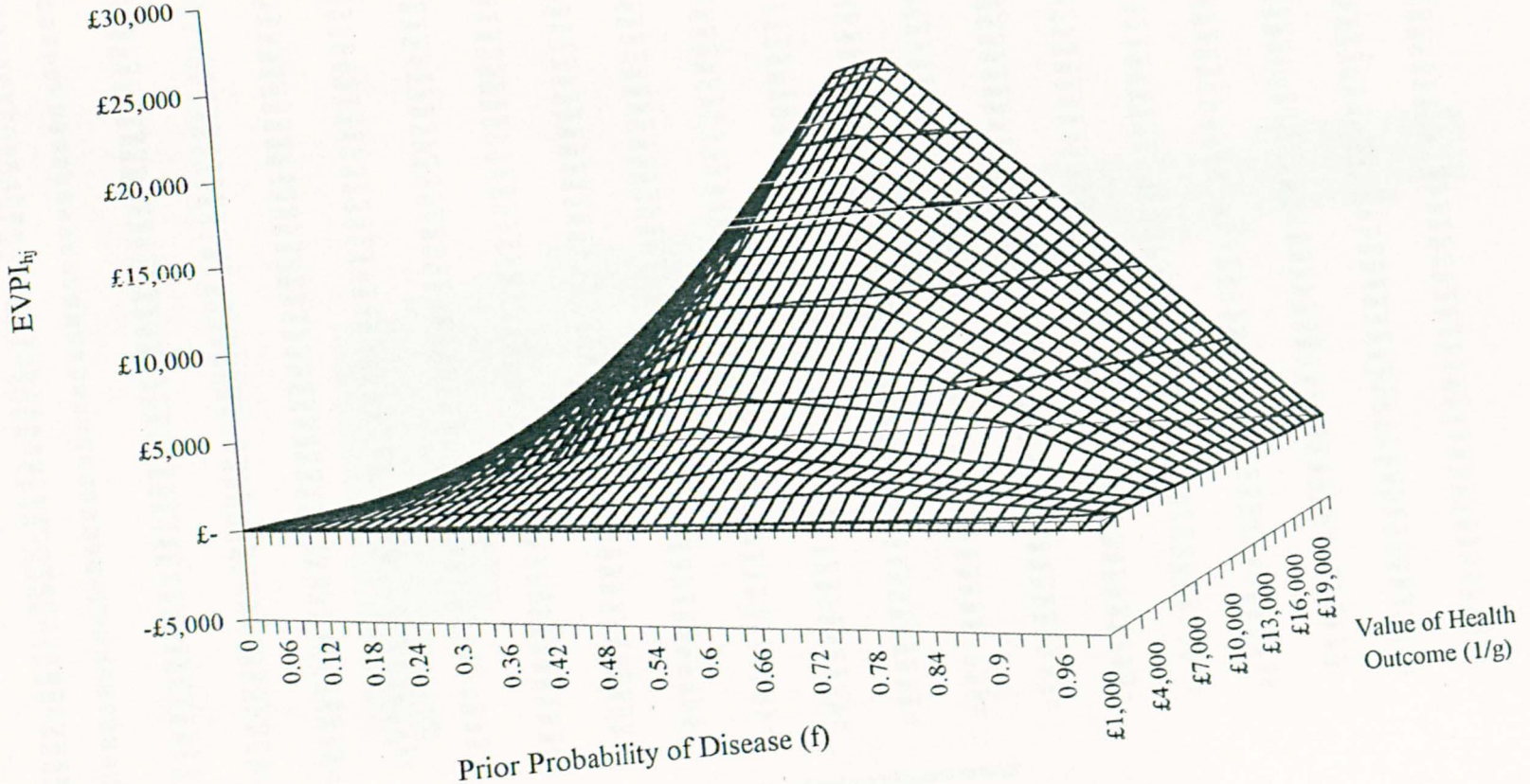


Table 3.2.2 Optimal Strategies with Perfect Information

Prior probability of disease (f)	Value of health outcome (1/g) £ 1,000	£ 2,000	£ 3,000	£ 4,000	£ 5,000	£ 6,000	£ 7,000	£ 8,000	£ 9,000	£ 10,000	£ 11,000	£ 12,000	£ 13,000	£ 14,000	£ 15,000	£ 16,000	£ 17,000	£ 18,000	£ 19,000	£ 20,000
0	02	02	02	02	01	01	01	01	01	01	01	01	01	01	01	01	01	01	01	01
0.02	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.04	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.06	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.08	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.1	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.12	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.14	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.16	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.18	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.2	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.22	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.24	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.26	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.28	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.3	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.32	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.34	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.36	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.38	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	2te1	2te1	2te1
0.4	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1
0.42	02	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1
0.44	02	0te2	0te2	0te2	0te1	0te1	0te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1
0.46	02	0te2	0te2	0te2	0te1	0te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1
0.48	02	0te2	0te2	0te2	0te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1
0.5	02	0te2	0te2	0te2	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.52	02	0te2	0te2	0te2	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.54	02	0te2	0te2	2te2	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.56	02	0te2	0te2	2te2	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.58	02	0te2	0te2	2te2	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.6	02	0te2	0te2	2te2	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.62	02	0te2	2te2	2te2	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.64	02	0te2	2te2	2te2	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.66	02	0te2	2te2	2te2	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.68	02	0te2	2te2	2te2	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.7	02	0te2	2te2	2te2	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.72	02	0te2	2te2	2te2	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.74	02	2te2	2te2	2te2	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.76	02	2te2	2te2	2te2	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.78	02	2te2	2te2	2te2	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.8	02	2te2	2te2	2te2	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.82	02	2te2	2te2	2te2	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.84	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.86	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.88	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.9	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.92	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.94	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.96	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.98	02	2te2	2te2	2te2	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
1	02	22	22	22	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11

2te1 indicates that treatment t2 is the fallback, testing is optimal, and treatment t1 is the optimal test treatment strategy

02 indicates that t0 is the fallback, no testing is optimal, and treatment t2 is the optimal test treatment strategy

Figure 3.2.4a Expected Value of Clinical Information ($EVCI_{hj}$) (1/g=£4,000)

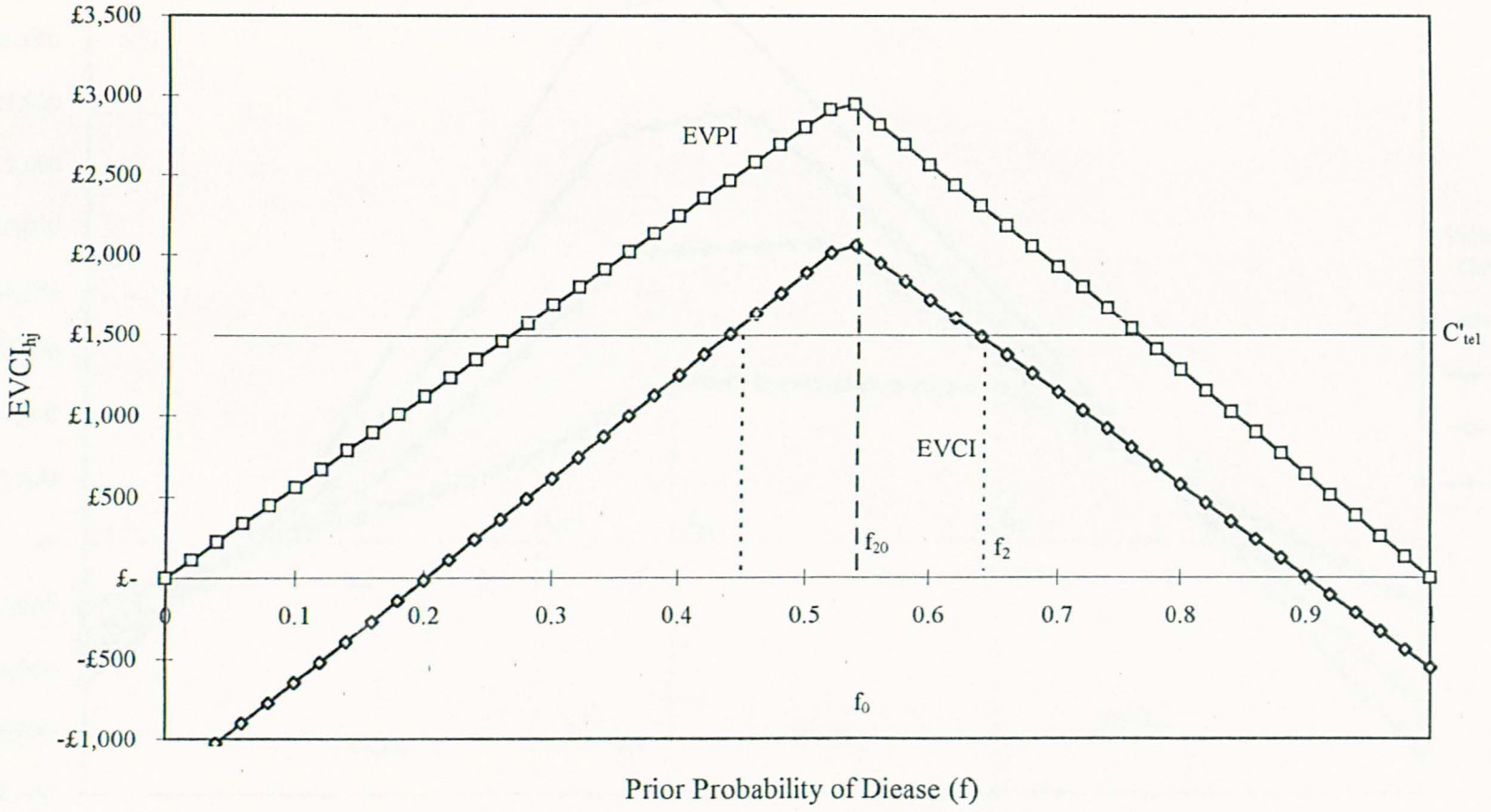


Figure 3.2.4b Expected Value of Clinical Information ($EVCI_hj$)

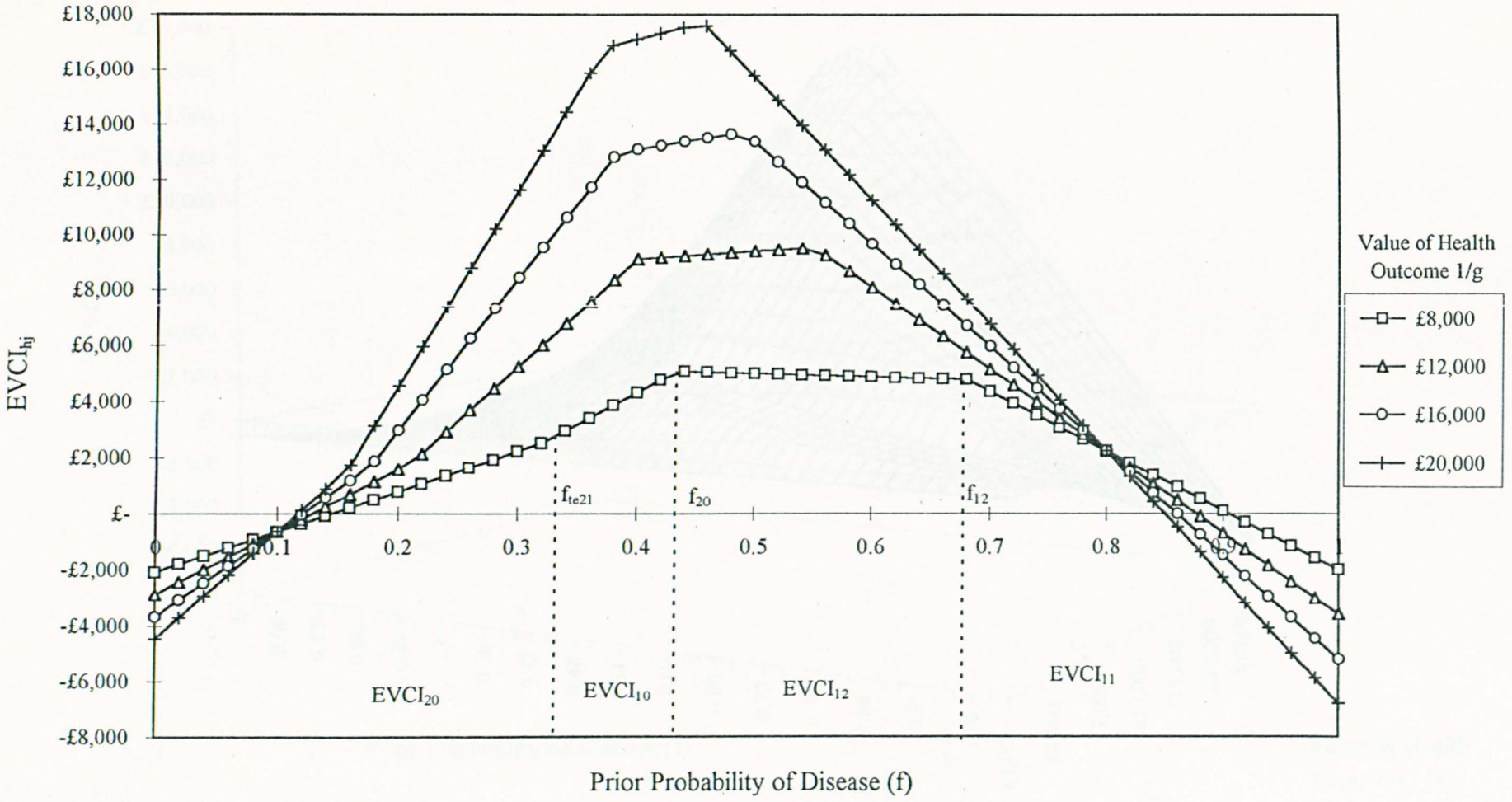


Figure 3.2.4c Expected Value of Clinical Information ($EVCI_{hj}$)

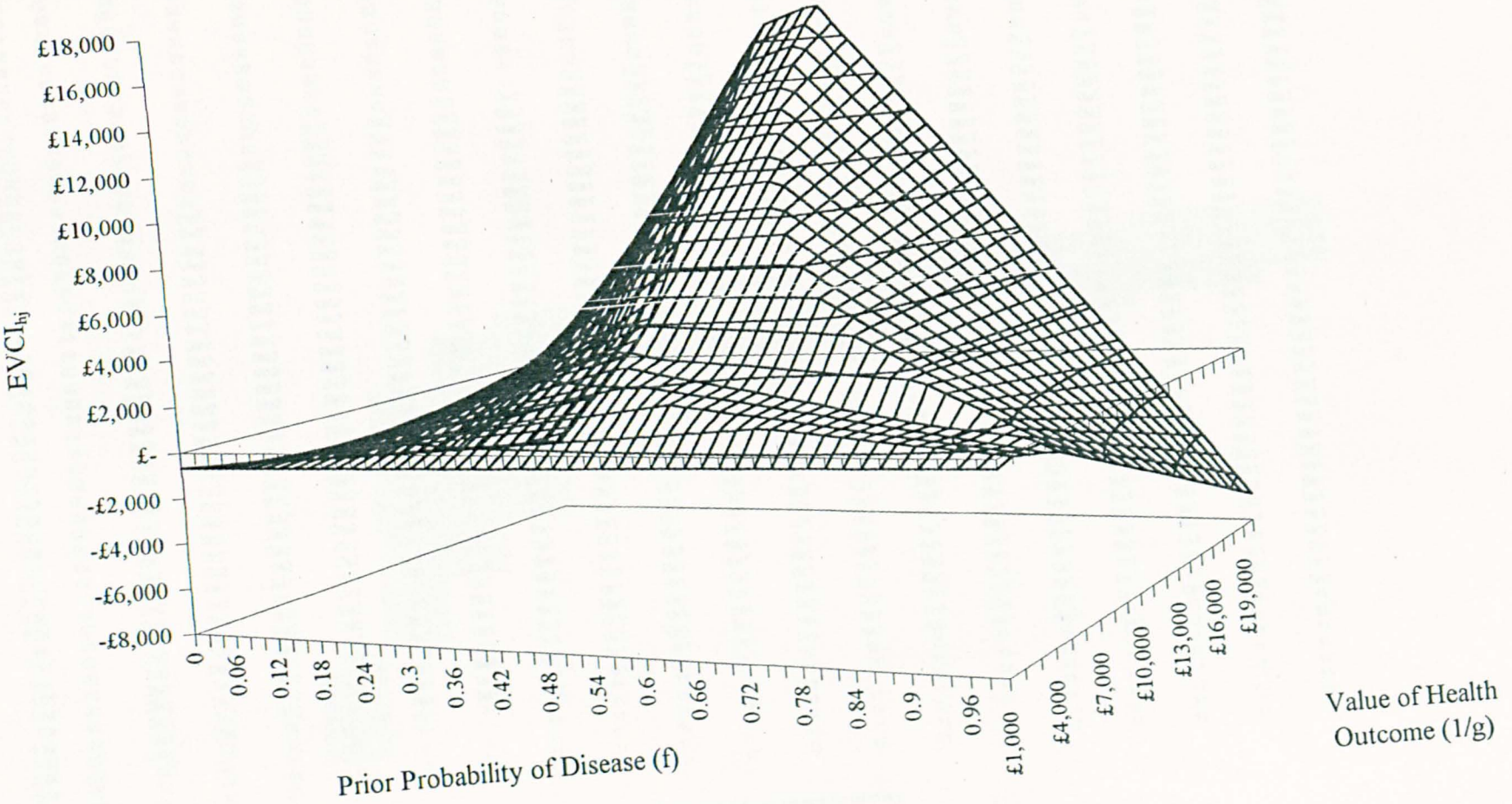


Table 3-2-3 Optimal Strategies with Clinical Information

Prior Probability of Disease (f)	Value of Health Outcome (1/g) £ 1,000	£ 2,000	£ 3,000	£ 4,000	£ 5,000	£ 6,000	£ 7,000	£ 8,000	£ 9,000	£ 10,000	£ 11,000	£ 12,000	£ 13,000	£ 14,000	£ 15,000	£ 16,000	£ 17,000	£ 18,000	£ 19,000	£ 20,000
0	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02
0.02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02
0.04	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02
0.06	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02
0.08	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02
0.1	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02
0.12	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	02	0te2	0te2	0te2
0.14	02	02	02	02	02	02	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2
0.16	02	02	02	02	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te1
0.18	02	02	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1
0.2	02	02	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.22	02	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.24	02	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.26	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.28	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.3	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.32	02	02	0te2	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.34	02	02	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.36	02	02	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1
0.38	02	0te2	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	0te1	2te1	2te1	2te1
0.4	02	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	0te1	0te1	0te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1
0.42	02	0te2	0te2	0te2	0te2	0te2	0te1	0te1	0te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1
0.44	02	0te2	0te2	0te2	0te2	0te2	0te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1
0.46	02	0te2	0te2	0te2	0te2	0te2	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1
0.48	02	0te2	0te2	0te2	0te2	2te2	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1
0.5	02	0te2	0te2	0te2	2te2	2te2	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1
0.52	02	0te2	0te2	0te2	2te2	2te1	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.54	02	0te2	0te2	2te2	2te2	2te1	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.56	02	0te2	0te2	2te2	2te2	2te1	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.58	02	0te2	0te2	2te2	2te2	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.6	02	0te2	0te2	2te2	2te2	2te1	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.62	02	0te2	2te2	2te2	2te2	2te1	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.64	02	0te2	2te2	2te2	2te2	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.66	02	0te2	2te2	2te2	2te2	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.68	02	0te2	2te2	2te2	2te2	2te1	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.7	02	0te2	2te2	2te2	2te2	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.72	02	0te2	2te2	2te2	2te2	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.74	02	2te2	2te2	2te2	2te2	2te1	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.76	02	2te2	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.78	02	2te2	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.8	02	2te2	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.82	02	2te2	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.84	02	2te2	2te2	2te2	2te2	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1
0.86	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	1te1	11	11	11	11
0.88	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	1te1	1te1	1te1	11	11	11	11	11	11	11	11	11
0.9	02	2te2	2te2	2te2	2te1	1te1	1te1	1te1	11	11	11	11	11	11	11	11	11	11	11	11
0.92	02	2te2	2te2	22	21	1te1	11	11	11	11	11	11	11	11	11	11	11	11	11	11
0.94	02	2te2	22	22	21	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
0.96	02	22	22	22	21	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
0.98	02	22	22	22	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
1	02	22	22	22	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11

2te1 indicates that treatment t2 is the fallback, testing is optimal, and treatment t1 is the optimal test treatment strategy

02 indicates that t0 is the fallback, no testing is optimal, and treatment t2 is the optimal test treatment strategy

Figure 3.3.1a Errors at the First Hurdle (when CCER=£4,000)

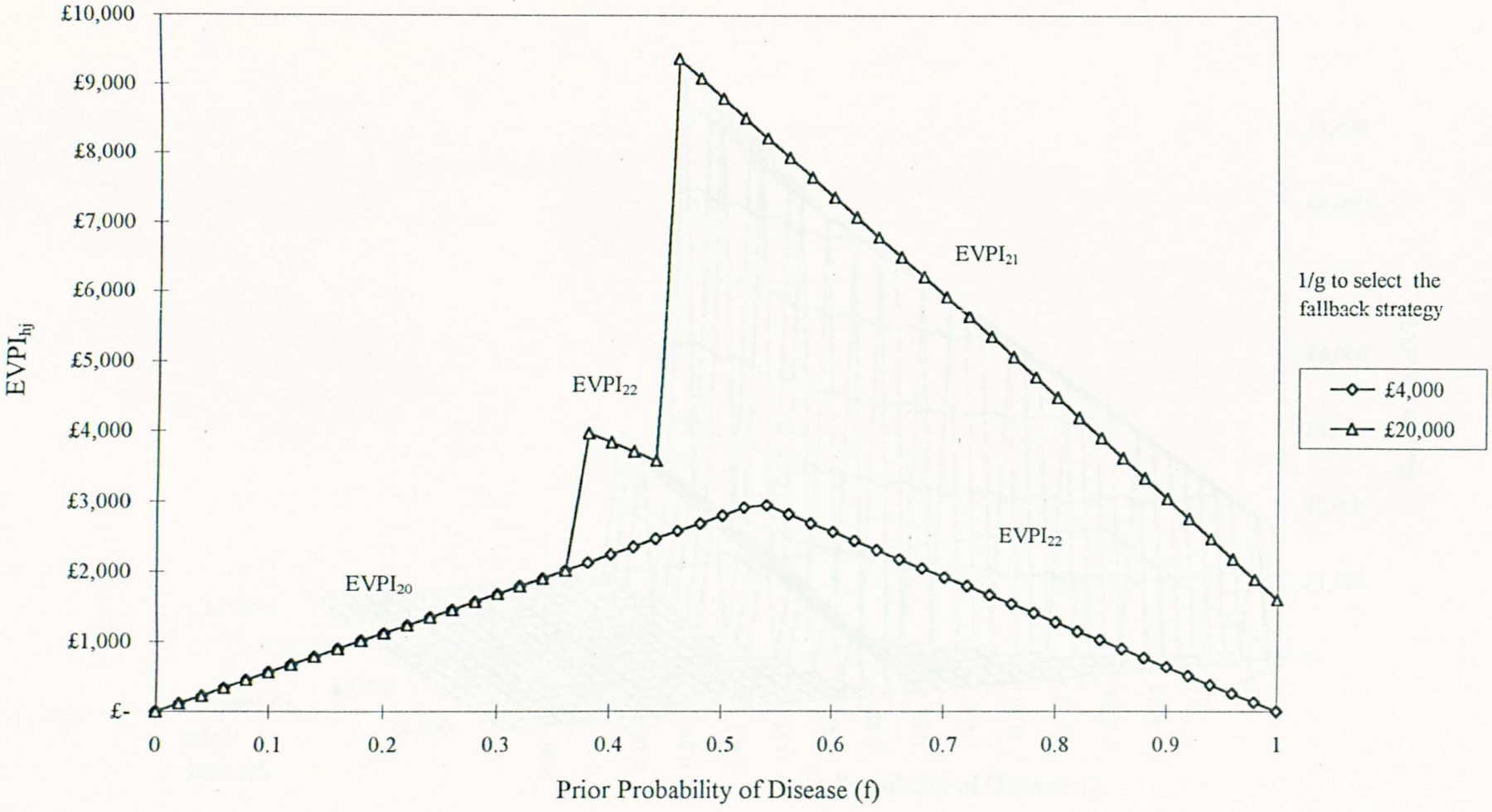


Figure 3.3.1b Errors at the First Hurdle (when 1/g>CCER=£4,000)

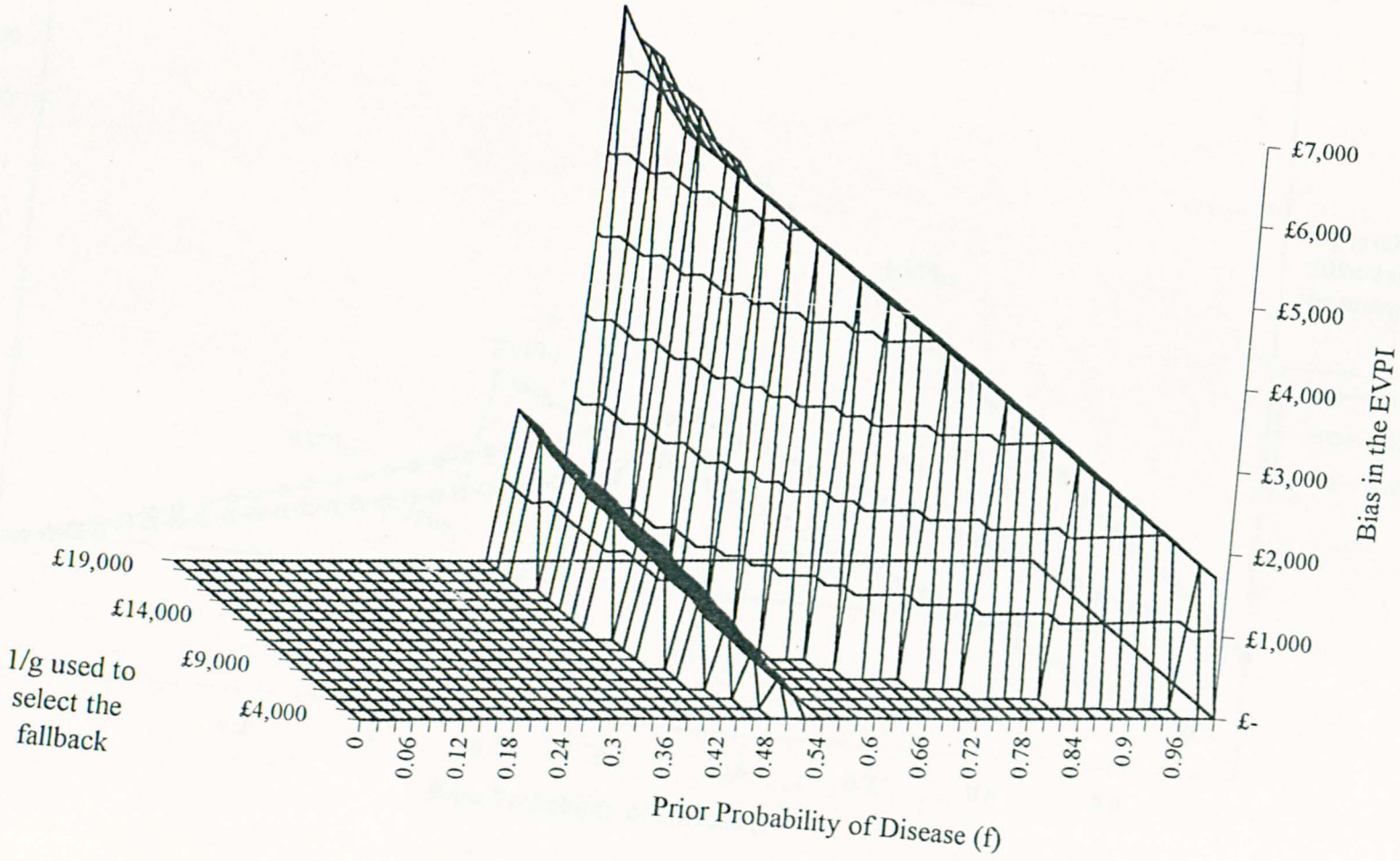


Figure 3.3.2a Errors at the First Hurdle (when CCER=£4,000)

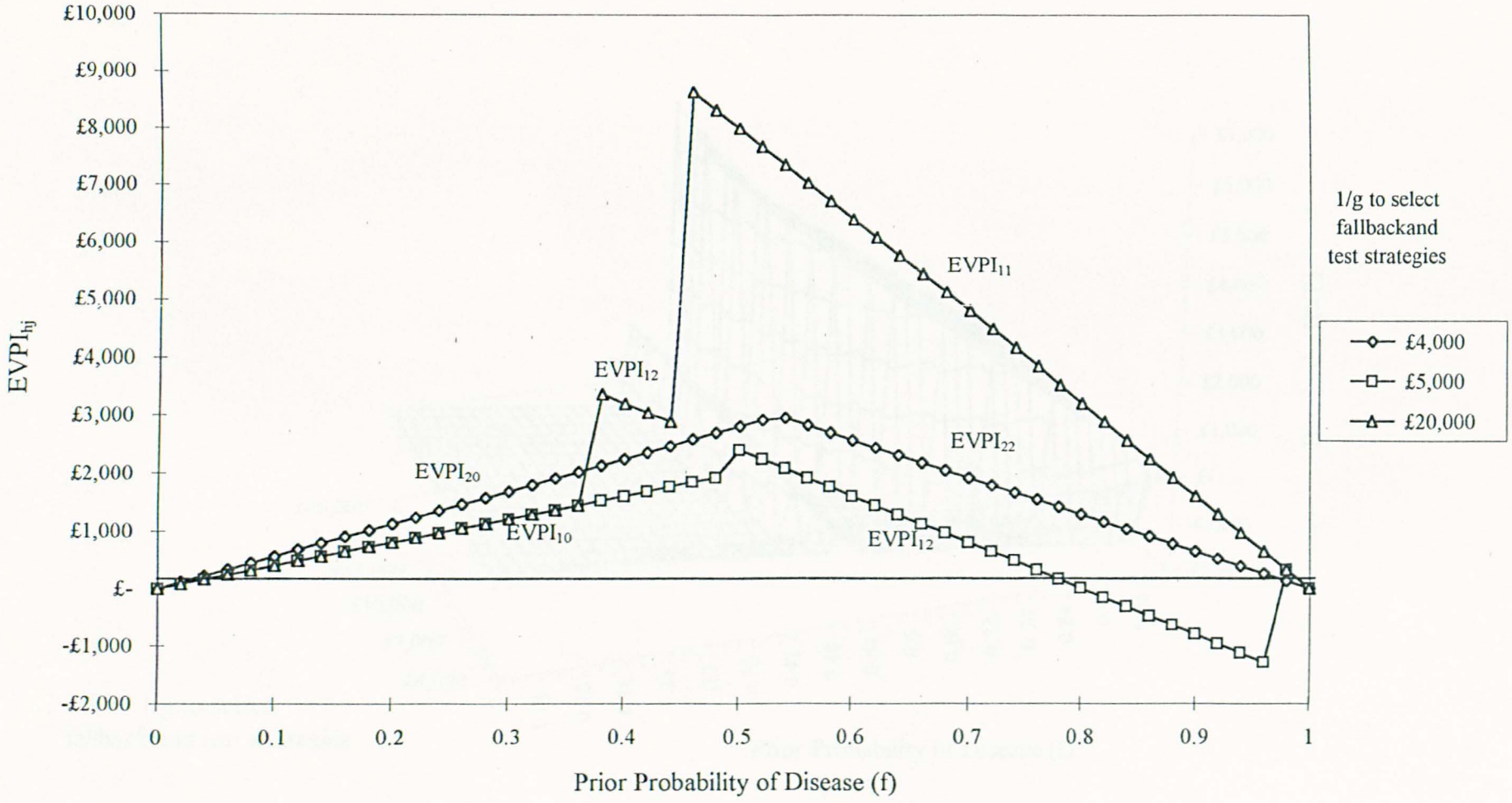


Figure 3.3.2b Errors at the first hurdle (when 1/g>CCER)

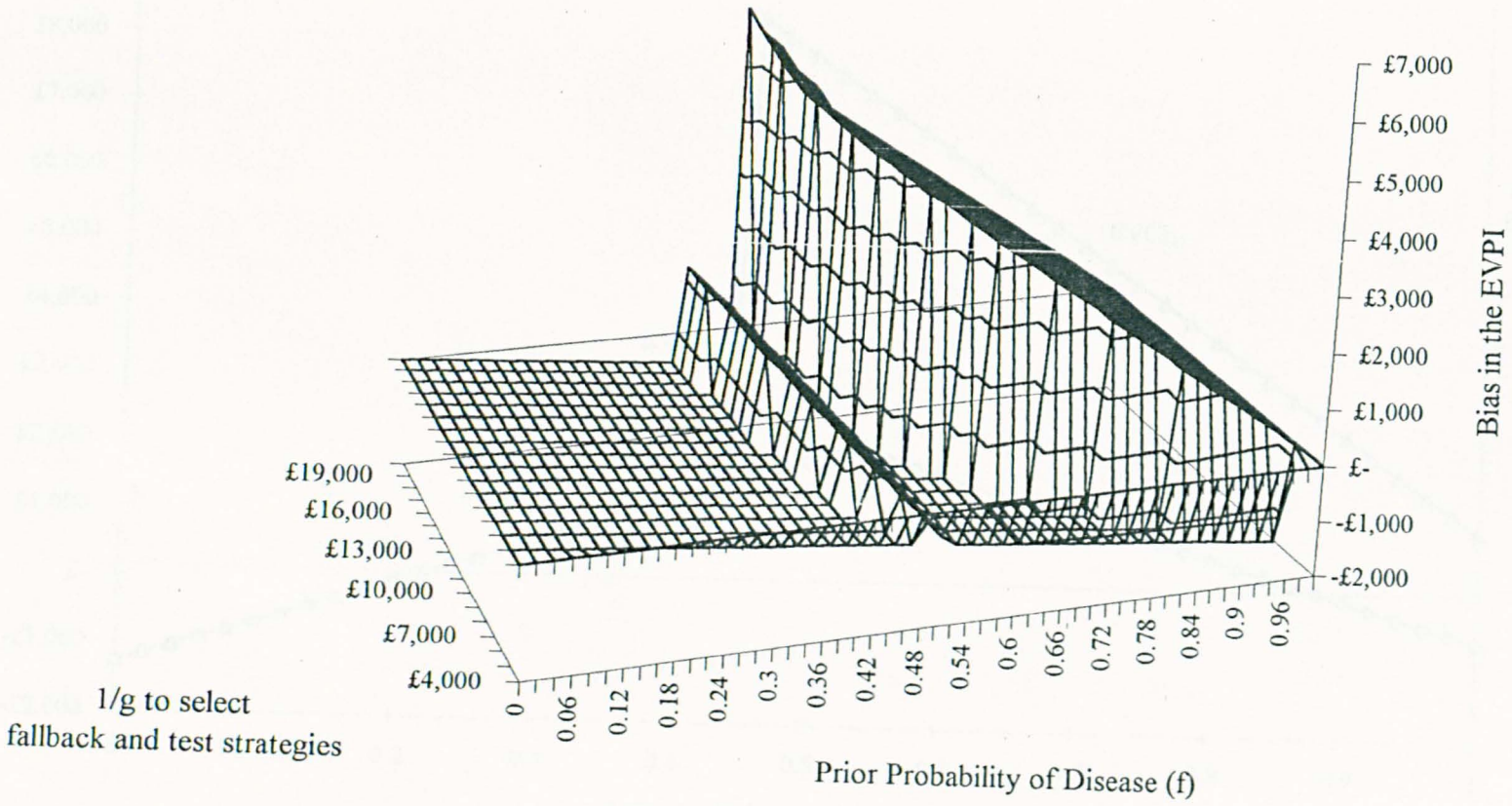


Figure 3.3.3 Errors at the Second Hurdle (when CCER=£4,000)

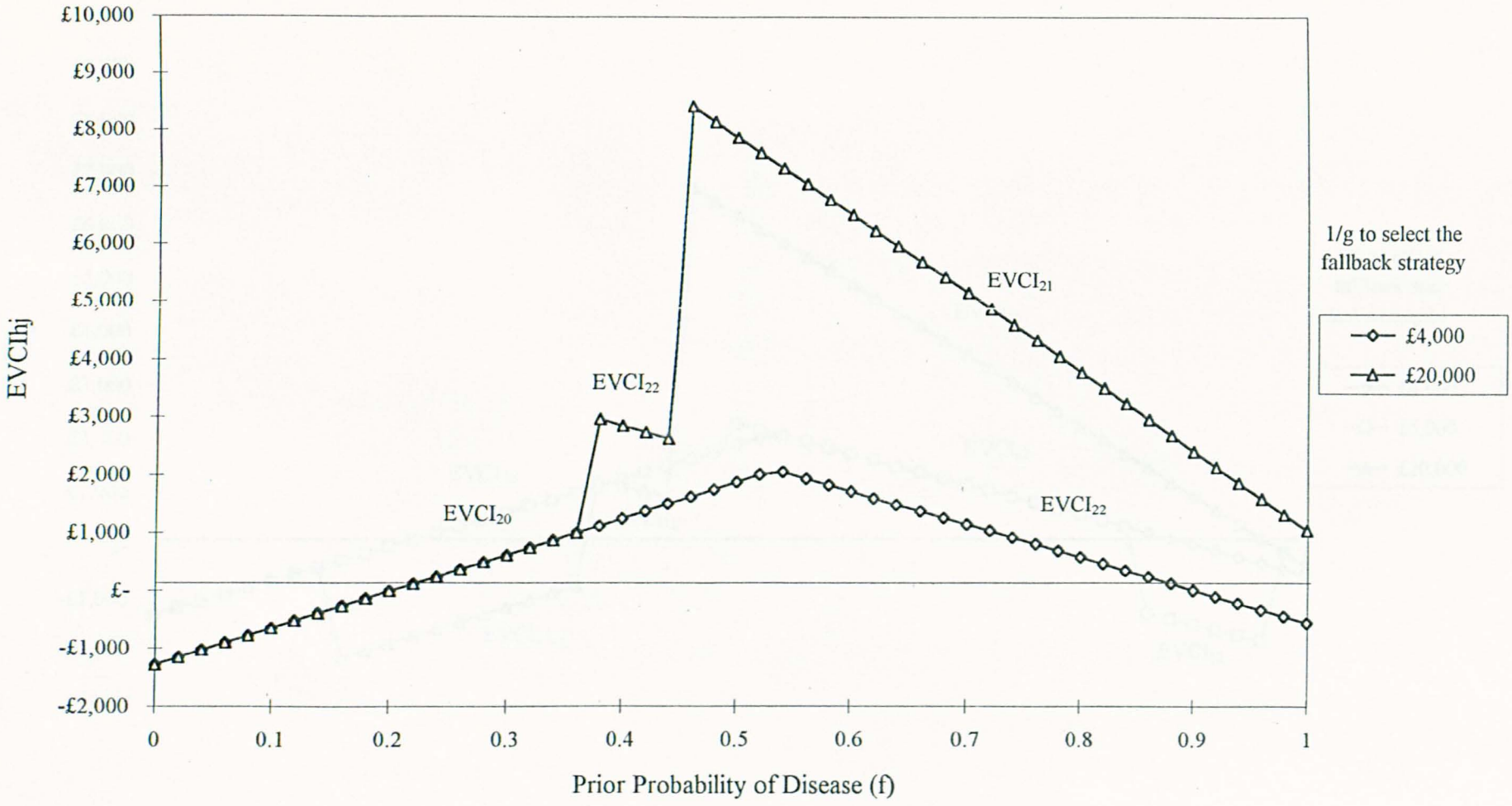


Figure 3.3.4a Errors at the Second Hurdle (when CCER=£4,000)

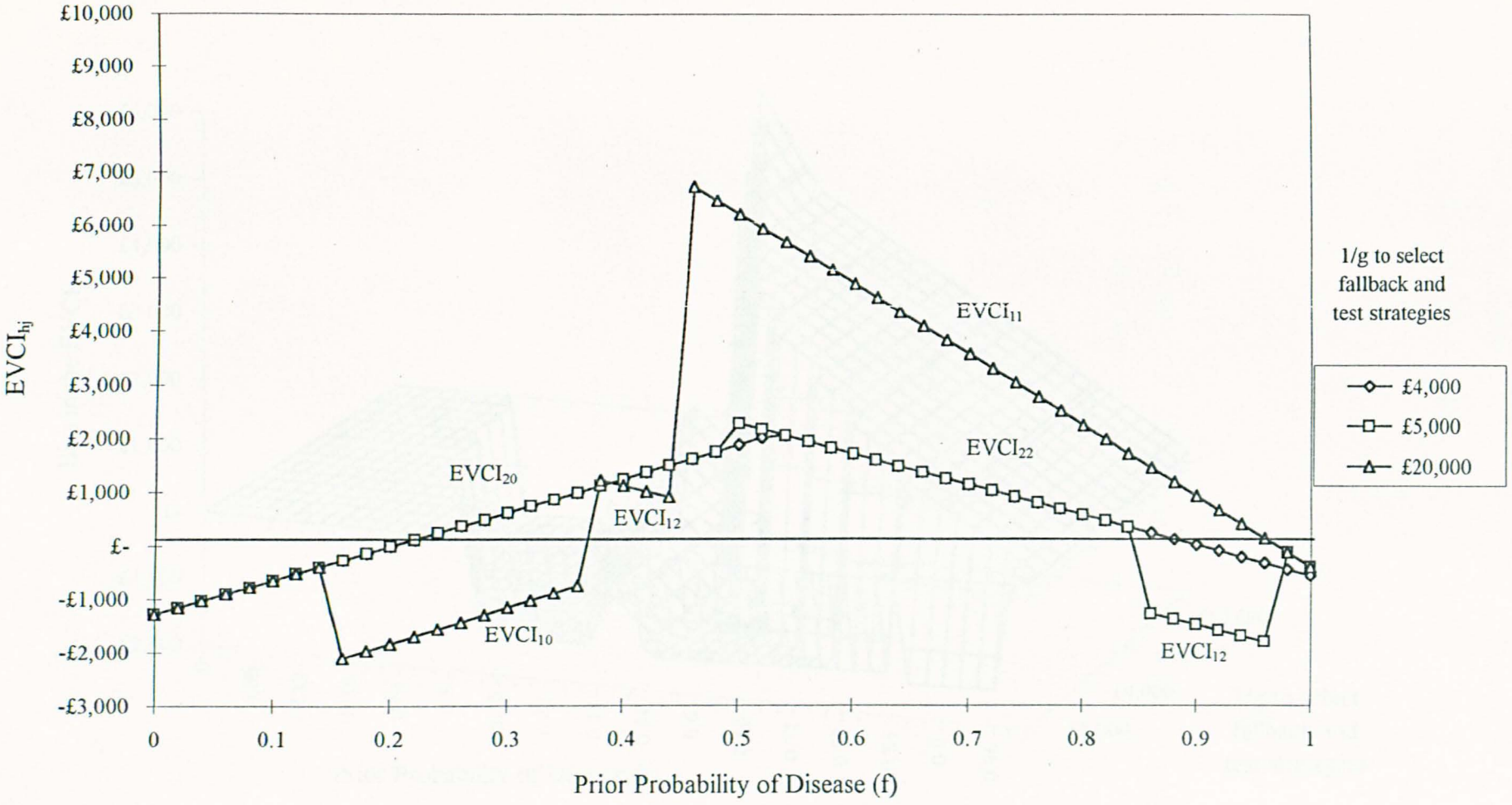
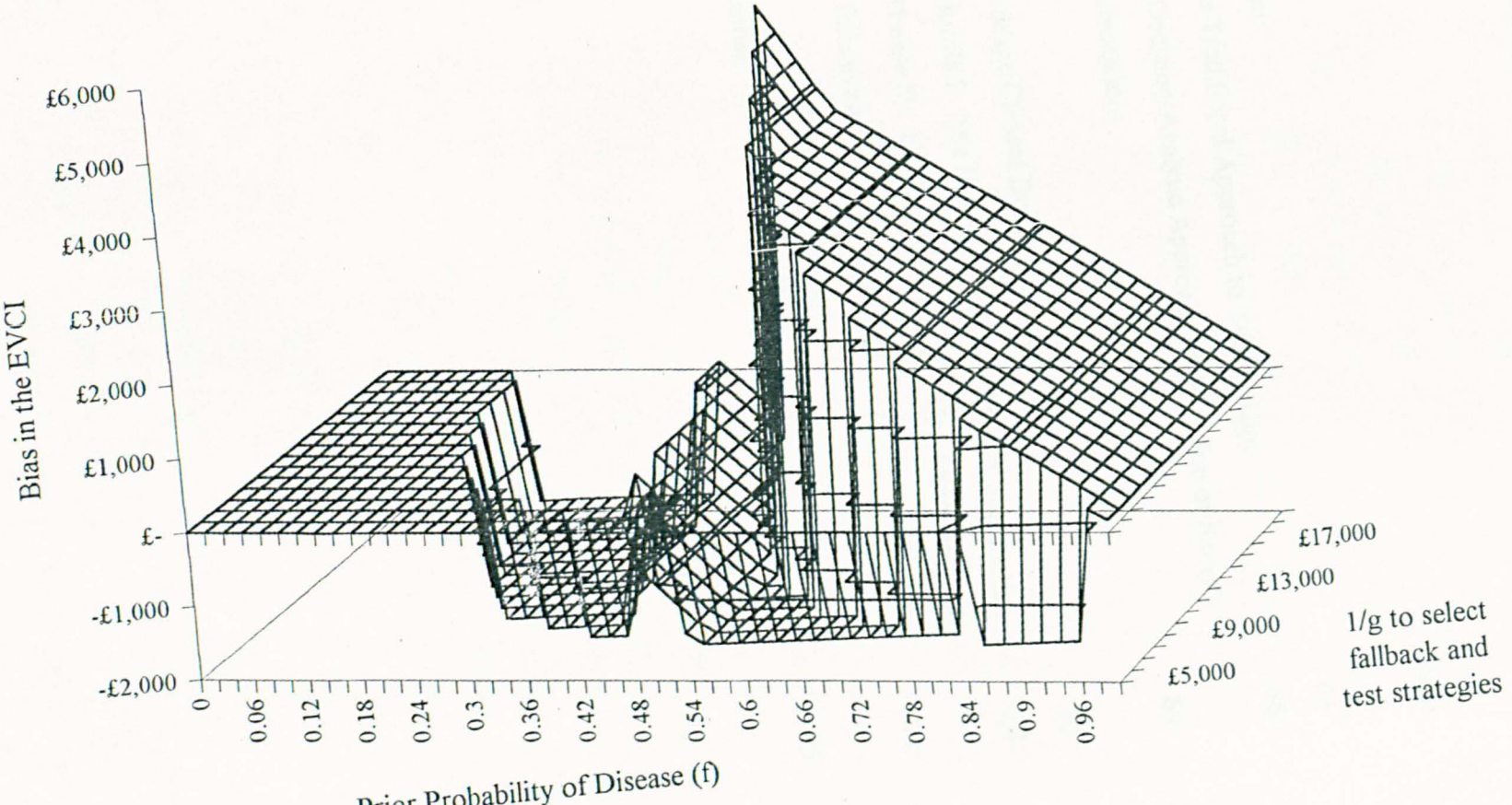


Figure 3.3.4b Errors at the Second Hurdle (when 1/g>CCER=£4,000)



Chapter 4 An Economic Approach to Clinical Trial Design and Research Priority-Setting

Contents: Page

4.1 Introduction

The analysis in the previous chapter demonstrated that a clinical trial may be unavoidable even in the evaluation of a non invasive diagnostic technology. This poses the problems of allocative efficiency across clinical research and technical efficiency in research design which were raised at the end of chapter 3. In this chapter it is argued that the traditional approach to clinical trial design is inconsistent with concepts of efficiency, leads to either infinite or arbitrary sample sizes, and cannot address the issues of allocative or technical efficiency in clinical research.

The methods developed in this chapter address these problems by constructing two hurdles that proposed research must overcome before it can be considered cost-effective. The first hurdle asks if the cost of proposed research exceeds the maximum possible benefits. If the cost does not exceed the maximum benefit then it is potentially cost-effective. Whether the proposed research is actually cost-effective can be established by constructing the second hurdle which explicitly considers the marginal cost and marginal benefits of sample information. The second hurdle ensures that the research is conducted at the technically efficient scale and provides a measure of the net present value of the proposed research. This approach is illustrated by considering the simple single-stage fallback treatment decision which was discussed in chapter 2. The approach is generalised to the more complex two and four-stage test/treatment decisions in chapter 5.

4.1.1 The Traditional Approach to Trial Design

The problems encountered when running an economic evaluation alongside a clinical trial are well documented [1,21, 42, 43, 45]. However, more fundamental is the fact that the traditional approach to the design of pragmatic clinical trials is inconsistent with concepts of efficiency, because an infinite value is implicitly

placed on the benefits of sample information. Furthermore, the traditional approach does not directly address the decision problem faced by clinicians, [122, 123] and cannot incorporate prior information explicitly and consistently. The purpose of this chapter is to show how the principles of economic evaluation can be used to develop a consistent approach to trial design and research priority-setting.

In the traditional approach (assuming a fixed sample design, where all the results are available at the same time at the end of the trial), the key design issue is the number of patients to recruit. Optimal sample size (n*) is determined by the reference improvement (δ_r), the working significance level (α), the power of the test ($1-\beta$), and the variance of population differences in effectiveness between interventions (σ^2).

$$n^* = (\gamma / e)^2 \qquad e = \text{effect size} = \delta_r / \sigma \qquad \gamma = f(\beta, \alpha) \qquad 4.1.1$$

Sample size is very sensitive to the reference improvement, and if the selection of δ_r is not well defined or is chosen in an arbitrary way, then sample size will also be arbitrary. The clinical reference improvement has been defined as the smallest worthwhile difference in effectiveness [84]. Very small improvements in effectiveness should be worthwhile, but as δ_r approaches zero, sample size tends to infinity. The justification given for δ_r substantially greater than zero (and finite sample size) is that practitioners require a large clinical difference before they can be convinced that the experimental treatment will improve health outcome, and that incurring the additional costs of the experimental treatment will be worthwhile.

These are two separate issues. If practitioners are sceptical of improved effectiveness then an appropriate response is to increase the level of significance and power by increasing sample size at each level of δ_r. The minimum improvement in effectiveness required to offset the additional costs of a new treatment can be established by rescaling the incremental costs into health outcome using the critical effectiveness-cost ratio [86,108, 140] (this is the effectiveness-

cost ratio of the marginal project which will be displaced by the new treatment). The costs of treatment are now included in the analysis and the appropriate reference improvement is the minimum improvement in efficiency (rather than effectiveness) that is considered worthwhile. But a reference improvement in effectiveness which would just offset the incremental cost will lead to a reference improvement in efficiency of zero, and sample size will again tend to infinity.

Although the costs of treatment can be taken into account when designing a clinical trial there is no consideration of the marginal cost of obtaining sample information. Any improvement in either effectiveness or efficiency will be worth detecting if it is assumed that the marginal cost of detecting such a difference is zero. An infinite value is implicitly placed on the benefits of sample information, leading to either infinite or arbitrary sample sizes. The approach is inconsistent with concepts of efficiency and with the original rationale for considering the cost of treatment alongside the trial.

It has been recognised for some time that the traditional approach to trial design does not directly address the decision problem facing clinicians. A minimum combination of α and β is stipulated which should be applied in all clinical settings irrespective of the relative costs of type I and type II errors. Schwartz and Lellouch (1967) have argued that in a pragmatic clinical trial type I errors impose no costs and are irrelevant [121,122]. A type I error will be made if the clinician concludes that there is a difference between treatments when no difference actually exists. In this case, since the treatments are equivalent, it does not matter which treatment is chosen, and the level of significance is irrelevant. It is the probability of making the wrong decision by concluding that the experimental treatment is superior to the control when the reverse is true (the probability of a type III error) which should be the issue of concern.

Finally, the traditional approach to clinical trial design is founded on the view that probability represents the relative frequency of repeated events. There is no

explicit role for prior information, although in practice it is implicit at each stage of design [14, 31] (including the choice of alternatives to be compared; and the selection of δr, α, and β), and during data monitoring in sequential trials [52, 53, 54, 55, 59, 76, 113, 126]. Because the role of prior information is not explicit, it cannot be handled consistently and is not open to criticism, alternative formulation and empirical falsification.

4.1.2 A Decision-Analytic Approach to the Value of Information

An approach to trial design is required which directly addresses the decision problem faced by clinicians; which takes account of the marginal costs and marginal benefits of sample information; which uses all of the information available prior to prospective research and can address the issue of allocative efficiency across clinical research; and technical efficiency in research design.

The decision-analytic approach presented in this chapter combines a Bayesian view of probability with a framework for decision-making which explicitly takes into account the consequences of making a type III error. The approach abandons traditional significance testing, confidence intervals,[58, 114] and their Bayesian counterparts [60, 127, 128] in favour of minimising the expected costs of making the wrong decision.

There have been a number of contributions to the literature which have proposed a decision-analytic approach to sequential clinical trial design [5, 6, 29, 30, 149]. These contributions have focused on clinical measures of efficacy rather than efficiency, and have been criticised [7, 27] because predicted sample sizes may become very large. There have been a number of contributions which have proposed data-dependent allocation, normal loss functions, and an explicit patient horizon to establish the optimal allocation of patients as a sequential trial progresses [7, 8, 9, 12, 13]. These approaches focus on clinical outcomes in sequential trial designs without

explicit consideration of economic criteria or the resource costs of obtaining sample information.

The following example illustrates the use of a decision-analytic approach when considering the more fundamental problem of a fixed sample design where the problems of optimal sequential allocation and optimal stopping do not arise [116, 115, 121]. It explicitly includes economic criteria at all stages of the design, including the costs of treatment and sampling,[108] and the timing of costs and benefits. This approach uses the same decision rules for cost-effectiveness and efficiency which are increasingly used to set priorities in service provision [146]. The objective is to promote consistency in decision-making and priority-setting between research and service provision.

4.2 A Single-Stage Clinical Decision Problem

The approach is illustrated using the simple numerical example of the fallback treatment decision considered in chapter 2. The example considers a fixed sample design of a pragmatic or phase III clinical trial for this single stage decision problem. The clinician faces a choice between two alternatives (t_j, j=0, 1) where t_0 is current practice (no treatment), and t_1 can be regarded as the experimental treatment for a well-defined patient population. This single-stage problem is illustrated in figure 4.2.1 and is identical the problem which was illustrated in figure 2.2.1. There are two disease states; no disease (i=0), and disease (i=1), with a prior probability of Pi. The health utilities (outcomes) can be regarded as measures of health related quality of life (U_{ij}), and the resource costs (C_{ij}) also depend on disease state and treatment.

Figure 4.2.1

Table 4.2

The prior mean and variance of the health utilities, probabilities and costs for the example are reported in table 4.2. It is assumed that the utilities and the probabilities are independent and normally distributed so the prior variance (Var_0) of $E(U_j)$:

$$Var_0(E(U_j)) = E(P_1^2).E(U_{1j}^2)-E(P_1)^2.E(U_{1j})^2 + E(P_0^2).E(U_{0j}^2)-E(P_0)^2.E(U_{0j})^2 \qquad 4.2.1a$$

$$E(P_i^2) = E(P_i)^2 + Var_0(P_i)$$

$$E(U_{ij}^2) = E(U_{ij})^2 + Var_0(Uij)$$

To simplify the example further it is assumed that each element of cost C_{ij} and the value of g are known, but expected cost ($E(C_j)$) is normally distributed because of the prior variance of the probability of disease.

$$Var_0(g.E(Cj)) = Var_0(P_1).C_{1j}^2.g^2+Var_0(P_0.).C_{0j}^2.g^2 \qquad 4.2.1b$$

These assumptions can be relaxed without loss of generality.

4.2.1 Hurdle I: The Expected Value of Perfect Information

Without sample information the decision-maker must choose between t_1 and t_0 using only prior information. If prior expected costs are rescaled to units of health outcome (using g, the critical effectiveness-cost ratio),[108, 140] the decision-maker should choose t_1 if the prior incremental net benefit of t_1 (δ_0) is positive.

$$\delta_{0(2)} = (E(U|t_1)-g.E(C|t_1))-(E(U|t_0)-g.E(C|t_0)) \qquad 4.2.2a$$

The prior net benefit can also be rescaled to monetary units (using 1/g, the critical cost-effectiveness ratio) so the decision-maker should choose t_1 when $k_1.\delta_0>0$ (where $k_1=1/g$), and should choose t_0 when $k_0.\delta_0>0$ (where $k_0= -1/g$). The decision-maker will be indifferent between t_1 and t_0 when δ_0 is equal to its break-even value ($\delta_b=0$). An alternative approach is to minimise the expected opportunity loss. Opportunity loss is the difference in incremental net benefit between the best choice and the alternative actually chosen (opportunity loss $=|k_1-k_0|.|\delta_0-\delta_b|= K_t.|\delta_0-\delta_b|$ where $K_t=2/g$). The loss functions for t_0 and t_1 are illustrated in figure 4.2.2.

Figure 4.2.2

The opportunity loss is minimised by choosing t_1 when $\delta_0>\delta_b$, and by choosing t_0 when $\delta_0<\delta_b$. However, the incremental net benefit of t_1 has a prior probability distribution with a prior mean of δ_0 and a prior variance of σ_0^2. Given the assumptions of normality and independence:

$$\sigma_0^2 = Var_0(E(U|t_1))+Var_0(g.E(C|t_1)) + Var_0(E(U|t_0))+Var_0(g.E(C|t_0)) \quad 4.2.2b$$

There is a probability that a decision based on the prior mean will be wrong, and opportunity losses will be incurred. The expected opportunity loss is the expected cost of the uncertainty surrounding the decision problem: this is the Expected Value of Perfect Information (EVPI).

The EVPI is determined by three factors: the slope of the loss function ($K_t = 2/g$), which determines the value of opportunity losses; the distance of the prior mean from break-even ($|\delta_0 - \delta_b|$) and the spread of the prior distribution (σ_0), both of which determine the chances of incurring opportunity losses. The expected opportunity loss (EVPI) is calculated based only on prior information and the unit normal loss integral [116, 121].

$$EVPI = K_t.\sigma_0.L(D_0) \quad 4.2.2c$$

$D_0 = (\delta_0 - \delta_b)/\sigma_0$

δ_0 = prior incremental net benefit

σ_0 = prior standard deviation of δ_0

(4.2.2c) gives the EVPI when faced with a choice between t_1 and t_0 for an individual patient. However, a decision-maker will face this same decision problem for a number of patients over a period of time. Given an estimate of the incidence of patients entering the decision problem in figure 1 in each period (h), the population EVPI can be calculated. The incidence in each period can be discounted at rate r to provide the present value of the population EVPI.

$$\text{Population EVPI} = \sum_{h=1}^{H} (EVPI.Incidence_{(h)}.1/(1+r)^h) \quad 4.2.2d$$

This is the maximum benefit that could be provided by additional information, and the maximum return to research effort in this area. This gives a method for

focusing research priorities. It can be used to identify those clinical decision problems (or areas of clinical research) where the costs of uncertainty are highest, and where the information from research will be most valuable. If the fixed costs of research are known, the EVPI can be used as an effective hurdle to eliminate proposals (where the costs exceed the EVPI) which will not be cost-effective. The EVPI can also be used in the same way to identify priority areas for scientific reviews and Meta-analysis: clinical decision problems where the costs of uncertainty are greatest derive the most benefit from a review of existing research. The EVPI is a powerful tool for identifying research priorities in support of a move towards evidence-based medicine [49, 119, 120]. Indeed this approach can set the limits to evidence based medicine and provide a framework within which it can be applied consistently [40, 41].

The expected value of perfect information can be calculated for any decision problem based only on prior information, including evidence from previous intervention and observational studies, but it can also include expert judgements. The decision-analytic framework focuses attention on those variables where evidence or judgement is required (in this example on the health outcomes, costs and path probabilities illustrated in figure 4.2.1 and shown in table 4.2). By making prior information and judgements explicit they are open to criticism, alternative formulation and empirical falsification. This is not necessarily the case in input/output models for assessing payback in clinical research [36, 48]. Delphic studies in research foresight which elicit preferences over research priorities [3, 4, 95, 102, 147] use expert judgements which are not open to criticism or empirical testing because assumptions about outcomes, costs, path probabilities and decision rules remain implicit.

Where expert judgements are used to establish the EVPI, the level of confidence in this prior information will be reflected in a higher prior variance. Prior information can be regarded as a quasi sample with a quasi sample size of n_0 (n_0= ratio of population to prior variance), where a smaller n_0 indicates a more sceptical

prior. This index of confidence is used in this example and is reported in table 4.2. The quasi sample size is higher for the utilities and probabilities associated with t_0, and this reflects the assumption that there may be more confidence in the prior information about current practice. This framework makes these judgements explicit and allows prior evidence form a variety of sources to be combined, and handled consistently using Bayes Theorem [67, 116, 121].

Value of Health Outcome

An example of the relationship between the EVPI and the value of 1/g (the critical cost-effectiveness ratio; the value of health outcome) is shown in figure 4.2.3, and this demonstrates the fact that the value of information is crucially dependent on the value of health outcome used to set priorities in service provision.

Figure 4.2.3

The slope of the loss function or loss constant (K_t=2/g) determines the value of opportunity losses when they occur. If the value of health outcome is greater the opportunity costs of making the wrong decision are valued more highly. This suggests a positive relationship between EVPI and value of health outcome.

However the prior variance of δ_0 partly determines the probability of incurring these opportunity losses, and σ_0 will fall as 1/g increases (4.2.1b). The probability of incurring opportunity losses is also determined by $|\delta_0 - \delta_b|$. When the prior cost-effectiveness ratio of t_1 (£6,500) is equal to 1/g the decision maker will be indifferent between t_1 and t_0. At this point $\delta_0 = \delta_b = 0$ and the decision maker is most uncertain. The standardised distance (D_0 in 4.2.2c) is equal to zero in figure 4.2.4a and $L(D_0)$ reaches a maximum in figure 4.2.4b. In this example the EVPI reaches a peak in figure 4.2.3 when 1/g=£6,500 and the clinician is most uncertain about the treatment decision.

Figure 4.2.4a

Figure 4.2.4b

The decision maker will choose t_0 when 1/g <£6,500, an increase in 1/g will reduce $|\delta_0 - \delta_b|$ and D_0 in figure 4.2.4a, and the probability of incurring opportunity losses will increase. Both $L(D_0)$ in figure 4.2.4b and K_t will rise with 1/g, and the EVPI will increase up to the point where 1/g=£6,500. The decision maker will choose t_1 when 1/g>£6,500, but now an increase in 1/g will increase $|\delta_0 - \delta_b|$ and D_0 in figure 4.2.4a and reduce the probability of incurring opportunity losses. $L(D_0)$ falls in figure 4.2.4b with a rise in 1/g, and the EVPI will fall if this off-sets the effect of the increase K_t. This occurs in figure 4.2.3 when 1/g is increased from £6,500 to £11,000 per unit of health outcome gained.

The value of 1/g is determined by the budget constraint faced by clinical practitioners in service provision. If the budget constraint is relaxed then more costly but effective health services can be provided, the cost-effectiveness ratio of the marginal service will increase, and the EVPI will rise (because the cost-effectiveness ratio of the new marginal service will always be greater than 1/g before the increase in the budget). If the budget is tightened the cost-effectiveness ratio of the marginal project will fall, service providers will be unable to take advantage of the information provided by clinical research, and the value placed on this information will diminish. The value of information, research priorities and the optimal level of research and development expenditure are all dependent on the budgetary constraint on the provision of health services.

The EVPI is also determined by the quality or confidence in the prior information. The confidence in the prior information is represented by the prior quasi sample size n_0 where a smaller quasi sample represents a more sceptical prior and less confidence in the prior mean. The impact on the EVPI of considering more or less sceptical prior is illustrated in figure 4.2.5. This demonstrates that when there is less confidence in the prior information (n_0=2) the EVPI is higher because there is

more uncertainty surrounding a decision based only on prior information. Similarly when the prior is less sceptical (n_0=18) the decision will be less uncertain and the EVPI is lower. The point at which $L(D_0)$ reaches a maximum where the clinician is indifferent between t_0 and t_1 does not change and the EVPI either reaches a peak or there is a discontinuity when 1/g=£6,500.

Figure 4.2.5

4.2.2 Hurdle II: The Expected Net Benefit of Sample Information

Proposed research which passes the first hurdle can be regarded as potentially cost-effective. To demonstrate that it will be cost-effective the optimal scale of the research (in this case sample size) must be established. Sample size will be optimal where the marginal benefit of additional sample information is equal to the marginal cost of sampling.

The expected benefit of sample information is measured by the reduction in expected opportunity loss, and this is given by the Expected Value of Sample Information ($EVSI_{(n)}$) [25,26]. This can be calculated for a particular sample size from the prior information already used to establish the EVPI and an estimate of the sample variance of the incremental net benefits of t_1 .

$$EVSI|n = K_t \,.\, \sqrt{Vn} \,.\sigma_0.L(D|n) \qquad 4.2.3a$$

$$D|n = (\delta_0-\delta_b)/ \sqrt{Vn}$$

$$\sqrt{Vn} = \sigma_0^2 / (\sigma_0^2+\sigma_n^2)$$

δ_0 = prior incremental net benefit of t_1

σ_0^2 = prior variance of δ_0

σ_n^2 = sample variance of the incremental net benefit of t_1 with sample size n

The EVSI|n is determined by four factors: the slope of the loss function; the prior

mean; the prior variance; and the variance fraction ($\sqrt{Vn}$). The variance fraction is determined by sample size and $\sqrt{Vn}$ approaches 1 as the sample size is increased. The EVSI|n approaches the EVPI as sample size tends to infinity, and this confirms the interpretation placed on the EVPI that it represents the maximum benefit that sample information can provide.

The population EVSI measures the benefits of sample information for current and future patients, and can be calculated for a particular sample size given an estimate of the incidence of patients entering the decision problem in each period.

$$\text{Population EVSI|n} = \sum_{h=1}^{H} (\text{EVSI|n.Incidence}_{(h)}.1/(1+r)^h) \qquad 4.2.3b$$

The Costs of Sampling

The cost of obtaining a sample of size n (Cs|n) takes the following simple form with fixed cost (C_f) and constant marginal cost (C_m)

$$Cs|n = C_f + C_m.n \qquad 4.2.3b$$

The marginal cost of sampling includes the additional cost of treatment when patients entering the trial are allocated to the experimental treatment. In this example patients are allocated equally between the control and experimental arms of the trial (an optimal allocation of patients in a fixed sample design is possible using dynamic programming techniques and is discussed in chapter 6) so in this example each observation on the incremental net benfit of t_1 requires two patients to enter the trial with each allocated to either t_1 or to t_0. The marginal costs of observing and recording the results of treatment are assumed to be negligible, and the marginal cost of an additional trial entrant is half the incremental cost of t_1.

$$C_m = (E(C|t_1)-E(C|t_0))/2 \qquad 4.2.3c$$

The maximum sample size that should ever be considered can now be established, because when $n = (EVPI-C_f)/C_m$ the cost of the research will be exactly equal to the maximum possible benefits.

Expected Net Benefit of Sampling and Optimal Sample Size

The Expected Net Benefit of Sampling (ENBS|n) is the difference between the total benefit and the total variable cost for a particular sample size.

$$ENBS|n = EVSI|n - Cs|n \qquad 4.2.3d$$

Sample size will be optimal (n*) when ENBS|n is positive and at a maximum. The relationship between sample size and ENBS|n is shown in figure 4.2.6a.

Figure 4.2.6a

In figure 4.2.6a 1/g=£4,000 and $D_0>0$. The decision-maker initially prefers the control treatment t_0. Small amounts of sample information are unlikely to change this decision, so that ENBS|n<0 when sample sizes are very small. However there is a range of sample size where EVSI|n-Cs|n>0 and the ENBS|n reaches a maximum when n*=92. At this point $\partial EVSI/\partial n = C_m$, and sample size is optimal.

The EVSI|n initially increases at an increasing rate with n, but ultimately declines because as n tends to infinity the EVSI approaches the EVPI. The variable costs of sampling continue to rise at rate C_m, and the optimal sample size will be finite. The problem of potentially infinite sample size associated with the traditional approach is avoided.

A second hurdle for potentially cost-effective research can now be constructed. The EVSI|n* represents the maximum that those commissioning research should be willing to pay given the budget constraint on service provision. In figure

4.2.6a, EVSI|n*= £1,202,021. If the total cost of the research is less than this amount then it is cost-effective and should be implemented. The ENBS|n* is the expected net present value of research. In figure 4.2.6a, EVSI|n*-C_m.n* = £723,621, and this is the maximum fixed cost of research which could be incurred if the research is to remain cost-effective. ENBS|n* can be used to prioritise research proposals. By implementing first those proposals with highest net present value, the maximum benefit can be obtained for a given research and development budget. The optimal level of research and development expenditure is given by the cost of implementing all proposals with a positive net present value. At the margin, ENBS|n* is zero.

The ENBS|n* also represents the opportunity cost of failing to implement cost-effective proposals. For example, if the fixed cost of this research proposal was estimated to be £100,000, the expected net benefit is £623,621, and it would pass the second hurdle. However, if this proposal was rejected on the grounds of medical ethics then the implicit opportunity cost of this ethical position is £623,621: equivalent to 156 units of health outcome (using 1/g=£4,000 per unit of health outcome gained). Consideration of medical ethics is an essential element in trial design and data monitoring, but this approach makes it possible to estimate the opportunity cost (to society as a whole) of particular concerns for the individuals involved in a clinical trial. In this way the inevitable trade-off between individual and collective ethics can be made explicit [38, 79, 91, 143, 144, 145]. If these trade-offs are explicit they can be made consistently, and be open to criticism and debate.

Value of Health Outcome

The expected net benefit of sampling and the optimal sample size are dependent on the budgetary restrictions on service provision or the value of health outcome. In this example when the value of health outcome is higher, at £10,000, the expected net benefit and optimal sample size is also higher. The ENBS, EVSI and

the variable costs of sampling are illustrated in figure 4.2.6b for a range of possible sample sizes. The ENBS reaches a maximum of £5,040,755 at and optimal sample size of 188 . At this value of health outcome the clinician should choose t_1. based only on prior information. The probability that this will be the wrong decision is higher than when 1/g=£4,000 (because $L(D_0)$ when 1/g=£10,000 is greater than $L(D_0)$ when 1/g=£4,000 in figure 4.3.4b) and the value placed on opportunity losses is also higher so the value of sample information is also higher

Figure 4.2.6b

The expected net benefit of sampling when the value of 1/g is increased to £20,000 is illustrated in figure 4.2.6c. The ENBS reaches a maximum of £5,425,760 at an optimal sample size of 246. The optimal sample size and the value of sample information is higher because although the prior decision to treat with t_1 is less uncertain ($L(D_0)$ when 1/g=£20,000 is less than $L(D_0)$ when 1/g=£10,000 in figure 4.3.4b) the value placed on opportunity losses is higher, and in this case the value of sample information is also higher. This demonstrates that the value of sample information and the technically efficient scale of clinical research is dependent on budgetary constraints on service provision and the issues of allocative and technical efficiency cannot be addressed before health outcome has been valued in monetary terms.

Figure 4.2.6c

In this example the optimal sample size increases with the value placed on health outcome and this is illustrated in figure 4.2.7. The relationship between optimal sample size and the value of 1/g for a more sceptical (n_0=2) and a less sceptical prior (n_0=18) is also illustrated in figure 4.2.6c. When the value of health outcome is low the optimal sample size is zero and the decision should be based only on prior information. In these circumstances the prior decision will be to reject the

experimental treatment and not treat without conducting a clinical trial. If there is less confidence in the prior information then sampling becomes optimal at a lower value of 1/g (if n_0=2 then n*=0 when 1/g<£2,000, but if n_0=18 then n*=0 when 1/g<£5,000) because the prior decision not to treat is less certain and sample information is more valuable.

Figure 4.2.7

Figure 4.2.7 also shows that when the prior is less sceptical (n_0=18) and the value of health outcome is high (1/g>£12,000) then the optimal sample size is zero. Decisions should once again be based only on prior information which is now to treat using the experimental treatment t_1. This suggests that there may be circumstances in which a new treatment should be adopted without gathering sample information through a clinical trial. If the prior incremental net benefit of the new treatment is sufficiently high and if there is sufficient confidence in this prior information, it will not be worth incurring the costs of a trial because these resources could be better used elseware, either in service provision or other areas of clinical research. This demonstrates that a decision-analytic approach can be used to set rational limits to evidence based medicine and provide a framework where new treatments of potentially great benefit [80, 134] can be adopted without incurring the cost (including the opportunity cost of the delay before the results are available) of a clinical trial.

Once the optimal sample size has been established the relationship between the value of 1/g and the maximum value of the ENBS|n can be considered. This is illustrated in figure 4.2.8 for three different priors. These estimates of the ENBS|n* represent the value of the second hurdle and show that the second hurdle is sensitive to both the value of health outcome and the strength of prior information. The ENBS|n* reaches a peak or shows a discontinuity when 1/g=£6,5000 because this is the prior cost-effectiveness ratio for this decision problem. At this point the prior decision is most uncertain, D_0=0 in figure 4.2.3a,

and small amounts of sample information is valuable.

Figure 4.2.8

Decision Rules

Once a proposal has passed both hurdles the decision rule which should be applied to the information provided by the sample must be established. The objective is to minimise the expected opportunity loss (maximise expected net benefits).

Before sample information is available the treatment decision can only be based on prior information. Once sample information is available this must be combined with prior information to produce a posterior distribution with mean δ_1 and variance σ_1^2. The posterior mean is a weighted average of the prior and sample information with the weights representing the informational content of each [87].

$$\delta_1 = (I_0.\delta_0+I_x.\delta_x)/(I_0+I_x) \qquad 4.2.4a$$

δ_x = sample mean

$I_0 = 1/\sigma_0^2$

$I_x = 1/(\sigma_{n*}^2)$

A decision rule which will minimise opportunity loss once sample information is available is to choose t_0 when $\delta_1<0$, and choose t_1 when $\delta_1>0$. The decision-maker will be indifferent when $\delta_1=\delta_b=0$. An equivalent decision rule based on sample results can be established by defining a critical value for the sample mean (δ_{x*}) which gives $\delta_1=\delta_b=0$.

$$\delta_{x*} = ((I_0+I_x).\delta_b - I_0.\delta_0)/I_x \qquad 4.2.4b$$

If the sample mean is less than δ_{x*} then $\delta_1<0$ and the decision maker should

choose t_0. In this example when 1/g =£4,000 and δ_0= -1, t_0 will be chosen on the basis of prior information alone. With an optimal sample size of 92, δ_{x^*}=0.504. A sample mean of at least 0.504 units of health outcome would need to be observed before the decision would be changed.

4.3 Conclusions

The traditional approach to clinical trial design is inconsistent with concepts of efficiency even when the cost of the treatment alternatives are considered alongside the trial. The traditional approach implicitly places an infinite value on the benefits of sample information, leading to either unbounded or arbitrary sample sizes. Because the marginal cost of acquiring sample information is not considered it is unable to provide a framework for setting priorities in clinical research or establishing efficient research design. It is not able to address the problems which were posed at the end of chapter 3.

A decision-analytic approach can provide practical tools for research priority-setting. The information generated by clinical research is valued in a way which is consistent with concepts of efficiency, and with the methods used to set priorities in service provision. The prior information, which is implicit in the traditional approach, is identified and handled consistently so that it is open to criticism, alternative formulation, and empirical testing. It is able to combine information from a variety of sources taking into account the variable quality of this information.

The simple example of a fixed sample phase III trial shows that these techniques can be used to identify areas of clinical practice where the cost of uncertainty is high, and where the potential benefits of clinical research will also be high. Estimates of the EVPI and the EVSI|n* can be used to construct two hurdles which proposed research must overcome before it can be considered cost-effective.

Hurdle I $\quad EVPI > C_f$

The first hurdle is based only on prior information, and asks if the EVPI (the cost of uncertainty or the maximum value of sample information) is greater than the

costs of the proposed research. This hurdle can eliminate proposed research which will not be cost-effective before issues of research design must be addressed. Those proposals which pass the first hurdle can be regarded as potentially cost-effective and can be considered at the second hurdle.

Hurdle II $\quad EVSI|n^* - Cs|n^* = ENBS|n^* > 0$

The second hurdle ensures that the design of potentially cost-effective research is technically efficient, and that it will be cost-effective when conducted at the optimal scale. The $ENBS|n^*$ represents the value of the proposed research to the providers and consumers of health services. It also represents the opportunity cost of rejecting cost-effective research proposals. Estimates of the expected net benefit of research can be used to rank proposed research. By implementing proposals with higher net benefit first, the maximum health benefits can be gained for limited research and development resources. This approach provides a means to decide which clinical decision problems are worth evaluating in a clinical trial and what is the technically efficient scale of this research.

All but one of the problems posed at the end of chapter 3 have been addressed using these techniques. However two substantial problems remain. In chapter 2 and 3 it was argued that many clinical decision problems are sequential and involve a choice between many competing alternative strategies. The approach has been illustrated using a single-stage decision problem and will be generalised to the two and four-stage sequential test/treatment decision problems in the next chapter. The second problem is the selection of relevant alternatives which should be compared in a clinical trial. This problem did not arise in this simple single-stage decision but this issue is considered in chapter 5 and will be addressed in detail in chapter 6.

Appendix C Tables and Figures for Chapter 4

Contents:

Figure 4.2.1 Decision Tree for the Single-Stage Decision Problem

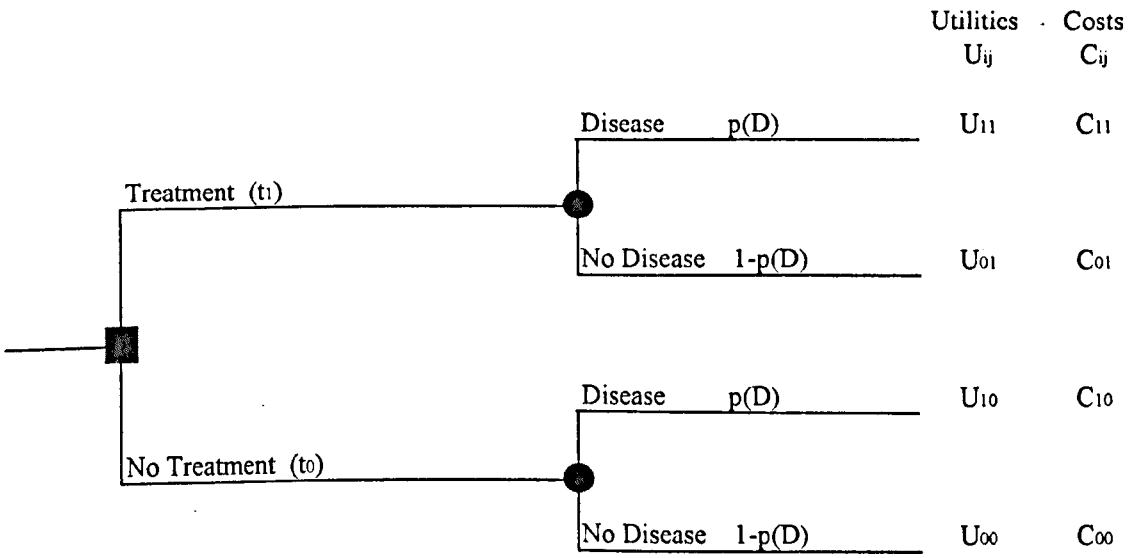


Table 4.2

Numerical Example for the Single-Stage Decision Problem

	Prior Mean	Prior SD	Population SD	Quasi prior sample size (n_0)
U_{11}	6	0.5164	1.2649	12
U_{01}	8	0.5164	1.2649	12
U_{10}	2	0.2582	0.8942	6
U_{00}	10	0.2582	0.8942	6
C_{11}	£12,000			
C_{01}	£8,000			
C_{10}	0			
C_{00}	0			
p(D)	0.6	0.1	0.4899	24

p(D) = prior probability of disease

In this example 1000 patients enter the decision problem in one year.

Figure 4.2.2 Opportinity Loss Functions for t_1 and t_0

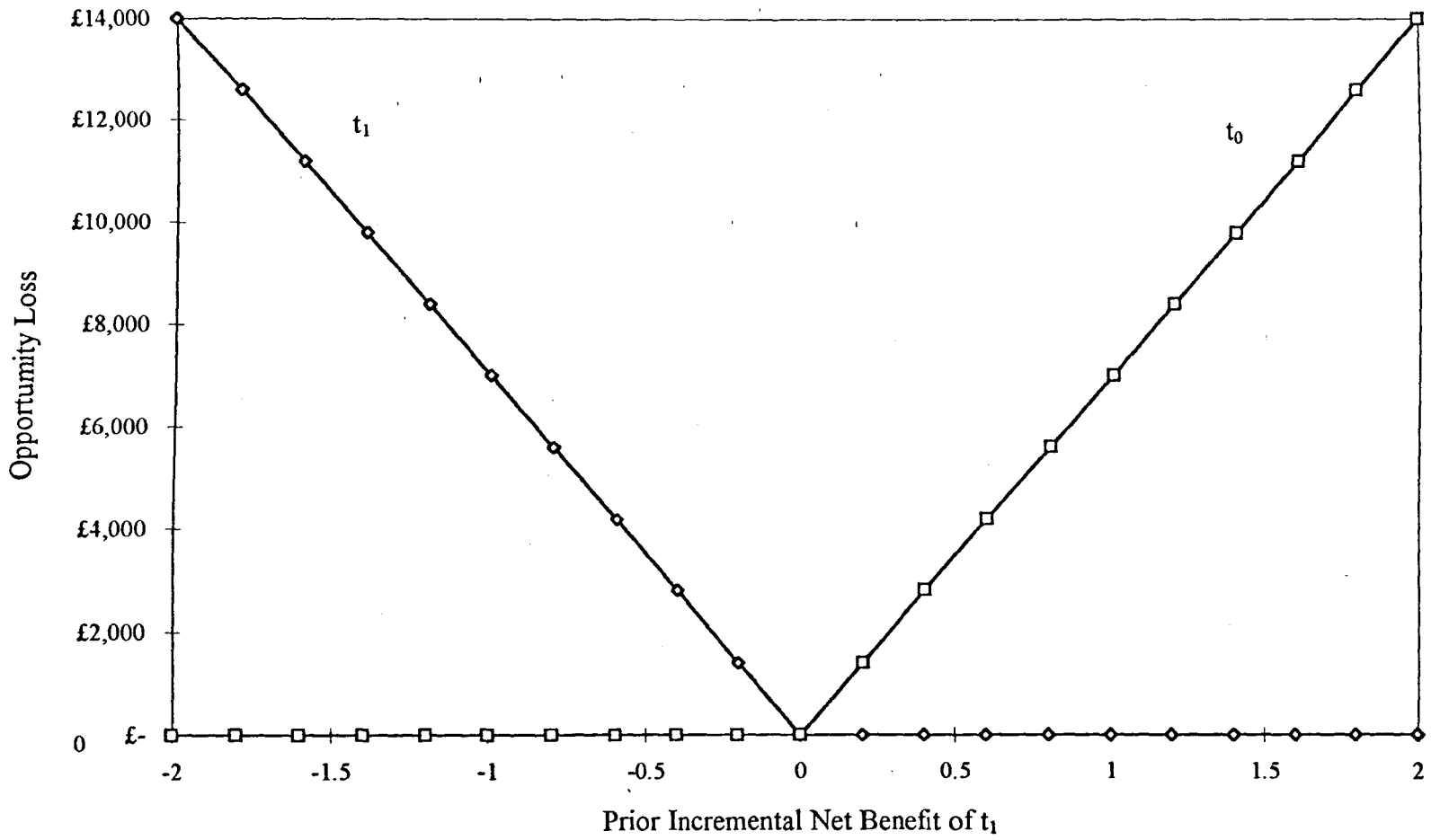


Figure 4.2.3 EVPI for the Single-Stage Decision Problem

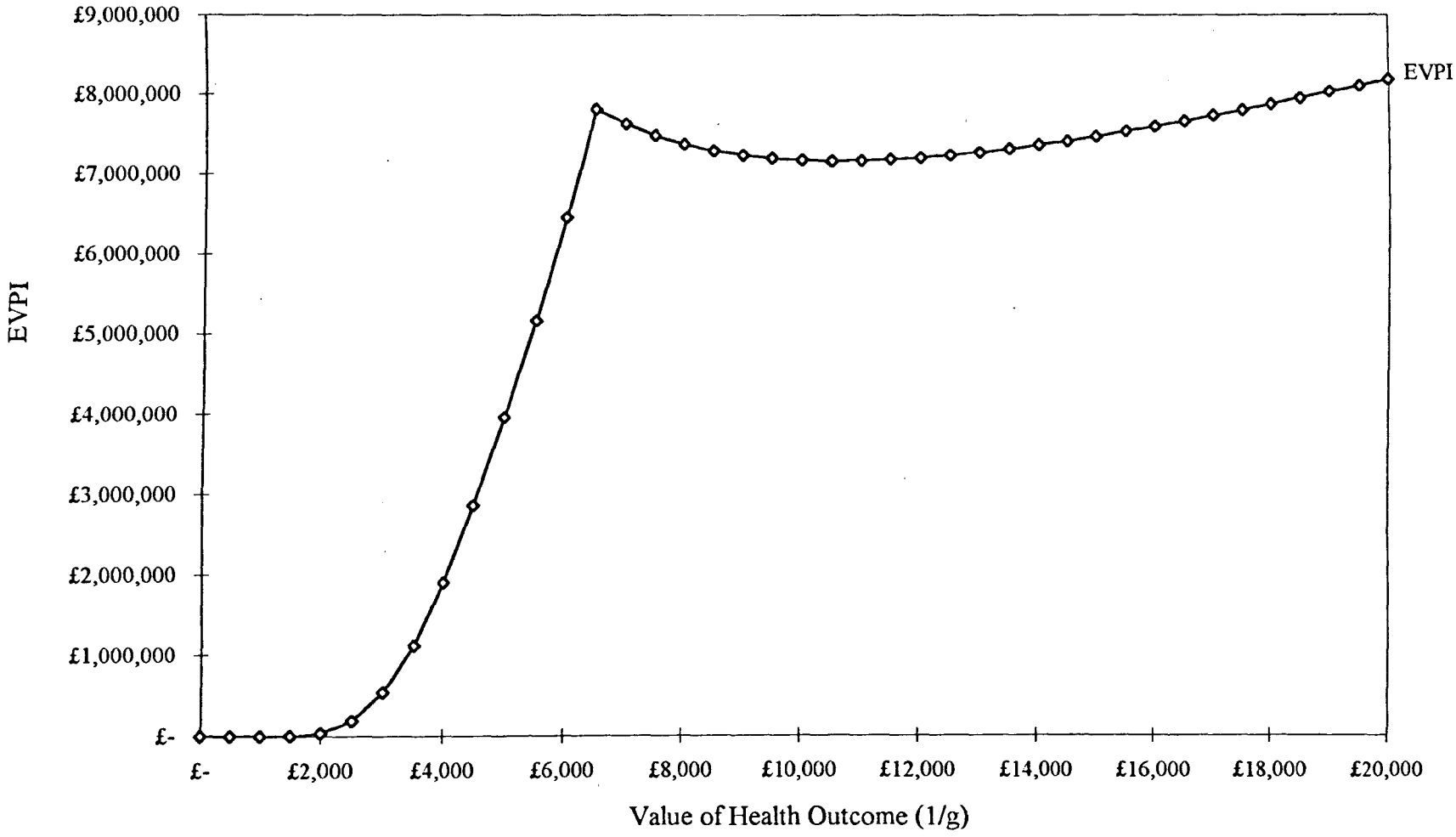


Figure 4.2.4a Standardised Distance (D_0) for the Single-Stage Decision Problem

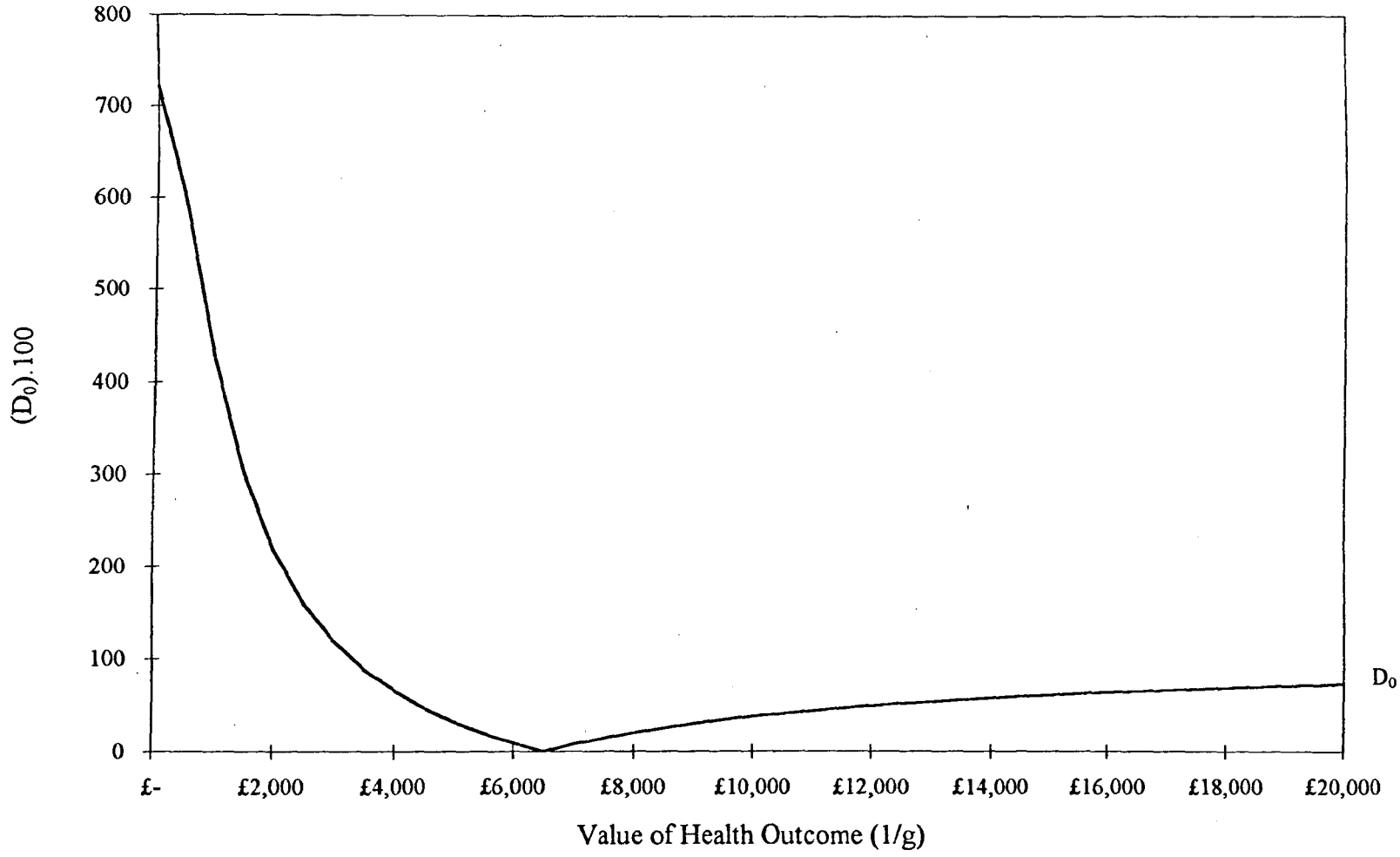


Figure 4.2.4b Loss Integral ($L(D_0)$ forthe Single-Stage Decision Problem

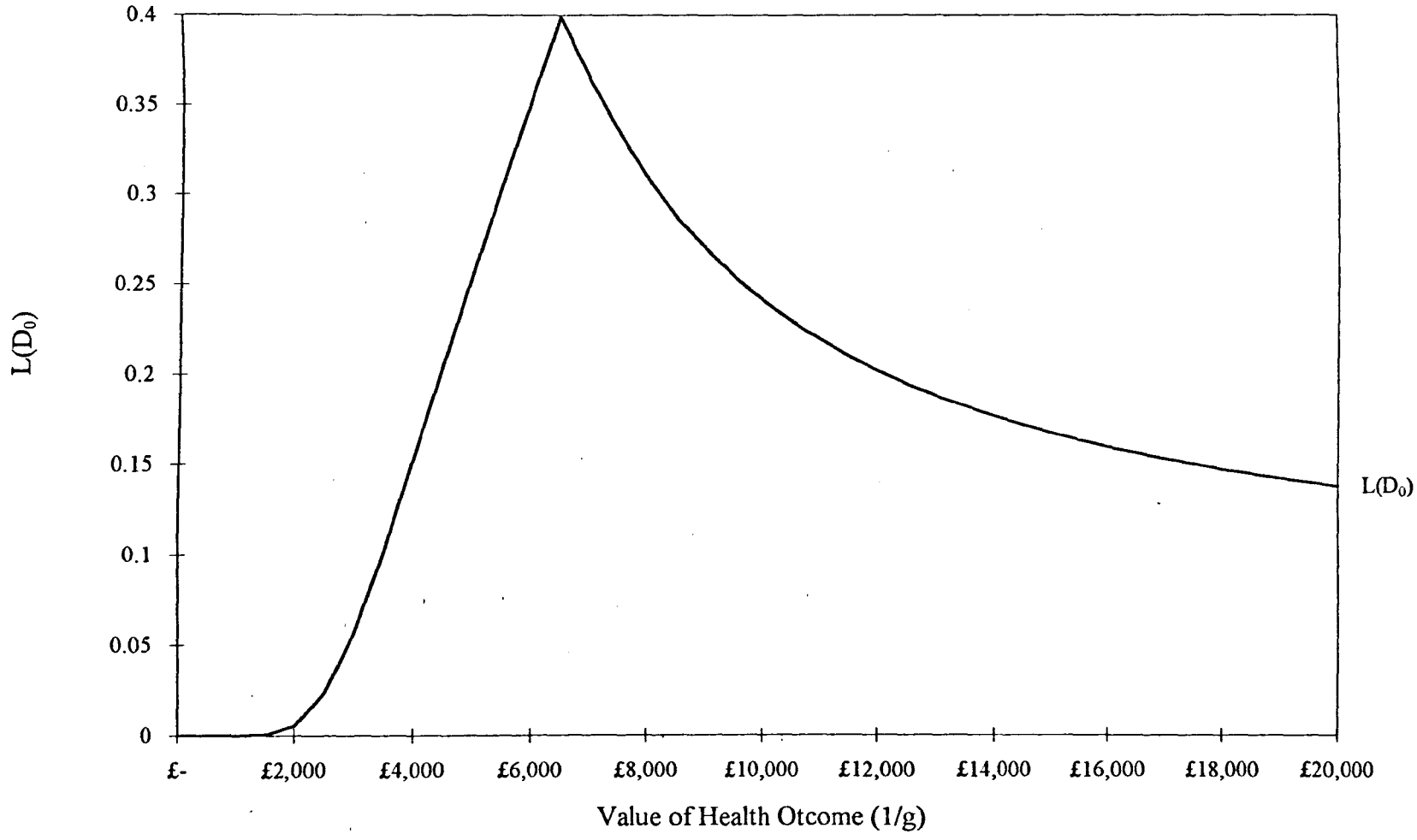


Figure 4.2.5 EVPI and the Strength of Prior Information

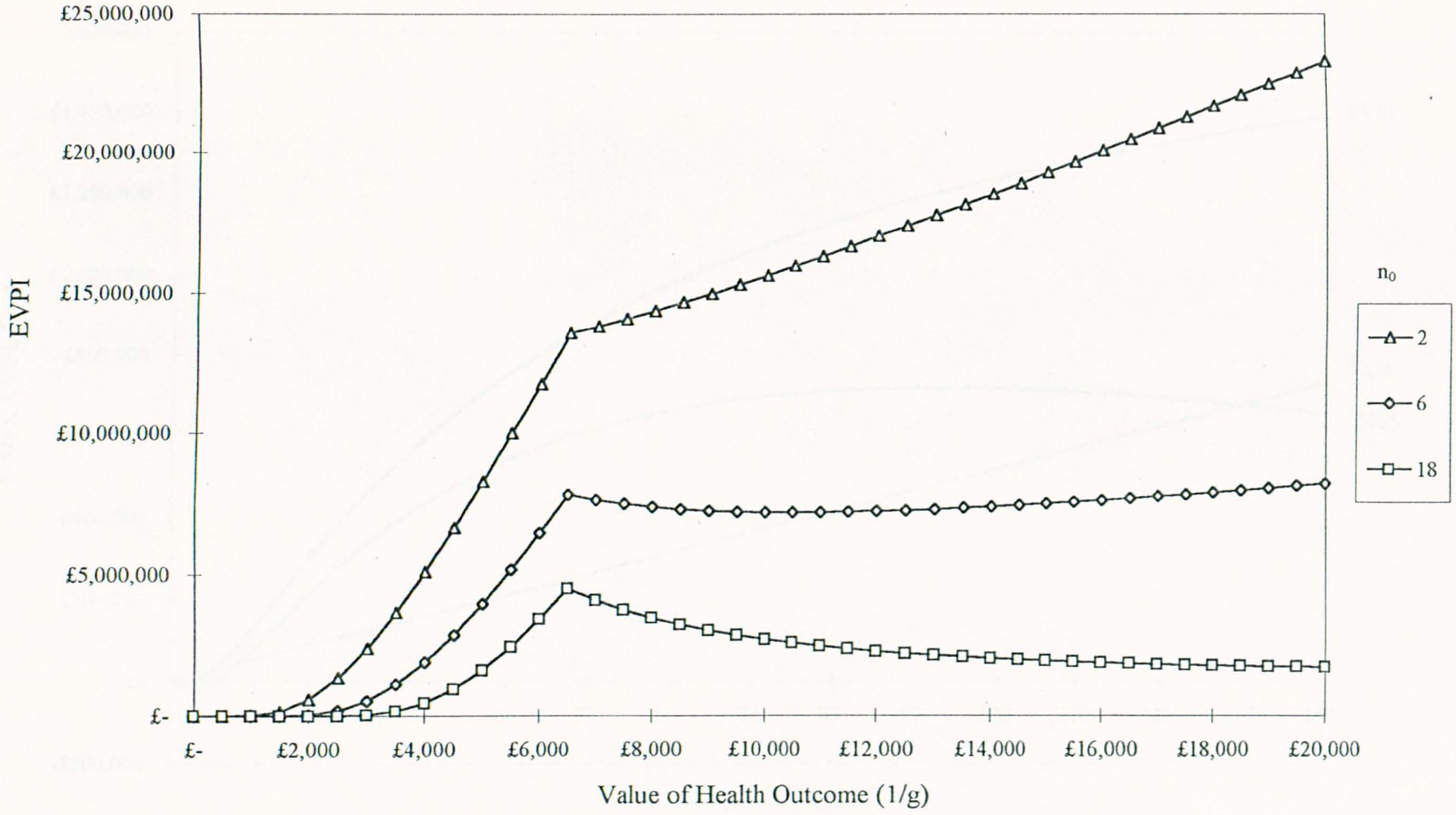


Figure 4.2.6a ENBS, EVSI, and $C_m.n$ (when 1/g=£4,000)

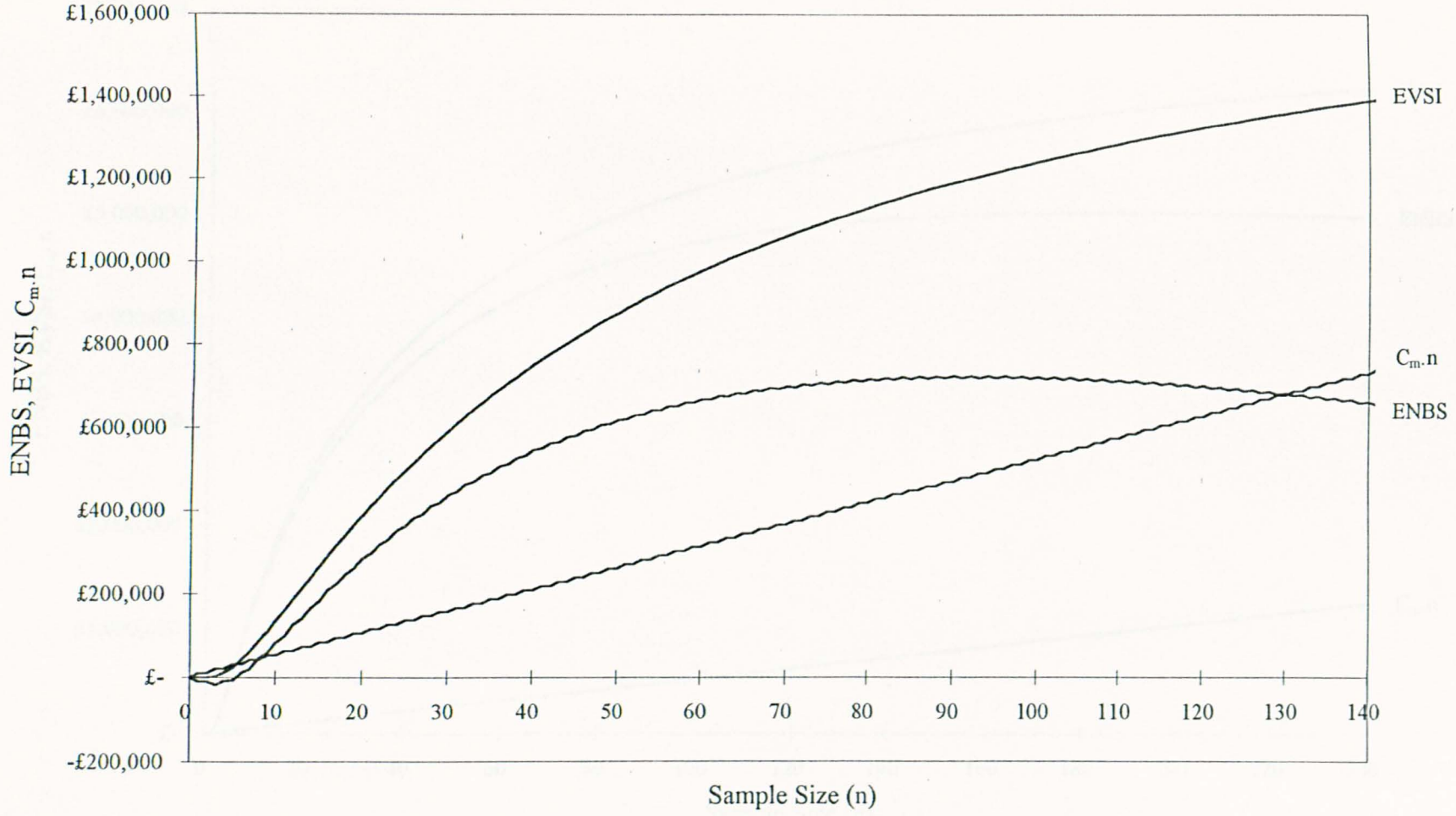


Figure 4.2.6b ENBS, EVSI, and $C_m.n$ (when 1/g=£10,000)

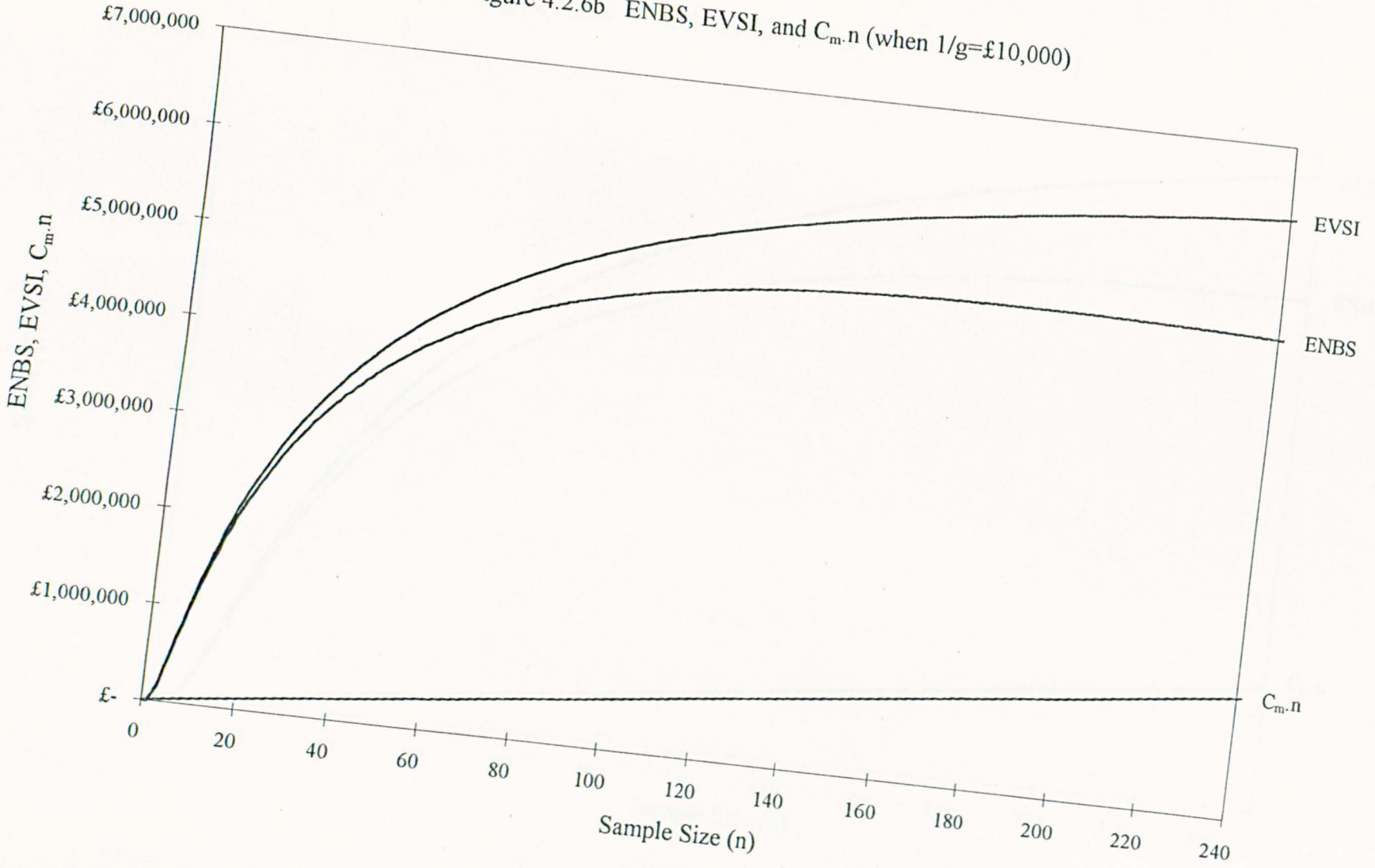


Figure 4.2.6c ENBS, EVSI, and $C_m.n$ (when 1/g=£20,000)

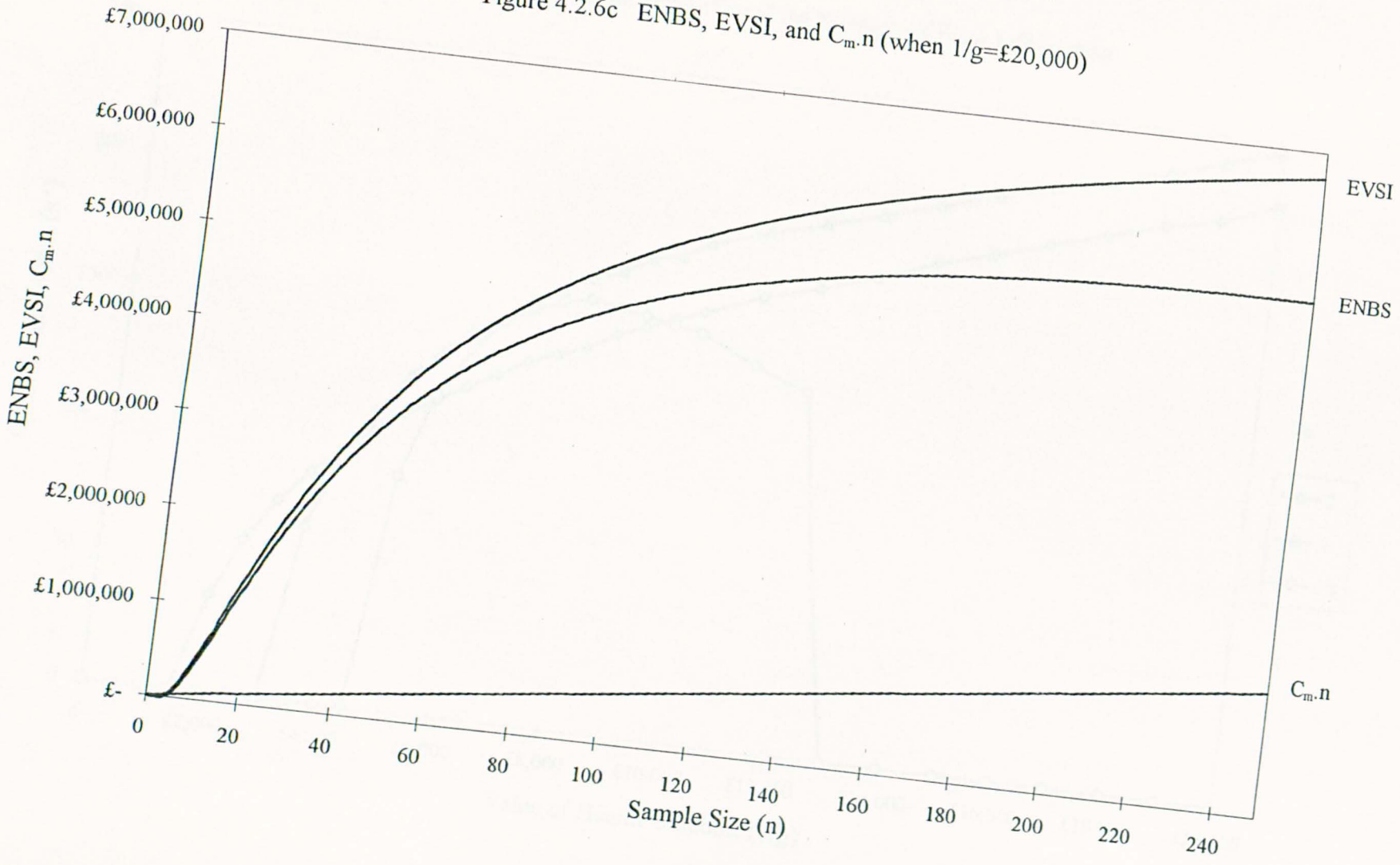


Figure 4.2.7 Optimal Sample Size (n*) and the Strength of Prior Information

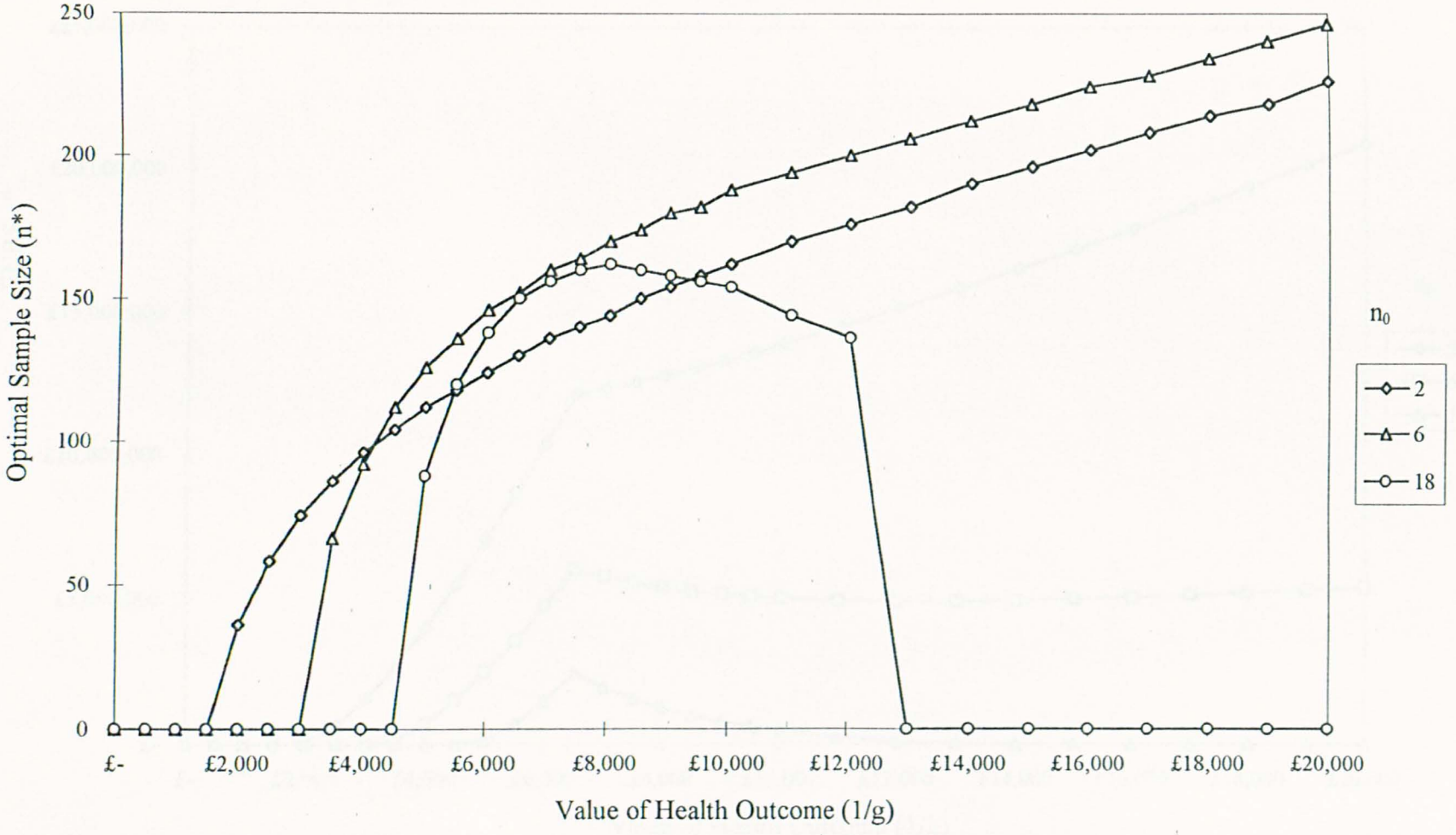
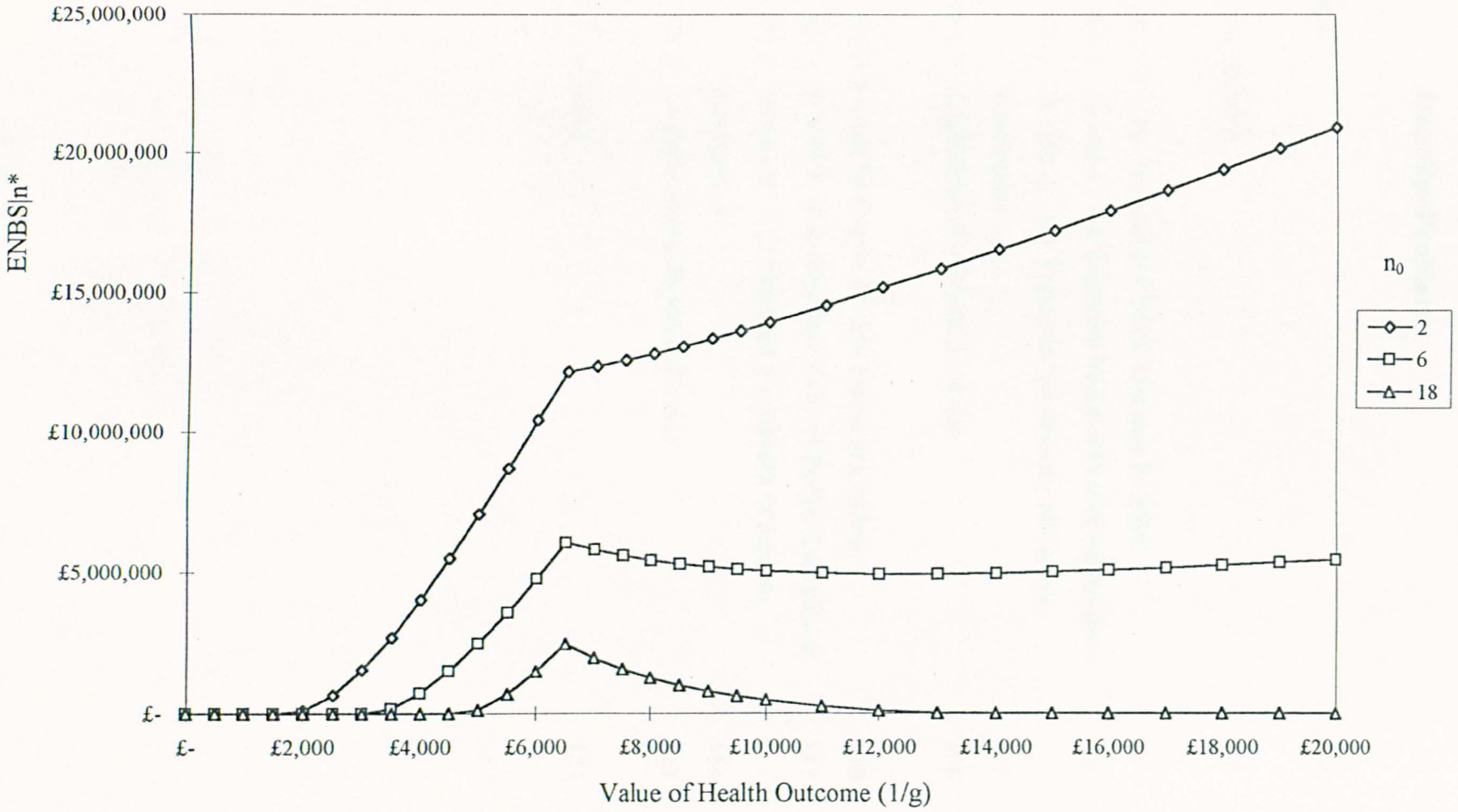


Figure 4.2.8 ENBS|n* and the Strength of Prior Information



Chapter 5 The Value of Information for Sequential Clinical Decision Problems

Contents: Page

5.1 Introduction

In the last chapter a decision analytic approach was applied to the simple single-stage decision problem of choosing between treatment (t_1) and no treatment (t_0). This single-stage problem is the choice between the fallback strategies which was discussed in chapter 2. However it was argued in chapter 2 that many clinical decision problems are sequential and involve a number of contingent decisions concerning diagnostic and treatment strategies. This chapter will demonstrate that the approach that was used in chapter 4 can also be applied to sequential clinical decision problems.

In section 5.2 this approach is applied to the test-treatment decision problem where the clinician faces a choice between using a diagnostic test and treating according to test results, or choosing not to test and selecting either t_1 or t_0. This is the same two-stage decision problem which was used to illustrate the strategy for the economic evaluation of diagnostic information proposed by Phelps an Mushlin in chapter 2 and the same numerical example is used to construct the first and second hurdles for proposed research. In section 5.3 this approach is also applied to the four-stage decision problem which was used to generalise the Phelps Mushlin strategy in chapter 3.

The value of perfect information (the cost of uncertainty) for these more complex decision problems can be established and this is used as the first hurdle that proposed research must overcome before it can be considered potentially cost-effective. If the expected value of perfect information exceeds the estimated fixed cost of proposed research the research is potentially cost-effective. The cost of uncertainty at particular points in a sequence of decisions can also be identified. Those contingent decisions where the cost of uncertainty is highest will be the points where additional information will be most valuable and this approach can be used to set priorities in acquiring information to inform particular contingent decisions.

The value of sample information at each point in the sequence of decisions can be established and compared to the cost of sample information. By establishing the expected net benefits of sample information at each stage in a sequential decision problem the optimal sample size entering the initial decision node can be identified, and the expected net benefit of prospective clinical research can be estimated. This can be used as the second hurdle that proposed research must overcome before it can be considered cost-effective. It provides a means of setting priorities in research and development across sequential clinical decision problems. The analysis in this chapter (like chapter 4) assumes a fixed and equal allocation of trial entrants between the different arms of the trial at each decision node. This assumption will be relaxed in chapter 6 where a dynamic programming approach is used to allocate patients efficiently.

5.2 A Two Stage Sequential Clinical Decision Problem

The same methods which were applied to the single-stage decision problem in chapter 4 can be applied to the two-stage sequential decision problem which is illustrated in figure 5.2.1. This is the same decision problem which was used to illustrate the approach to the economic evaluation of diagnostic information proposed by Phelps and Mushlin in chapter 2, but it has been structured following the correct chronology and includes Bayesian probability revisions. This chapter also uses the same numerical example but with a prior distribution for each of the health outcomes and path probabilities. The prior distributions reflect the quality of prior information or confidence in the prior mean. In this example the quasi prior sample (n_0) is higher for current practice (t_0) where more prior information may be available. The values of the prior mean and variance and population variance for these variables are reported in table 5.2. The assumption of normality and independence, which was also made in chapter 4, allows normal loss functions to be used and covariance terms to be zero when calculating the variance of expected net benefits at different stages.

Figure 5.2.1

Table 5.2

Figure 5.2.1 illustrates the decision problem facing the clinician. If the clinician decides not to use the diagnostic test then at the second stage the decision problem is to choose to treat with t_1 or not to treat (t_0). This is identical to the single-stage decision problem in chapter 4, but it is now a contingent treatment decision because a diagnostic device is available. The initial decision at the first stage is to choose to test and treat according to test results, with t_1 following a positive test and t_0 following an negative result, or to not test and follow the contingent treatment decision at the second stage.

The problem is to establish the value of perfect information for this sequential

decision problem to construct the first hurdle and then to estimate the expected net benefits of sampling and optimal sample size so that the second hurdle can be constructed. This will allow an efficient allocation of research and development resources between technically efficient research designs.

5.2.1 Hurdle I: The Expected Value of Perfect Information

The clinician must make a decision at two points in this model. Without any additional information this choice can be based only on the prior information. The clinician should choose the alternative with the highest prior net benefit, but when estimates of net benefit are based on prior information there is a possibility that this decision will be wrong and opportunity losses will be incurred. Following the intuition of chapter 4, the expected value of opportunity losses is the expected cost of uncertainty surrounding the decision problem, or the Expected Value of Perfect Information. In general the EVPI for a sequential decision problem with S stages and a choice between two alternatives at each stage (s=1,..,S), is the sum of the $EVPI_{(s)}$ at each stage or at each point where the clinician faces an uncertain decision.

$$EVPI=\sum_{s=1}^{S} EVPI_{(s)} \qquad 5.2.1a$$

$$EVPI_{(s)} = K_t.\sigma_{0(s)}.L(D_{0(s)}) \qquad 5.2.1b$$

$D_{0(s)} = (\delta_{0(s)} - \delta_b)/\sigma_{0(s)}$

$\delta_{0(s)}$ = prior incremental net benefit at stage (s)

$\sigma_{0(s)}$ = prior standard deviation of $\delta_{0(s)}$

The decision problem in figure 5.2.1 indicates that the clinician will face an uncertain choice at two points: firstly when either test or no test must be selected and secondly where t_1 or t_0 must be selected (contingent on choosing not to test), so in this decision problem (s=1, 2) and the $EVPI_{(s)}$ must be estimated at both

stages.

EVPI at Stage 2

The clinician must make a decision at stage 2 of whether to treat with t_1 or t_0 given that it has been decided not to use the test. This contingent decision is based on the prior incremental net benefit of t_1 ($\delta_{0(2)}$) and must be made before the initial diagnostic decision can be taken at stage 1. The decision problem at stage 2 is identical to the decision problem considered in chapter 4 and the $EVPI_{(2)}$ at what is now a contingent treatment decision in this larger sequential problem is identical to that reported in chapter 4 where:

$$\delta_{0(2)} = (E(U|t_1)-g.E(C|t_1))-(E(U|t_0)-g.E(C|t_0)) \qquad 5.2.2$$

EVPI at Stage 1

At the initial diagnostic decision the clinician must choose between testing and treating according to test results and not testing and following the contingent treatment decision at stage 2. This decision will also be based on prior information and there is a chance that choices based on the prior incremental net benefit of testing at this first stage will be wrong and opportunity losses will be incurred. The expected opportunity loss at stage 1 (the $EVPI_{(1)}$) is in addition to the expected costs of uncertainty at stage 2 ($EVPI_{(2)}$) and the EVPI for the full decision problem will include both the EVPI at the contingent treatment decision and at the initial diagnostic decision . The $EVPI_{(1)}$ at the initial decision is calculated in the same way as stage 2 but now $\delta_{0(1)}$ is the prior incremental net benefit of testing. This is the difference between the prior net benefits of testing ($E(U|t_e)-g.E(C|t_e)$) and not testing and selecting either t_1 or t_0 at stage 2. When $1/g>£6,500$ t_1 is selected at stage 2 and:

$$\delta_{0(1)} = (E(U|t_e)-g.E(C|t_e)) - (E(U|t_1)-g.E(C|t_1)) \qquad 5.2.3a$$

But when 1/g<£6,500, t_0 is selected at stage 2:

$$\delta_{0(1)} = (E(U|t_e)-g.E(C|t_e)) - (E(U|t_0)-g.E(C|t_0)) \qquad 5.2.3b$$

EVPI for the Two-Stage Problem

The EVPI for this sequential decision problem is illustrated in figure 5.2.2 and is the sum of the EVPI at stage 1 and stage 2. The EVPI rises with the value of health outcome because the value placed on opportunity losses (the slope of the loss function (K_i=2/g)) increases with 1/g, but there are now two discontinuities in this relationship which are due to discontinuities in the $EVPI_{(1)}$ and $EVPI_{(2)}$. The discontinuity in the $EVPI_{(2)}$ occurs when the clinician is indifferent between t_1 and t_0, at 1/g=£6,500. The $EVPI_{(2)}$ reaches a peak at this point because $D_{0(2)}$=0 in figure 5.2.3a and $L(D_{0(2)})$ reaches a maximum in figure 5.2.3b in exactly the same way as in chapter 4.

Figure 5.2.2

The $EVPI_{(1)}$ at the initial decision rises with the value of health outcome but there are two discontinuities in this relationship. The first also occurs when the clinician is indifferent between t_1 and t_0, because when 1/g<£6,500 the clinician compares the net benefit of testing with the net benefit of t_0 ($\delta_{0(1)}$=5.2.3b) but when 1/g>£6,500 the net benefit of testing is compared to the net benefit of t_1 ($\delta_{0(1)}$=5.2.3a). The second discontinuity in the $EVPI_{(1)}$ occurs when the clinician is indifferent between the test and no test alternative at stage 1. In this numerical example it is when 1/g=£11,800 and at this point the clinician is indifferent between testing and treatment t_1. This where the initial decision is most uncertain and $D_{0(1)}$=0 in figure 5.2.3a and $L(D_{0(1)})$ reaches a maximum in figure 5.2.3b. As the value of 1/g increases the clinician will prefer testing ($\delta_{0(1)}$ >0) and the prior incremental net benefit of testing will increase ($D_{0(1)}$ increases), reducing the probability of incurring opportunity losses ($L(D_{0(1)})$ falls). This is offset by the

increase in value placed on opportunity losses when they occur (K_t=2/g rises) and the $EVPI_{(1)}$ increases with 1/g.

Figure 5.2.3a

Figure 5.2.3b

It was argued in chapter 2 and chapter 3 that many clinical decision problems are sequential and involve a choice between a number of competing strategies. This example shows that the EVPI can be established for sequential clinical decision problems by estimating the cost of uncertainty at each stage of the decision problem. The estimates of the EVPI in figure 5.2.2 can be used in the same way as in chapter 4 as a first hurdle that proposed clinical research must overcome. If the estimated cost of research is less than the EVPI then the proposed research is potentially cost-effective. Estimates of the EVPI can be used to set priorities across different sequential clinical decision problems by identifying those decision problems which may benefit most from information generated by prospective research as well as systematic reviews of existing literature and non-experimental research designs [40, 41, 119, 120].

This numerical example also illustrates that the first hurdle may be very sensitive to the value of health outcome and in this example doubling the value of 1/g leads to an approximately four-fold increase in the EVPI. This simply demonstrates that the value of information depends on the value placed on opportunity losses when they occur, which is double the value of health outcome (K_t=2/g). Just as in chapter 4 the relationship between the EVPI and 1/g demonstrates that the value of information and research priorities cannot be separated from the budgetary restrictions on service provision. If the budget is relaxed then the cost-effectiveness ratio of the marginal project (1/g) will increase and the value of information will increase. Similarly if the budget for service provision is tightened then the cost-effectiveness ratio of the marginal project (1/g) will fall and the EVPI will fall.

The example also demonstrates that the cost of uncertainty for a clinical decision problem will be underestimated if some alternatives are ruled out as not relevant and a sequential decision problem is simplified to a single-stage problem. This is because the EVPI for the whole decision problem will be the sum of EVPI at each of the contingent decisions and at the initial decision. An analysis which simplified this sequential clinical problem to a single-stage problem by excluding the diagnostic process as not relevant (because it is not cost-effective when 1/g<£11,800) would underestimate the EVPI because the process of simplification excludes some alternatives which are feasible and relevant and in certain circumstances could become the preferred strategy.

Finally it demonstrates that by calculating the EVPI at each stage of a sequential decision problem, those points in the sequence of decisions where the cost of uncertainty is highest can be identified. This is not necessarily the case with conventional sensitivity analysis [18, 19] because the prior distributions for the key variables and the value placed on opportunity losses at sensitive decisions are not necessarily taken into account. This can be illustrated in figure 5.2.2. If 1/g=£7,000 a traditional approach may regard the treatment decision at stage 2 to be more sensitive than the diagnostic decision at stage 1 because this value of 1/g is very close to the prior cost-effectiveness ratio at stage 2 (£6,500) and small changes to the key variables could change the treatment decision. The prior cost-effectiveness ratio at stage 1 (£11,800) is further from this value of 1/g and the decision not to use the diagnostic test at stage 1 may be regarded as less sensitive. However even when 1/g=£7,000 the $EVPI_{(1)}$ at the initial diagnostic decision is greater than the $EVPI_{(2)}$ at the contingent treatment decision. The cost of uncertainty at stage 1 is greater than at stage 2, demonstrating that simple measures of sensitivity may be misleading particularly if they are used to identify those points where information may be most valuable

Figure 5.2.4

The relationship between the EVPI for the full decision problem and the quality of or confidence in, the prior information is illustrated in figure 5.2.4. The level of confidence is measured by n_0, which is the ratio of population to prior variance (the quasi sample size of prior information). When the confidence in prior information is reduced from n_0=6 to n_0=2 the EVPI increases because there will be more uncertainty surrounding a decision based on prior information. As the confidence in prior information is increased (n_0=18) the EVPI falls because the probability of incurring opportunity losses will decline. Clearly the first hurdle is sensitive to strength of the prior, and in this example a three-fold increase in the strength of prior information leads to an approximately three-fold decline in the EVPI.

5.2.2 Hurdle II: The Expected Net Benefits of Sample Information

If the EVPI exceeds the estimated fixed costs of prospective research then research is potentially cost-effective. The next step is to estimate the benefit of sample information and the marginal cost of acquiring sample information to establish the expected net benefits of sampling. The scale of proposed research will be technically efficient and sample size will be optimal when the ENBS reaches a maximum. The ENBS is the second hurdle that proposed research must overcome before it can be regarded as cost-effective. It operates in the same way as the second hurdle for the single-stage decision problem in chapter 4 and proposed research will be cost-effective if the ENBS at the optimal sample size exceeds the fixed cost of research.

The sequential clinical decision problem in figure 5.2.1 provides two points where trial entrants will be allocated to the different arms of this trial. At stage 1 in figure 5.2.1 the trial entrants ($n_{(1)}$) are allocated equally to the test and no test arms of the trial, so that $n_{(1)}/2$ patients will be assigned to the test arm of the trial, $(n_{(1)}/2).p(t_e^+)$ will receive t_1 following a positive test and $(n_{(1)}/2).p(t_e^-)$ will receive t_0

following a negative test result. The patients enrolled in the trial who are assigned to the no test arm at stage 1 will enter stage 2 ($n_{(2)}=n_{(1)}/2$) and will be allocated equally between t_1 and t_0. Other optimal allocation rules are considered in chapter 6.

The Expected Value of Sample Information is calculated using the same methods which were described in chapter 4, but now in a sequential decision problem there are benefits from sample information at each stage (s) given the sample size entering and allocated at that stage ($n_{(s)}$)

$$EVSI_{(s)}|n_{(s)} = K_t \cdot \sqrt{Vn_{(s)}} \cdot \sigma_{0(s)} \cdot L(D_{(s)}|n_{(s)}) \qquad 5.2.4a$$

$$D_{(s)}|n_{(s)} = (\delta_{0(s)}-\delta_b)/\sqrt{Vn_{(s)}}$$

$$\sqrt{Vn_{(s)}} = \sigma_{0(s)}^2 / (\sigma_{0(s)}^2 + \sigma_{n(s)}^2)$$

$\delta_{0(s)}$ = prior incremental net benefit at stage s

$\sigma_{0(s)}^2$ = prior variance of $\delta_{0(s)}$ at stage s

$\sigma_{n(s)}^2$ = sample variance of $\delta_{(s)}$ with sample size $n_{(s)}$ at stage s

The marginal cost of a sample entering stage s ($C_{m(s)}$) is the additional treatment costs (compared to current practice) of allocating patients to the alternatives at that stage. The expected net benefits given a sample of $n_{(s)}$ at stage s is the difference between the expected benefits ($EVSI_{(s)}|n_{(s)}$) and the total variable cost of sampling ($Cm_{(s)}.n_{(s)}$).

$$ENBS_{(s)}|n_{(s)} = EVSI_{(s)}|n_{(s)} - C_{m(s)}.n_{(s)} \qquad 5.2.4b$$

The ENBS for the decision problem is the sum of the $ENBS_{(s)}$ at each stage:

$$ENBS|n=\sum_{s=1}^{S} ENBS_{(s)}|n_{(s)} \qquad 5.2.4c$$

Sample size will be optimal (n^*) when the ENBS|n reaches a maximum. If the fixed cost of research is less than the ENBS|n^* then the research is cost-effective at this technically efficient scale. The difference between the ENBS|n^* and the fixed cost of research can be used to set priorities in research and development and those proposals where the additional net benefits of research are greatest should be implemented first. In this way allocative efficiency in research and development across different sequential clinical decision problems can be achieved. The second hurdle for the two-stage decision problem in figure 5.2.1 can be constructed using the approach detailed above. The problem is to establish the EVSI and ENBS at stage 2 and stage 1 for a range of possible sample sizes, and then select the sample size entering the trial at stage 1 which will generate the maximum expected net benefits of sampling.

EVSI at Stage 2

The EVSI at stage 2 is calculated for the number of patients entering the trial at stage 1 who are assigned to the no test arm and enter stage 2 ($n_{(2)} = (n_{(1)}/2)$). The $EVSI_{(2)}|n_{(2)}$ is calculated in the same way as the single-stage decision problem in chapter 4, and the from 5.2.4a

$$EVSI_{(2)}|n_{(2)} = K_t \cdot \sqrt{Vn_{(2)}} \cdot \sigma_{0(2)} \cdot L(D_{(2)}|n_{(2)}) \qquad 5.2.5a$$

The prior incremental net benefit of t_1 ($\delta_{0(2)}$) is the difference between the prior net benefit of t_1 and the prior net benefit of t_0.

$$\delta_{0(2)} = (E(U|t_1)\text{-}g.E(C|t_1)) - (E(U|t_0)\text{-}g.E(C|t_0)) \qquad 5.2.5b$$

Cost of Sampling at Stage 2

The marginal cost of sampling at this point is the additional costs of assigning half the trial entrants who enter stage 2 to t_1 rather than current practice (t_0).

$$C_{m(2)} = (E(C|t_1) - E(C|t_0))/2 \qquad 5.2.5c$$

As before it is assumed that the marginal reporting cost are negligible and initially the fixed cost of research are zero because the fixed element will have no impact on optimal sample size. The expected net benefits of sampling given a sample of $n_{(2)}$ entering stage 2 ($ENBS_{(2)}|n_2$) is simply the difference between the expected benefits and the total variable costs of sampling, and is identical to the single stage problem considered in chapter 4:

$$ENBS_{(2)}|n_{(2)} = EVSI_{(2)}|n_{(2)} - C_{m(2)}.n_{(2)} \qquad 5.2.5d$$

EVSI at Stage 1

The benefits of sampling at the initial diagnostic decision cannot be separated from the sample which enters stage 2 because the fixed allocation rule means that a sample of $n_{(1)}$ at stage 1 will generate a sample of $n_{(2)} = n_{(1)}/2$ at stage 2, with $n_{(2)}/2$ allocated to t_1 and $n_{(2)}/2$ allocated to t_0. So the $ENBS_{(1)}|n_{(1)}$ cannot be calculated simply based on prior mean and variance at stage 1 because this would assume that those allocated to the no test arm would not be allocated <u>between</u> t_1 and t_0 at stage 2 but would all be allocated to <u>either</u> t_1 or t_2. A sample at stage 1 implies acquiring information about the contingent treatment decision at stage 2 and this will change the prior information about the no test alternative at stage 1. It will change the expected prior net benefits of not testing and reduce the uncertainty surrounding the initial diagnostic decision because more will be known about the no test alternative.

This problem can be solved by making contingent sampling decisions at stage 2 before calculating The ENBS at stage 1. Considering a sample of n_1 at stage 1 implies a sample of $n_{(2)} = n_{(1)}/2$ at stage 2, and the expected net benefit and posterior variance from stage 2, with sample of $n_{(2)}$, is used as the prior mean and variance of the no test alternative at stage 1. Both the expected net benefits and

posterior variance from stage 2 depend on the size of the sample entering stage 2 and this approach is consistent with the principles of backward induction, where contingent sampling decision must be solved at stage 2 and expected posterior values calculated, before the value of sampling at stage 1 can be estimated. From 5.2.4a the $EVSI_{(1)}|n_{(1)}$ can be calculated as follows:

$$EVSI_{(1)}|n_{(1)} = K_t \cdot \sqrt{Vn_{(1)}} \cdot \sigma_{0(1)} \cdot L(D_{(1)}|n_{(1)}) \qquad 5.2.6a$$

The prior incremental net benefit at stage 1 ($\delta_{0(1)}$) is the difference between the prior net benefit of testing ($E(U|t_e)$-$g.E(C|t_e)$) and the expected net benefit of not testing ($E(U|nt_e)$-$g.E(C|nt_e)$), which is dependent on sample size.

$$\delta_{0(1)} = (E(U|t_e)-g.E(C|t_e)) - (E(U|nt_e)-g.E(C|nt_e)) \qquad 5.2.6b$$

Prior Net Benefit at Stage 1

The value of $E(U|nt_e)$-$g.E(C|nt_e)$ is dependent on the sample size at stage 2, because there is a chance that a sample of $n_{(1)}$ which implies a sample $n_{(2)}=n_{(1)}/2$ at stage 2 will generate a posterior mean net benefit which will lead to t_1 being selected with net benefits of $E(U|t_1)$-$g.E(C|t_1)$. There is also a chance that the same sample may generate a posterior mean which will lead to t_0 being selected with net benefits of $E(U|t_0)$-$g.E(C|t_0)$. The expected net benefits of not testing depends on the posterior values at stage 2 which are a combination of prior and sample information. In general a posterior mean at stage s ($\delta_{1(s)}$) is simply a weighted average of the prior ($\delta_{0(s)}$) and sample mean ($\delta_{x(s)}$) with the weights representing the informational content of each.

$$\delta_{1(s)} = (I_{0(s)}.\delta_{0(s)} + I_{x(s)}.\delta_{x(s)})/(I_{0(s)} + I_{x(s)}) \qquad 5.2.7a$$

$$I_{0(s)} = 1/\sigma_{0(s)}^2$$

$$I_{x(s)} = 1/\sigma_{n(s)}^2$$

Before sample information is available and when $\delta_{0(s)}<\delta_b$ the clinician should select t_0 based on the prior mean, but once sample information is available the clinician should change this prior decision and select t_1 if $\delta_{1(s)}>\delta_b$. The critical value of the sample mean ($\delta_{x(s)}^*$) is the sample mean which generates a posterior mean that changes the prior decision. This can be found by setting $\delta_{1(s)}=\delta_b$ and $\delta_{x(s)}=\delta_{x(s)}^*$ and rearranging 5.2.7a:

$$\delta_{x(s)}^* = ((I_{0(s)} + I_{x(s)})\delta_b - I_{0(s)}.\delta_{0(s)})/I_{x(s)} \qquad 5.2.7b$$

If the sample mean is greater than this critical value ($\delta_{x(s)}>\delta_{x(s)}^*$) the posterior mean will be greater than δ_b and the clinician should select t_1 at stage 2, but when $\delta_{x(s)}<\delta_{x(s)}^*$ the posterior mean is less than δ_b and the clinical should select t_0. When a sample enters stage 2 there is a probability that the sample mean will lead to t_1 being selected ($p(\delta_{x(2)}>\delta_{x(2)}^*)$), with prior net benefits of $E(U|t_1)$-g.$E(C|t_0)$. There is also a probability that the same sample will lead to t_0 being selected ($1-p(\delta_{x(2)}>\delta_{x(2)}^*)$) with prior net benefits of $E(U|t_0)$-g.$E(C|t_0)$. The expected net benefits of not testing given a sample of $n_{(2)}$ at stage 2 is the prior net benefits of the no test arm of the trial at stage 1, and in 5.2.6b:

$$E(U|nt_e)\text{-g.}E(C|nt_e) = p(\delta_{x(2)}>\delta_{x(2)}^*).(E(U|t_1)\text{-g.}E(C|t_1)) + 1\text{-}p(\delta_{x(2)}>\delta_{x(2)}^*).(E(U|t_0)\text{-g.}E(C|t_0)) \qquad 5.2.8a$$

Since the choice of sample size must be made before any sample information is available the null hypothesis is that the sample mean is normally distributed, centred on the prior mean, with sample variance of $\sigma_{n(2)}^2$:

$$p(\delta_{x(2)}>\delta_{x(2)}^*) = p(Z>((\delta_{0(2)}>\delta_{x(2)}^*)/\sigma_{n(2)})) \qquad 5.2.8b$$

Prior Variance at Stage 1

Sampling at stage 1 generates sample information at stage 2 which will reduce the

uncertainty surrounding the no test alternative at stage 1. The prior variance at stage 2 does not reflect this additional information and it is the posterior variance from stage 2 which is used to establish the prior variance of the expected net benefits of not testing at stage 1. Posterior variance is a combination of prior and sample variance and will be less than either the prior variance or the sample variance. The uncertainty surrounding the no test alternative is reduced by taking the sample information generated at stage 2 into account in this way. In general the posterior variance ($\sigma_{1(s)}^2$) with a sample of $n_{(s)}$ is a combination of prior ($\sigma_{0(s)}^2$) and sample variance ($\sigma_{n(s)}^2$):

$$\sigma_{1(s)}^2 = (\sigma_{0(s)}^2/(\sigma_{0(s)}^2 + (\sigma_{n(s)}^2))).(\sigma_{n(s)}^2) \qquad 5.2.9$$

The prior variance of $E(U|nt_e)$-g.$E(C|nt_e)$ at stage 1 is a combination of the posterior variance of $E(U|t_1)$-g.$E(C|t_1)$ and $E(U|t_0)$-g.$E(C|t_0)$ given a sample of $n_{(2)}=n_{(1)}/2$ entering stage 2 with $n_{(2)}/2$ allocated to t_1 and $n_{(2)}/2$ allocated to t_0. This approach ensures that the information generated at stage 2 is taken into account when calculating the benefits of sampling at stage 1. The population variance of $E(U|nt_e)$-g.$E(C|nt_e)$ is also dependent on the sample entering stage 2 because this determines the value of $p(\delta_{x(2)}>\delta_{x(2)}^*)$ which is regarded as a constant when calculating the population and the prior variance of $E(U|nt_e)$-g.$E(C|nt_e)$ at stage 1.

The $EVSI_{(1)}|n_{(1)}$ from 5.2.6a can now be established taking into account the relationship between the sample considered at stage 1 and sample information it will generate at stage 2. It measures the additional value of sample information at this initial decision given that a sample of $n_{(1)}$ at stage 1 will generate a sample of $n_{(2)}$ at stage 2, which will reduce the uncertainty surrounding the no test alternative and will also change the expected prior net benefits of choosing not to test. Increasing the sample considered at stage 1 will change the expected net benefits of not testing from stage 2 and therefore change $\delta_{0(1)}$. It will also reduce the posterior variance from stage 2 and therefore reduce the prior variance at stage 1. So unlike the single-stage decision problem, the prior mean and prior variance for

the initial decision is not independent of the sample size considered.

Cost of Sampling at Stage 1

The marginal cost of sampling at stage 1 is the additional treatment costs of allocating trial entrants to the testing arm of the trial. The marginal cost of allocating a trial entrant to the no test arm at stage 1 is zero because the additional treatment cost of assigning patients to t_1 at stage 2 has already being taken into account in the calculation of the $ENBS_{(2)}|n_{(2)}$. So the marginal cost of sampling at the initial diagnostic decision is the additional cost of assigning half the entrants to the test alternative:

$$C_{m(1)} = (E(C|t_e)-E(C|t_0))/2 \qquad 5.2.10$$

The expected net benefit of sampling at the initial diagnostic decision is simply the difference between the expected benefits and the total variable cost of sampling:

$$ENBS_{(1)}|n_{(1)} = EVSI_{(1)}|n_{(1)} - C_{m(1)}.n_{(1)} \qquad 5.2.11$$

The $ENBS_{(1)}|n_{(1)}$ is the net benefit of comparing the testing strategy to the no test strategy using sample information, given that a sample at stage 1 will generate a sample of $n_{(2)}$ at stage 2. The sample generated at stage 2 will provide net benefits from comparing t_1 to t_0 using sample information and this is measured by $ENBS_{(2)}|n_{(2)}$, but sampling at stage 2 will also provide information about the net benefit of the no test alternative at stage 1. The methods outlined above ensures that the ENBS for the full decision problem can be estimated by the sum of $ENBS_{(1)}|n_{(1)}$ and $ENBS_{(2)}|n_{(2)}$ without the danger that the benefits of sampling will be overestimated by double counting benefits at stage 1 and 2, or underestimating the benefits of proposed research by simplifying this two stage problem to a single stage problem and only comparing the testing strategy to either t_1 or t_0.

ENBS and Optimal Sample Size

The ENBS for both stages of this decision problem is illustrated in figure 5.2.5a when the value of 1/g=£4,000 and for a range of possible sample sizes entering stage 1. The ENBS is the sum of the net benefits at stage 1 and 2 and reaches a maximum at $n^*=92$. At this optimal sample size 46 entrants will be allocated to the test and 46 to the no test arm at stage 1, and 23 will be allocated to t_1 and 23 to t_0 at stage 2. The maximum $ENBS|n^*$ =£442,340, and this is the second hurdle that proposed research must overcome. If the fixed cost of the research is less than the expected net benefit then the proposed research is cost-effective when conducted at this technically efficient scale.

Figure 5.2.5a

The expected net benefits of sampling are greater at stage 2 than stage 1, indeed the $ENBS_{(1)}|n_{(1)} <0$. This is because the marginal sampling cost at stage 1 is high, due to the additional costs of the diagnostic test, and because at 1/g=£4,000 the prior decision not to test is less uncertain than the prior decision to choose t_0 at stage 2 ($L(D_{0(2)})>L(D_{0(1)})$ in figure 5.2.3b). Although the $ENBS_{(1)}|n_{(1)} <0$ it is still worth taking a sample at stage 1 because it enables a sample to enter stage 2 which will produce positive net benefits of $ENBS_{(2)}|n_{(2)}$. The $ENBS_{(2)}|n_{(2)}$ is identical to the expected net benefits of the single stage treatment decision problem considered in chapter 4, but because of the negative net benefits at stage 1 the optimal sample size of 46 at stage 2 is less than the optimal sample size for the single-stage decision problem.

Value of Health Outcome

The expected net benefit of sampling and optimal sample size is dependent on the value placed on health outcome. In this example when 1/g is higher at £10,000, the expected net benefit and the optimal sample size is also higher. This is

illustrated in figure 5.2.5b where the ENBS reaches a maximum at $n^*=260$. The fixed allocation rule dictates that 130 entrants will be allocated to the testing and 130 to the no test arm, so that $n_{(2)}^*=130$, and 70 entrants will be allocated to t_1 and 70 to t_0 at stage 2. Now the expected net benefits of sampling are greater at stage 1 than at stage 2 because there is less weight attached to the higher cost of sampling at stage 1 and the prior decision not to test is less certain than the prior decision to select t_1 at stage 2 ($L(D_{0(1)})>L(D_{0(2)})$ in figure 5.2.3b). The $ENBS_{(2)}|n_{(2)}$ reaches a maximum when n=376, but the $ENBS_{(1)}|n_{(1)}$ reaches a maximum at n=201 due to the higher marginal sampling cost.

Figure 5.2.5b

The expected net benefits of sampling when the value of 1/g is increased to £20,000 is illustrated in figure 5.2.5c. Optimal sample size is greater ($n^*=372$) and the maximum ENBS|n* is also higher. Again the expected net benefits at stage 1 are substantially greater than net benefits at stage 2, because even less weight is placed on the higher marginal sampling costs at stage 1, and because the prior decision (which is now to test), is less certain than the prior decision to treat at stage 2 ($L(D_{0(1)})>L(D_{0(2)})$ in figure 5.2.3b).

Figure 5.2.5c

In this example of a sequential decision problem optimal sample size increases with the value placed on health outcome. The relationship between 1/g and optimal sample size is illustrated in figure 5.2.6. Optimal sample size with a more sceptical ($n_0=2$) and a less sceptical prior ($n_0=18$) is also illustrated in figure 5.2.6. When the value of health outcome is low the optimal sample size is zero and decisions should be based only on prior information, which in this example would be not to test and not to treat. Sampling becomes optimal at lower values of 1/g when the prior is more sceptical (if $n_0=2$ then $n^*=0$ when 1/g<£2,000) because the prior decision is more uncertain and less information is required to change the

prior decision than with a more confident prior (if n_0=18 then n*=0 when 1/g<£5,000).

Figure 5.2.6

The discontinuities in the relationship between 1/g and n^* when n_0=2, and n_0=6 is at the point where the clinician is indifferent between the testing and the expected net benefits of not testing from stage 2. This point changes with n* because the expected net benefit of not testing depends on the sample generated at stage 2. The discontinuity in the relationship between n^* and 1/g when n_0=18 occurs when the net benefits of sampling at stage 2 are less than zero.

Once optimal sample size is established for each value of 1/g the relationship between the ENBS|n^* and the value of 1/g can be considered. Figure 5.2.7a illustrates this relationship for each stage of the decision problem. The ENBS|n^* is the second hurdle that proposed research must overcome before it can be considered cost-effective, and figure 5.2.7a demonstrates that the second hurdle is sensitive to the value of 1/g and budgetary restrictions on service provision. There are two discontinuities in the relationship between ENBS|n* and 1/g. The $ENBS_{(2)}|n_{(2)}^*$ reaches a peak when the clinician is indifferent between t_1 and t_0 based only on prior information. The discontinuity in the $ENBS_{(1)}|n_{(1)}^*$ occurs when the clinician is indifferent between the test and no test alternatives which (from 5.2.8a) is partly determined by sample size.

Figure 5.2.7a

The ENBS|n^* for three different priors is illustrated in figure 5.2.7b. The ENBS|n^* increases with the value placed in health outcome, but the value of sample information is also sensitive to the confidence in prior information. When the prior is more sceptical (n_0=2) then the expected cost of uncertainty surrounding this decision problem is higher and the value placed on additional sample information

is also high. However when there is more confidence in prior information (n_0=18) the cost of uncertainty is lower and the value of additional sample information is also low. This confirms the interpretation placed on the EVSI that the benefit of sample information is the reduction in the costs of uncertainty surrounding a decision problem.

Figure 5.2.7b

5.2.3 Implications for Research Design

This numerical example has demonstrated that the expected net benefits of sample information and the optimal sample size can be established and the second hurdle can be constructed for a sequential decision problem. An analysis which did not recognise the sequential nature of the decision problem by only considering the single-stage treatment decision and excluding the diagnostic process would bias efficient research design and cause errors at the second hurdle. This illustrates the dangers of ruling out alternatives from consideration based on implicit decision rules and inconsistent judgements. For example if the value of 1/g is £6,500 then the clinician would be indifferent between t_1 and t_0 at stage 2. The clinician is most uncertain about this treatment decision and prospective research may only be considered for this single-stage treatment decision. The diagnostic process may be excluded and regarded as not relevant because prior information suggests that testing will not be cost-effective (prior cost-effectiveness ratio = £11,800). However by arbitrarily excluding the diagnostic process from prospective research the expected net benefits of sample information and optimal sample size will be underestimated. This is illustrated in figure 5.2.8a and figure 5.2.8b where the ENBS|n* and n* for this sequential problem are compared to the single stage treatment decision considered in chapter 4 .

Figure 5.2.8a

Figure 5.2.8b

The ENBS|n^* for this two-stage problem is greater than the single-stage problem, and excluding the diagnostic process from the design of a proposed trial will underestimate the expected net benefits of sample information and the optimal scale of the research. The second hurdle will be biased and there will be a danger that cost-effective research will be rejected at the second hurdle. The optimal sample size will also be underestimated and research design will be technically inefficient. The bias in the estimates of the expected net benefits of sampling may also lead to inefficient allocation between clinical decision problems and the allocation of research and development resources to technically inefficient designs.

5.3 A Four-Stage Sequential Clinical Decision Problem

The same approach to the value of information can be applied to more complex decision problems, and in this section the first and second hurdles can be constructed for a four-stage decision problem. This is the same problem which was used to generalise the Phelps and Mushlin strategy for the economic evaluation of diagnostic information in chapter 3. This decision problem is identical to the two-stage problem in section 5.2 except that an alternative treatment t_2 is available.

Figure 5.3.1

The decision problem is illustrated in figure 5.3.1 and is structured to follow the correct chronology and includes Bayesian probability revision. Now that an alternative treatment (t_2) is available the clinician must choose between two testing and three fallback strategies. At stage 4 the clinician must choose to treat with either t_1 or t_2 following a positive test result, and those with negative test results are not treated (t_0). If the clinician decides not to test at stage 1 they must decide whether to treat (t_r) or not treat (t_0) at stage 2, and if they decide to treat they must then decide whether to treat with either t_1 or t_2 at stage 3. In this section the same numerical example as chapter 3 is used but with prior distributions for the health outcomes and path probabilities. These are reported in table 5.3. Again the assumptions of normality and independence allow normal loss functions to be used and covariance terms to be zero when calculating the prior and population variance at each stage.

Table 5.3

The problem is to construct the first hurdle for this decision problem by establishing the expected value of perfect information at each stage. If proposed research passes this first hurdle the expected net benefits of sample information

must be established and the technically efficient scale of the research must be identified to construct the second hurdle .

5.3.1 Hurdle I: The Expected Value of Perfect Information

The clinician is faced with an uncertain choice between two alternatives at each stage of this decision problem. Without any additional information these choices must be made based on prior information and there will be a chance that opportunity losses will be incurred. The expected cost of uncertainty at each stage is the expected value of perfect information and the EVPI for this four-stage decision problem will be the sum of the $EVPI_{(s)}$ at each stage (s=1, .. ,4).

EVPI at Stage 4

If the clinician decides to use the test a contingent treatment decision must be made at stage 4 in figure 5.2.1, because now if a patient has a positive test result the clinician must decide to treat with either t_1 or t_2 (prior CER=£5,800). This contingent decision is based on the prior incremental net benefit of t_1 ($\delta_{0(4)}$) and is the difference between the prior net benefit of t_1 given a positive test result ($E(U|t_e^+, t_1)$-g.$E(C|t_e^+, t_1)$), and the prior net benefit of t_2 given a positive test result ($E(U|t_e^+, t_2)$-g.$E(C|t_e^+, t_2)$).

$$\delta_{0(4)} = (E(U|t_e^+, t_1)\text{-g.}E(C|t_e^+, t_1))-(E(U|t_e^+, t_2)\text{-g.}E(U|t_e^+, t_2)) \quad 5.3.1$$

EVPI at Stage 3

The clinician also faces an uncertain contingent decision at stage 3 and must choose either t_1 or t_2. This decision is based on the prior incremental net benefits of t_1 at stage 3 ($\delta_{0(3)}$), and is the difference between the prior net benefits of t_1 ($E(U|t_1)$-g.$E(C|t_1)$) and the prior net benefits of t_2 ($E(U|t_2)$-g.$E(U|t_2)$) (prior

CER=£10,000).

$$\delta_{0(3)} = (E(U|t_1)\text{-}g.E(C|t_1))\text{-}(E(U|t_2)\text{-}g.E(U|t_2)) \qquad 5.3.2$$

EVPI at Stage 2

If the clinician decides not to test they also face a choice between treatment (t_r), following the contingent treatment decision at stage 3, and no treatment (t_0) at stage 2. This contingent decision is also uncertain and opportunity losses may be incurred. The $EVPI_{(2)}$ at stage 2 is in addition to the $EVPI_{(3)}$ because t_0 must be compared with either t_1 or t_2 depending on the contingent decision made at stage 3. The decision at stage 2 is based on the prior incremental net benefits of treatment ($\delta_{0(2)}$) which is the difference between the prior net benefits of treatment ($E(U|t_r)\text{-}g.E(C|t_r)$), and the prior net benefits of no treatment ($E(U|t_0)\text{-}g.E(U|t_0)$).

$$\delta_{0(2)} = (E(U|t_r)\text{-}g.E(C|t_r))\text{-}(E(U|t_0)\text{-}g.E(U|t_0)) \qquad 5.3.3$$

EVPI at Stage 1

Finally the clinician faces an uncertain choice at the initial diagnostic decision. At stage 1 the clinician must choose whether to test and follow the contingent decision at stage 4, or choose not to use the test and follow the contingent treatment decisions at stage 2 and stage 3. The uncertainty surrounding this initial decision is in addition to the uncertainty surrounding each of the three contingent decision problems at stages 2, 3, and 4. The decision is based on the prior incremental net benefit of testing ($\delta_{0(1)}$) which is the difference between the prior net benefits of testing ($E(U|t_e)\text{-}g.E(C|t_e)$) given the contingent decisions at stage 4 and the prior net benefits of not using the test ($E(U|nt_e)\text{-}g.E(C|nt_e)$) given the contingent decisions at stage 2 and stage 3.

$$\delta_{0(1)} = (E(U|t_e)\text{-}g.E(C|t_e)) - (E(U|nt_e)\text{-}g.E(C|nt_e)) \qquad 5.3.4$$

EVPI for the Four-Stage Problem

The EVPI for this sequential decision problem is the sum of the $EVPI_{(s)}$ at each stage (each point where the clinician faces an uncertain decision) and this is illustrated for this numerical example in figure 5.3.2a. The EVPI rises with the value of 1/g because the value placed on opportunity losses increases (K_t =2/g). This is the first hurdle for this decision problem and if the EVPI exceeds the estimated cost of proposed research then it is potentially cost-effective and passes the first hurdle. Figure 5.3.2a. demonstrates that this first hurdle is sensitive to the value of health outcome, and (just as in chapter 4 and section 5.2) the value of information is dependent on budgetary restrictions on service provision. If the budget constraint is relaxed then the value of 1/g and the value of information will increase, similarly when the budget constraint is tightened the value of 1/g will fall and the value of information will also fall.

Figure 5.3.2a

This approach allows the value of perfect information to be identified at particular points in a sequential decision. This can be used to indicate where additional information about particular contingent decisions will be valuable. Figure 5.3.2b illustrates the $EVPI_{(s)}$ at each stage in this decision problem and it demonstrates that the points in a sequential decision problem where information is most valuable will depend on the value of 1/g. In this numerical example when 1/g is low (£3,000) the expected cost of uncertainty is highest at stage 2 where the clinician must choose between t_0 and t_2. But when the value of 1/g is higher (£14,000) the point where information may be most valuable will now be at stage 1 where the clinician faces a choice between using the test and treatment without testing.

Figure 5.3.2b

The relationship between the $EVPI_{(s)}$ at each stage and the value of 1/g contains a

number of discontinuities at the points where the prior contingent decisions change. The discontinuity in the $EVPI_{(4)}$ occurs at 1/g=£5,800 which is where the clinician will be indifferent between t_1 and t_2 at stage 4. At this point $D_{0(4)}=0$ in figure 5.3.3a and $L(D_{0(4)})$ reaches a maximum in figure 5.3.3b. The discontinuity in the $EVPI_{(3)}$ occurs at 1/g=£10,000 where the clinician is indifferent between t_1 and t_2 at stage 3. At this point the $D_{0(3)}=0$ in figure 5.3.3a and the $L(D_{0(3)})$ reaches a maximum in figure 5.3.3b. There are two discontinuities in the $EVPI_{(2)}$. The first occurs when 1/g=£3,000 where the clinician is indifferent between t_0 and t_2. At this point $D_{0(2)}=0$ in figure 5.3.3a and $L(D_{0(2)})$ reaches a maximum in figure 5.3.3b. The second discontinuity occurs at 1/g=£10,000 when the decision facing the clinician at stage 2 changes from a choice between t_0 and t_2 to a choice between t_0 and t_1. The discontinuities the $EVPI_{(1)}$ reflect the changes in the contingent decisions at stages 2, 3, and 4 outlined above, but in addition there is a discontinuity in the $EVPI_{(1)}$ where the clinician will be indifferent between testing and not testing at 1/g=£11,800. At this point the initial diagnostic decision is most uncertain and $D_{0(1)}=0$ in figure 5.3.3a and the $L(D_{0(1)})$ reaches a maximum in figure 5.3.3b.

Figure 5.3.3a

Figure 5.3.3b

The EVPI depends on the confidence in prior information. Figure 5.3.4 illustrates the EVPI for 3 different levels of confidence in prior information. When the prior is more sceptical (the quasi sample size is lower ($n_0=2$)) then the uncertainty surrounding each decision will be greater and the EVPI is higher. Similarly when the prior is less sceptical (the prior sample size is higher ($n_0=18$)) then the uncertainty surrounding each decision will be lower and the EVPI is also lower.

Figure 5.3.4

The EVPI for this four-stage decision problem and for the two-stage decision

problem discussed in section 5.2 is illustrated in figure 5.3.5. These decision problems are identical except that treatment alternative t_2 is not available in the two-stage problem. An analysis of this clinical problem which simplified what is a four-stage decision problem to a two-stage problem by excluding t_2 as a relevant alternative would clearly underestimate the EVPI. The first hurdle would be biased and there would be a danger that potentially cost-effective research would be rejected. In this numerical example if the value of 1/g is greater than £10,000 then t_2 will not be selected at any stage based on prior information. In these circumstances it may be tempting to conclude that t_2 is not a relevant alternative and can be excluded from the analysis. However this process of simplification and arbitrarily excluding feasible alternatives as not relevant may introduce serious bias into the analysis. It is possible that a feasible alternative is not relevant if it will never be selected. This could be established by comparing the EVPI with and without the alternative and if the alternative is not relevant then both estimates of the EVPI should coincide.

Figure 5.3.5

5.3.2 Hurdle II: The Expected Net Benefit of Sample Information

If the EVPI exceeds the estimated costs of proposed research then it is potentially cost-effective and the expected net benefits of sample information must be established to construct the second hurdle and identify the optimal sample size. The second hurdle can be constructed for this four-stage decision problem using the same methods which were used to construct the second hurdle for the two-stage decision problem in sections 5.2. The same fixed allocation rule is used at each stage where equal numbers of trial entrants are assigned to each arm at each stage. Following the principles of backward induction the expected net benefits and the posterior variance at each contingent decision is uaed as the prior information at earlier stages.

The expected net benefits of sampling for this decision problem is the sum of the $ENBS_{(s)}$ at each stage, taking into account that sampling at an early stage implies a sample at later stages which will reduce the uncertainty surrounding the alternatives considered at these earlier stages. The optimal number of trial entrants at stage 1 is the sample size that maximises the ENBS and it is this maximum value which is used as the second hurdle.

ENBS at Stage 4

At stage 4 the clinician must decide whether to use treatment t_1 or t_2 following a positive test result. The fixed and equal allocation of trial entrants means that if a sample of $n_{(1)}$ enters stage 1 then the number of entrants entering stage 4 will be $n_{(4)}= (n_{(1)}/2).p(t_e^+)$. The $EVSI_{(4)}|n_{(4)}$ given a sample of $n_{(4)}$ entering stage 4 can be calculated and the marginal cost of sampling at stage 4 is the additional cost (compared to t_0) of assigning half the entrants to t_1 and half to t_2 .

$$C_{m(4)} = (E(C|t_1)-E(C|t_0))/2 +(E(C|t_2)-E(C|t_0))/2 \qquad 5.3.5$$

The costs of testing at this point are sunk, but will be taken into account when estimating the expected net benefits of allocating entrants to the testing arm at stage 1.

ENBS at Stage 3.

At stage 3 the clinician must choose between t_1 and t_2. The fixed and equal allocation of trial entrants means that a sample of $n_{(1)}$ at stage 1 generates a sample of $n_{(2)}= (n_{(1)}/2)$ entering stage 2, with $(n_{(2)}/3).2=n_{(3)}$ assigned to the treatment arm and entering stage 3. This allocation rule ensures that equal numbers of entrants will be allocated to t_0, t_1, and t_2. The $EVSI_{(3)}|n_{(3)}$ given a sample $n_{(3)}$ entering stage 3 can be calculated and the marginal sampling cost is the additional treatment cost of assigning half the entrants at stage 3 to t_1 and half to t_2.

$$C_{m(3)} = (E(C|t_1)-E(C|t_0))/2 + (E(C|t_2)-E(C|t_0))/2 \qquad 5.3.6$$

ENBS at Stage 2

At stage 2 the clinician must decide whether to treat (tr) and follow the contingent treatment decision at stage 3 or not treat (t_0). If $n_{(1)}$ patients enter the trial at stage 1 then $n_{(2)}=n_{(1)}/2$ will enter stage 2 with $(n_{(2)}/3)2$ assigned to tr and $n_{(2)}/3$ will be allocated to t_0. The $EVSI_{(2)}|n_{(2)}$ must take into account that a sample of $n_{(2)}$ at stage 2 implies a sample of $n_{(3)}=(n_{(2)}/3)2$ at stage 3. This will change the expected net benefits of t_r and will reduce the uncertainty surrounding the treatment arm at stage 2. The $EVSI_{(2)}|n_{(2)}$ can be established using the same approach that was used at stage 1 in section 5.2. The prior incremental net benefit of selecting t_r ($\delta_{0(2)}$) is the difference between the prior net benefit of treatment ($E(U|t_r)$-g.$E(C|t_r)$) and the prior net benefit of t_0 ($E(U|t_0)$-g.$E(Ct_0)$). The expected net benefit of treatment and therefore $\delta_{0(2)}$ is dependent on sample size because for each sample considered at stage 2 there is a probability that the sample which enters stage 3 will generate a posterior mean which will lead to t_1 being selected ($p(\delta_{x(3)}>\delta_{x(3)}^{*})$) with net benefits of ($E(U|t_1)$-g.$E(C|t_1)$). There is also a probability that the same sample will lead to t_2 being selected ($1-p(\delta_{x(3)}>\delta_{x(3)}^{*})$) at stage 3 with net benefits of ($E(U|t_2)$-g.$E(C|t_2)$).

$$(E(U|t_r)-g.E(C|t_r)) = p(\delta_{x(3)}>\delta_{x(3)}^{*}).(E(U|t_1)-g.E(C|t_1)) + 1-p(\delta_{x(3)}>\delta_{x(3)}^{*}).(E(U|t_2)-g.E(C|t_2)) \qquad 5.3.7$$

$$p(\delta_{x(3)}>\delta_{x(3)}^{*}) = p(Z>((\delta_{0(3)}-\delta_{x(3)}^{*})/(\sigma_{n(3)})))$$

The prior net benefits of the treatment arm at stage 2 ($E(U|t_r)$-g.$E(C|t_r)$) is the expected net benefits from stage 3 and depends on the sample size considered at stage 2. The prior variance of $E(U|t_r)$-g.$E(C|t_r)$ at stage 2 is dependent on the sample assigned to the treatment arm which will enter stage 3, because it is a combination of the posterior variance of $E(U|t_1)$-g.$E(C|t_1)$ and $E(U|t_2)$-g.$E(C|t_2)$ at stage 3 given a sample $n_{(3)}=(n_{(2)}/3)2$. The population variance of $E(U|t_r)$-g.$E(C|t_r)$

is also dependent on the sample entering stage 3 because this determines the value of $p(\delta_{x(3)}>\delta_{x(3)}^{*})$.

The marginal sampling cost at stage 2 is zero because the additional treatment cost of assigning an entrant to t_0 is zero and the additional costs of assigning an entrant to the treatment arm is also zero because the additional costs of treatment with t_1 or t_2 have been taken into account at stage 3.

ENBS at Stage 1

At the initial diagnostic decision the clinician must decide whether to test (t_e) and follow the contingent treatment decision at stage 4, or not test (nt_e) and follow the contingent treatment decisions at stage 2 and 3. The $EVSI_{(1)}|n_{(1)}$ is the additional benefit of sample information at stage 1 given that each sample considered will generate sample information at the contingent treatment decisions. A sample of $n_{(1)}$ implies a sample of $n_{(2)}=n_{(1)}/2$ entering stage 2 which will change the expected net benefits of not testing and reduce the uncertainty surrounding the no test arm. It also implies a sample of $n_{(4)}= (n_{(1)}/2).p(t_e^{+})$ entering stage 4 which will change the prior expected net benefits of the test arm and will reduce the uncertainty surrounding the testing arm at stage 1. The reduction in uncertainty surrounding this initial decision due to sampling at contingent decisions will reduce the additional benefits of sample information at this stage. This can be taken into account in the same way as section 5.2 by using the expected net benefit and posterior variance from stage 4 and stage 2 as the prior information at stage 1.

The prior incremental net benefit of testing ($\delta_{0(1)}$) is dependent on sample size because it is the difference between the prior net benefits of testing ($E(U|t_e)$-$g.E(C|t_e)$, which is partly determined by sample entering stage 4, and the prior net benefit of not testing ($E(U|nt_e)$-$g.E(C|nt_e)$), which is also determined by sample entering stage 2.

Each sample considered at stage 1 generates a sample of $n_{(2)}=n_{(1)}/2$ at stage 2 and there is a probability that this sample will generate a posterior mean at stage 2 which will lead to t_r being selected ($p(\delta_{x(2)}>\delta_{x(2)}^{*})$) with net benefits of ($E(U|t_r)$-$g.E(C|t_r)$). There is also a probability that the same sample will generate a posterior mean which will lead to t_0 being selected at stage 2 ($1-p(\delta_{x(2)}>\delta_{x(2)}^{*})$) with net benefits of ($E(U|t_0)$-$g.E(C|t_0)$).

$$E(U|nt_e)-g.E(C|nt_e) = p(\delta_{x(2)}>\delta_{x(2)}^{*}).(E(U|t_r)-g.E(C|t_r)) \qquad 5.3.8a$$
$$+1-p(\delta_{x(2)}>\delta_{x(2)}^{*}).(E(U|t_0)-g.E(C|t_0))$$
$$p(\delta_{x(2)}>\delta_{x(2)}^{*}) = p(Z>((\delta_{0(2)}>\delta_{x(2)}^{*})/(\sigma_{n(2)})))$$

Each sample at stage 1 generates a sample of $n_{(4)}=(n_{(1)}/2).p(t_e^{+})$ at stage 4 which will lead to t_1 being selected following a positive result ($E(U|t_e^{+},t_1)$-$g.E(C|t_e^{+},t_1)$) with a probability of $p(\delta_{x(4)}>\delta_{x}^{*}{}_{(4)})$, and t_2 will be selected following a positive result ($E(U|t_e^{+},t_2)$-$g.E(U|t_e^{+},t_2)$) with a probability of $1-p(\delta_{x(4)}>\delta_{x}^{*}{}_{(4)})$. No treatment follows a negative test result ($p(t_e^{-})$) with net benfits of ($E(U|t_e^{-},t_0)$-$g.E(C|t_e^{-},t_0)$).

$$E(U|t_e)-g.E(C|t_e) = p(t_e^{-}).(E(U|t_e^{-},t_0)-g.E(C|t_e^{-},t_0)) \qquad 5.3.8b$$
$$+p(t_e^{+}).(p(\delta_{x(4)}>\delta_{x(4)}^{*}).(E(U|t_e^{+},t_1)-g.E(C|t_e^{+},t_1))$$
$$+1-p(\delta_{x(4)}>\delta_{x(4)}^{*}).(E(U|t_e^{+},t_2)-g.E(U|t_e^{+},t_2)))$$
$$p(\delta_{x(4)}>\delta_{x(4)}^{*}) = p(Z>((\delta_{0(4)}>\delta_{x(4)}^{*})/(\sigma_{n(4)})))$$

The marginal sampling cost at stage 1 only includes the cost of the diagnostic test (C_{te}) because the additional costs of treatment are taken in to account in the ENBS at stage 3 and at stage 4, and the additional cost of t_0 for those with negative test results is zero.

$$C_{m(1)} = (C_{te}-E(C|t_0))/2 \qquad 5.3.8c$$

ENBS and Optimal Sample Size

The expected net benefits of sample information for this decision problem is the sum of the ENBS at each stage. The approach taken at stage 1 and 2 of using the expected net benefits from contingent decisions as the prior information at these earlier stages ensures that the benefits of sampling at each stage will not be overestimated. The ENBS for this numerical example is illustrated in figure 5.3.6a when the value of 1/g=£4,000. The ENBS reaches a maximum (ENBS|n*) at an optimal sample size of n*=191 and is the second hurdle that the this proposed research must overcome before it can be considered cost-effective. At this value of 1/g the ENBS at stage 2 is greater than at the other stages. In particular it is greater than at stage 1 where the $ENBS_{(1)}<0$. This is because of the high marginal sampling cost at stage 1 but also because the prior decision not to test is less uncertain than the decision to treat with t_2 at stage 2 ($L(D_{0(2)}) > L(D_{0(1)})$ in figure 5.3.3b). Although the $ENBS_{(1)} <0$ it is still worth taking at sample at stage 1 because it means that the positive net benefits at the contingent decisions can be realised.

Figure 5.3.6a

The ENBS and optimal sample size for this decision problem depends on the value of 1/g, and figure 5.3.6b illustrates the ENBS for this numerical example when the value of 1/g is higher at £10,0000. In this case the optimal sample size is higher (n^*=432) and the ENBS|n* is also higher. Now the ENBS at stage 3 and stage 1 is higher than at stage 2. This is because less weight is placed on the marginal sampling cost at stage 1, and because the prior decision to treat at stage 2 is less uncertain than the decision not to use the test at stage 1 ($L(D_{0(2)})<L(D_{0(1)})$ in figure 5.3.3b) and it is also less uncertain than the choice between t_1 and t_2 at stage 3 ($L(D_{0(2)}) < L(D_{0(3)})$ in figure 5.3.3b).

Figure 5.3.6b

The ENBS when the value of 1/g is increased to £20,000 is illustrated in figure 5.3.6c. The optimal sample size is greater (n^*=666) and the maximum ENBS is also higher. The highest $ENBS_{(s)}$ is at stage 1 and the lowest is at stage 2, this is because the prior decision to treat at stage 2 is less uncertain than the prior decision at stage 1, which is now to use the test ($L(D_{0(2)}) < L(D_{0(1)})$ in figure 5.3.3b).

Figure 5.3.6c

These examples demonstrate that it is possible to construct the second hurdle for more complex sequential decision problems and this hurdle operates in the same way as in the previous section and in chapter 4. If the ENBS|n* exceeds the fixed cost of proposed research then it is cost-effective when conducted at the technically efficient scale. Similarly the difference between the ENBS and the fixed cost of research can be used to set priorities across research proposal that pass the second hurdle. If those proposals where the difference between ENBS|n* and fixed cost are greatest are implemented first then the maximum benefits can be gained for a fixed research and development budget. So again this approach and construction of the second hurdle can be used to ensure the optimal allocation of research and development resources among technically efficient research designs.

The relationship between the value of 1/g and the optimal sample size for this numerical example is illustrated in figure 5.3.7. Optimal sample size is illustrated for a range of values of 1/g and for three quasi prior samples (n_0=2, 6, 18). This illustrates that when the value of 1/g is low there will be a point where the optimal sample size is zero and decisions should be based only on prior information which in this example would be not to treat. However sampling becomes optimal at lower values of 1/g when the prior is more sceptical (n_0=2) just as in the analysis of the two stage decision problem . This is because if there is less confidence in prior information then the prior decision will be more uncertain and the benefits of

sample information will be greater. In this numerical example the optimal sample size increases with 1/g and is not particularly sensitive to the strength of the prior information. The discontinuities in this relationship occur where the clinician is indifferent between the alternatives at each stage. These points of indifference at stage 1 and stage 2 where information will be most valuable change with sample size.

Figure 5.3.7

Once the optimal sample size has been established the relationship between ENBS|n^* and the value of 1/g can be illustrated in figure 5.3.8a. The discontinuities in the relationship between the value of 1/g and the ENBS are at those values of 1/g where the clinician would be indifferent between the alternatives. For example the discontinuity in the $ENBS_{(3)}$ occurs when 1/g =£10,000 and the clinician is indifferent between treatment t_1 and t_2. Figure 5.3.8a also illustrates that the stages where information will be most valuable will also depend on the value of 1/g. When the value of 1/g is high (greater than £11,000) the ENBS is highest at stage 1 and lowest at stage 2, however if the value of 1/g is lower (less than £6,000) then the ENBS is highest at stage 2 and lowest at stage 1. This approach can identify those points in sequential decision problems where information will be most valuable and this can be used to set priorities in acquiring additional information, however the point at which information is most valuable will depend on the value of 1/g.

Figure 5.3.8a

Figure 5.3.8b

The ENBS|n^* for this decision problem using three quasi prior samples (n_0=2, 6, 18) is illustrated in figure 5.3.8b, and demonstrates that the second hurdle is also sensitive to the quality or strength of prior information. When the prior is more sceptical (n_0=2) then the ENBS is higher at each value of 1/g because the prior

decisions are more uncertain and additional information will be more valuable. Similarly when the prior is less sceptical the ENBS is lower for each value of 1/g, because the prior decisions are less uncertain and the value of additional information will be lower.

5.3.3 Implications for Research Design

This numerical example demonstrates that this approach can be applied to more complex sequential decision problems. This four-stage decision problem is identical to the two stage decision problem in considered in section 5.2 of this chapter except that an alternative treatment t_2 is available. An analysis which either did not recognise that t_2 is a feasible alternative or ruled it out as not relevant would bias the second hurdle and lead to errors in efficient research design by excluding arms of the trial, and biasing the optimal sample size.

These dangers can be illustrated by comparing the ENBS$|n^*$ for the four-stage problem considered in this section and the ENBS$|n^*$ for the two-stage problem considered in section 5.2. This comparison is illustrated in figure 5.3.9a. The ENBS for the two-stage problem represents the results of an analysis which has ruled out t_2 as an irrelevant alternative. Treatment t_2 may have been regarded as not relevant because it may not be part of current practice. If the value of 1/g is greater than £10,000 it will not be selected at any stage based on prior information, and if purely clinical decision rules are used to select existing patient management strategies it will never be selected at any prior probability of disease. This was discussed in more detail in chapter 2. However if t_2 is excluded for what ever reason figure 5.3.9a demonstrates that the ENBS will be underestimated, the second hurdle will be biased. There will be a danger that a cost-effective research proposal will be rejected at this second hurdle.

Figure 5.3.9a

Figure 5.3.9b

The design of the proposed trial will be inefficient and will exclude the random allocation of trial entrants at stage 4 and stage 3. The optimal sample size will also be biased and this is illustrated in figure 5.3.9b where the optimal sample size for the four stage problem is greater than for the two stage problem which excludes t_2. Once again this illustrates that arbitrarily excluding feasible alternatives and using implicit rules and inconsistent judgements to identify which alternatives are regarded as relevant will bias research design. In this example the value of information and the technically efficient scale of proposed research will be underestimated. There is a danger that cost-effective proposals will be rejected and those that are accepted will be conducted at less than the technically efficient scale.

5.4 Conclusions

The examples considered in this chapter have demonstrated that the decision analytic approach to valuing the information generated by clinical research can be applied to sequential clinical decision problems. The first and second hurdles can be constructed and they can be used to achieve technically efficient research design and allocative efficiency across proposed clinical research.

In a sequential clinical decision problem the clinician faces an uncertain choice at each stage (decision node) and with no sample information these decisions must be made based only on prior information. There is a chance that these decisions will be wrong and opportunity losses will be incurred at each stage. The EVPI, or the expected opportunity loss, for the full decision problem will be the sum of the EVPI at each stage (at each point where the clinician faces an uncertain decision). The EVPI can be used as the first hurdle for proposed research which will be potentially cost-effective if the EVPI exceeds its fixed cost. Just as in chapter 4 the EVPI is sensitive to the value of health outcome and demonstrates once again that the value of information and research priorities cannot be separated from the budgetary restrictions on service provision. This approach to sequential decision problems allows the EVPI to be established at each stage and can be used to identify those points in a sequence of decisions where the cost of uncertainty is greatest and where additional information may be most valuable. This may provide very different results to a simple sensitivity analysis which is potentially misleading if it is used to identify points where additional information will be valuable.

If proposed research passes the first hurdle then it is potentially cost-effective. The next step is to establish the expected net benefit of sample information at the optimal sample size. In a sequential decision problem trial entrants are allocated to each arm of the trial at each stage. The problem is that the benefits of sampling at an initial decision cannot be separated from the sample which will be generated

at later stages. Therefore the ENBS is not simply the sum of the ENBS calculated separately at each stage. A sample at the initial stage implies acquiring sample information about contingent decisions which will change the prior incremental net benefit and reduce the uncertainty surrounding the initial decision. This problem is solved by making contingent sampling decisions and calculating the posterior net benefits and posterior variance at later stages which is used as the prior information at earlier decisions. This approach is consistent with the principles of backward induction where contingent sampling decisions must be made and posterior values calculated before the ENBS can be established at the initial stage. These methods ensure that the ENBS for the full decision problem can be established by taking the sum of the ENBS at each stage without the danger that the benefit of sampling will be overestimated by double counting at each stage, or underestimated by simplifying a sequential problem to a single-stage decision.

The ENBS at the optimal sample size is the second hurdled that proposed research must overcome. If the ENBS exceeds the fixed cost of the research then it can be regarded as cost-effective and should be implemented. The optimal sample size or the technical efficient scale of research is where the ENBS reaches a maximum and this is the value of the second hurdle. These numerical examples demonstrate that the value of the second hurdle is dependent on the value of health outcome, and reaffirms the conclusion that the value of information and research priorities can not be separated from the budgetary restrictions on service provision.

The comparison of the results of the two-stage problem in section 5.2 and the single-stage problem in chapter 4 demonstrated that an analysis which did not recognise the sequential nature of this problem by excluding the diagnostic test would seriously bias efficient research design and underestimate the value of the research. Similarly, an analysis which simplified the four-stage decision problem in section 5.3 to the two-stage problem in section 5.2 by excluding t_2 would also lead to serious errors at the second hurdle and a technically inefficient design. These examples demonstrate the dangers of using implicit and inconsistent

decision rules to identify relevant alternatives and exclude others. This is an issue which was also discussed in chapter 3. The examples used in this chapter are the same examples which were used in chapter 2 and 3 to show that valid inferences may not necessarily be made by observing current clinical practice. At the end of chapter 3 this conclusion posed a number of problems:

> "If valid inferences cannot be based on observing current clinical practice, but the prospective evaluation of all possible alternatives in a sequential clinical decision problem is not possible, efficient, or ethical, then: (a) how should information of different quality from different sources be combined consistently and explicitly; (b) which clinical decision problems will be worth evaluating in a clinical trial; (b) if a clinical decision problem is worth evaluating which of the competing alternatives should be compared in a clinical trial; and (c) what is the optimal scale of this prospective research?"

The approach taken in this chapter and the last has solved, at least in principle, all but one of these problems. The Bayesian view of probability and the prior distributions assigned to the key variables can consistently and explicitly incorporate prior information from different sources and of different quality (with the quasi sample size representing a more or less sceptical prior). The first and second hurdles can identify which clinical decision problems should be considered for prospective clinical research, and the construction of the second hurdle identifies the efficient scale of this research.

However there remains the problem of which of a number of competing alternatives should be compared within a clinical trial. This has not been addressed, because the analysis in this chapter and in chapter 4 assumed an fixed and equal allocation of trial entrants at each stage. This means that a sample is allocated to each arm of the trial irrespective of the cost and benefit. This arbitrary rule forces part of the sample to be allocated to each alternative and does

not provide a method to identify which of the alternative are irrelevant. In the next chapter this arbitrary fixed and equal allocation rule is relaxed and a simple dynamic programming approach is used to establish optimal patient allocation and provide an explicit and consistent method to identify relevant alternatives which should be compared in the trial.

Appendix D Tables and Figures for Chapter 5

Contents:

Figure 5.2.1 Decision Tree for the Two Stage Decision Problem

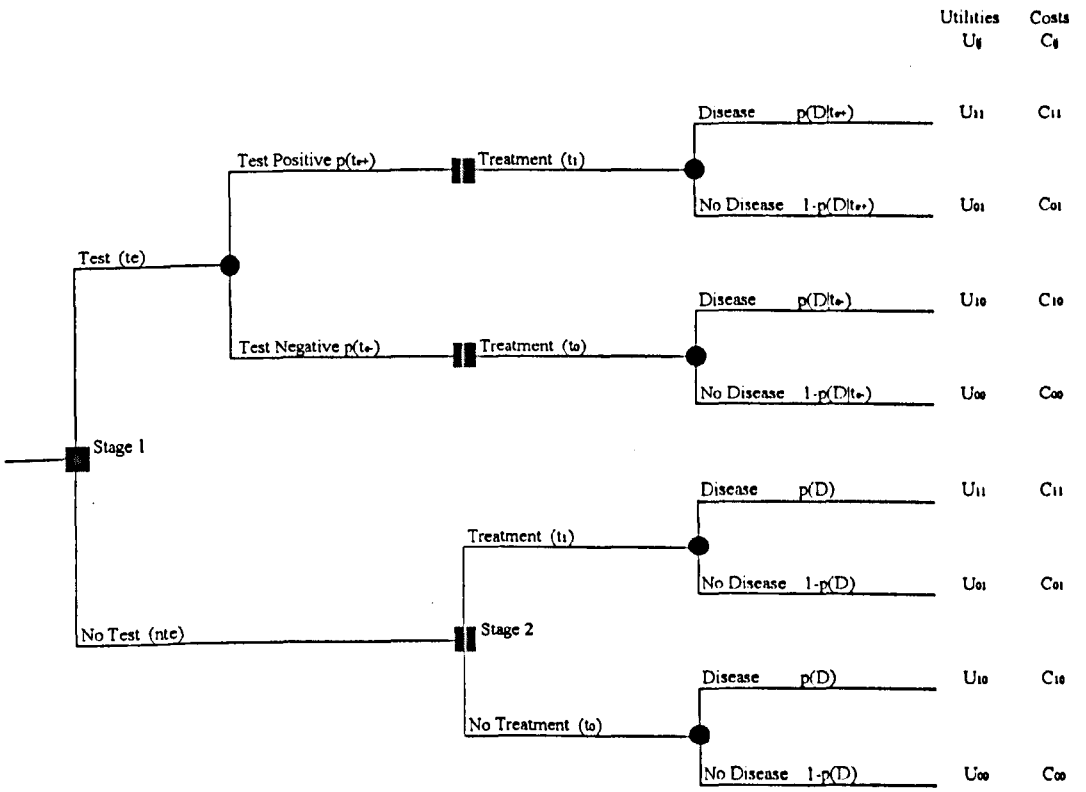


Table 5.2

Numerical Example for the Two-Stage Decision Problem

	Prior Mean	Prior SD	Population SD	Quasi prior sample size (n_0)
U_{11}	6	0.5164	1.2649	12
U_{01}	8	0.5164	1.2649	12
U_{10}	2	0.2582	0.8942	6
U_{00}	10	0.2582	0.8942	6
C_{11}	£12,000			
C_{01}	£8,000			
C_{10}	0			
C_{00}	0			
C_{te}	£8,000			
p(D)	0.6	0.1	0.4899	24
$p(t_e^+\|D)$	0.9	0.0866	0.3	12
$p(t_e^-\|ND)$	0.8	0.1155	0.4	12

$p(t_e^+|D)$ = probability of a true positive result (sensitivity)

$p(t_e^-|ND)$ = probability of a true negative result (specificity)

In this example 1000 patients enter the decision problem in one year.

Figure 5.2.2 EVPI for the Two-Stage Decision Problem

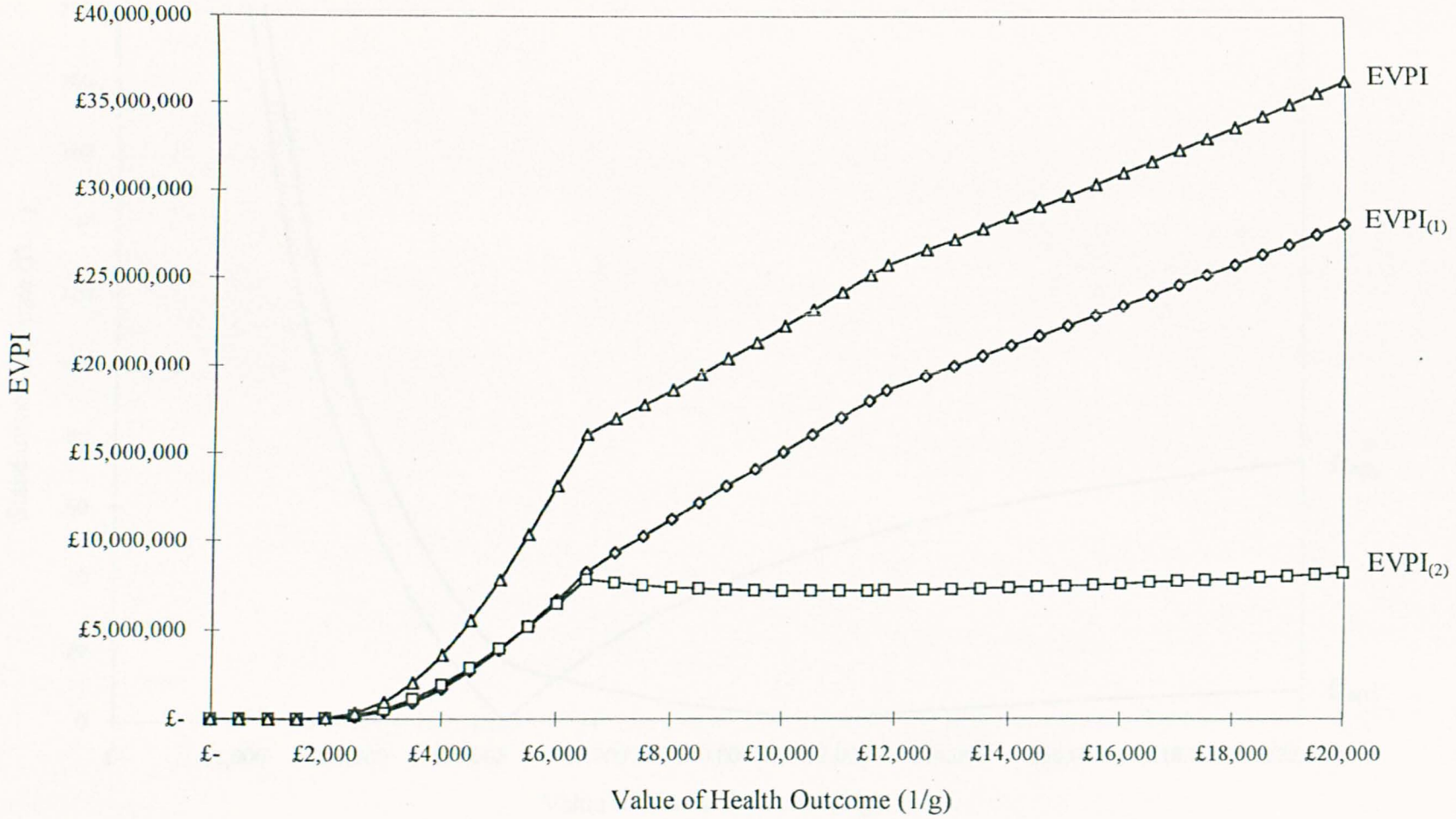


Figure 5.2.3a Standardised Distance ($D_{0(s)}$) at Stage 1 and Stage 2

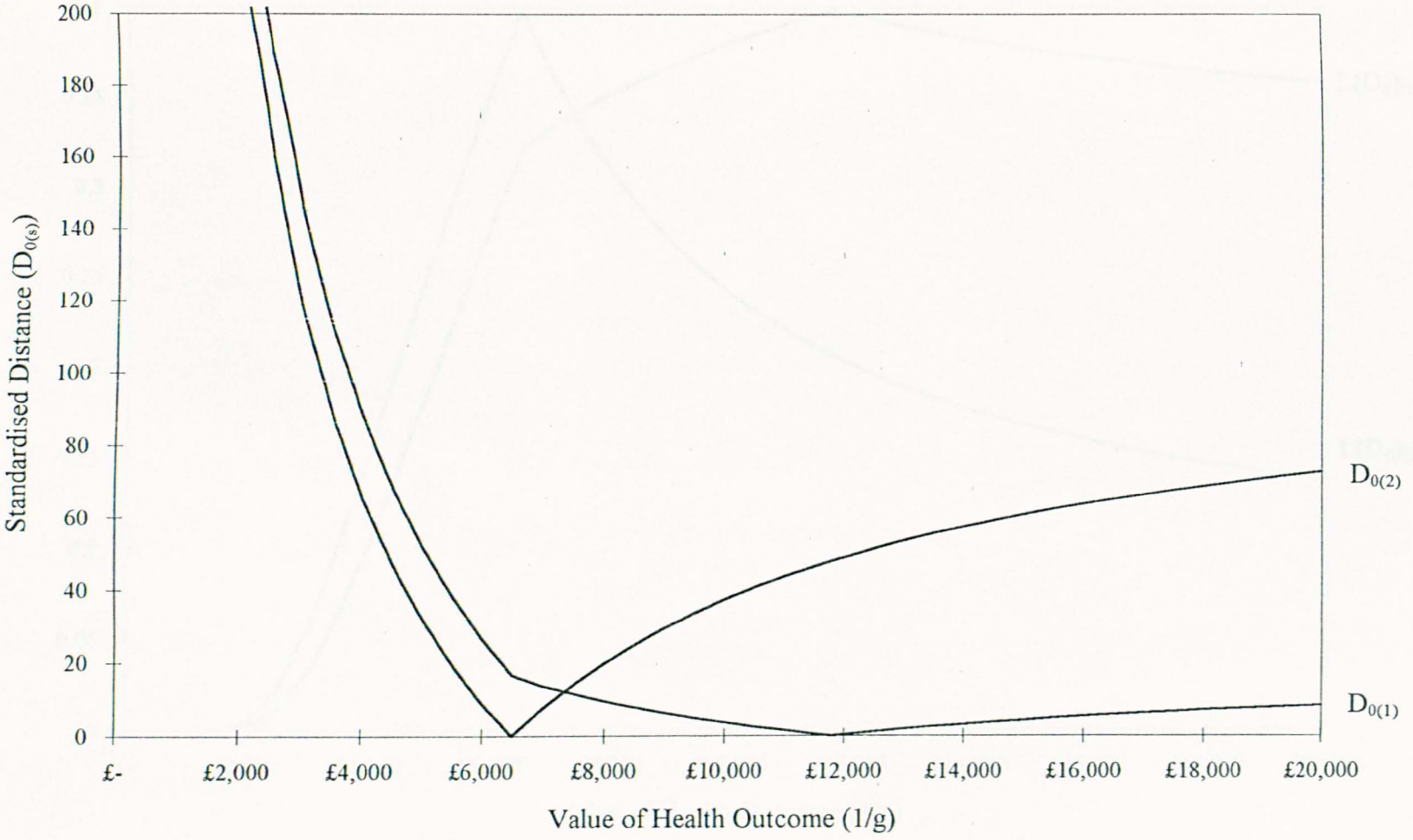


Figure 5.2.3b Loss Integral ($L(D_0)$) at Stage 1 and Stage 2

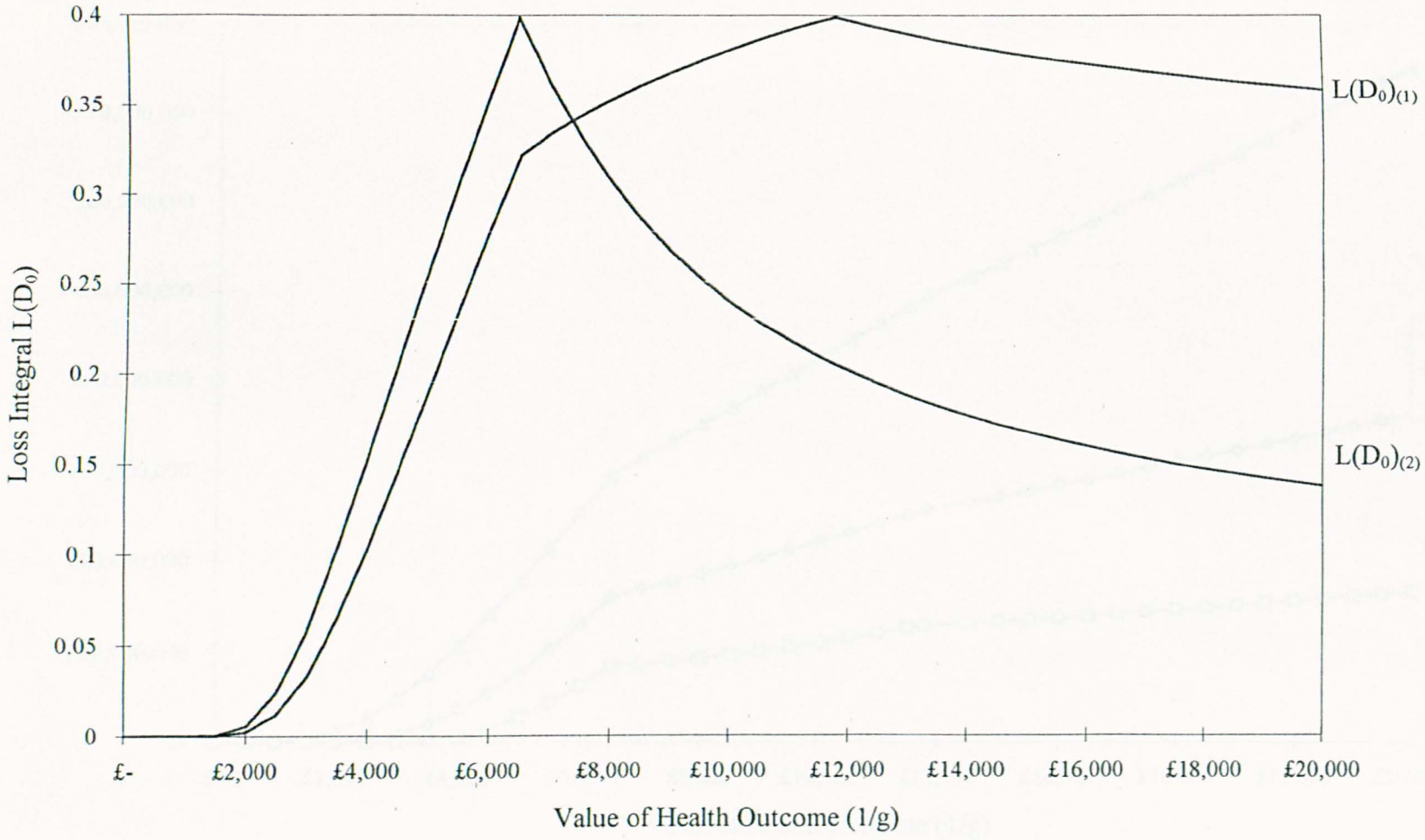


Figure 5.2.4 EVPI and the Strength of Prior Information

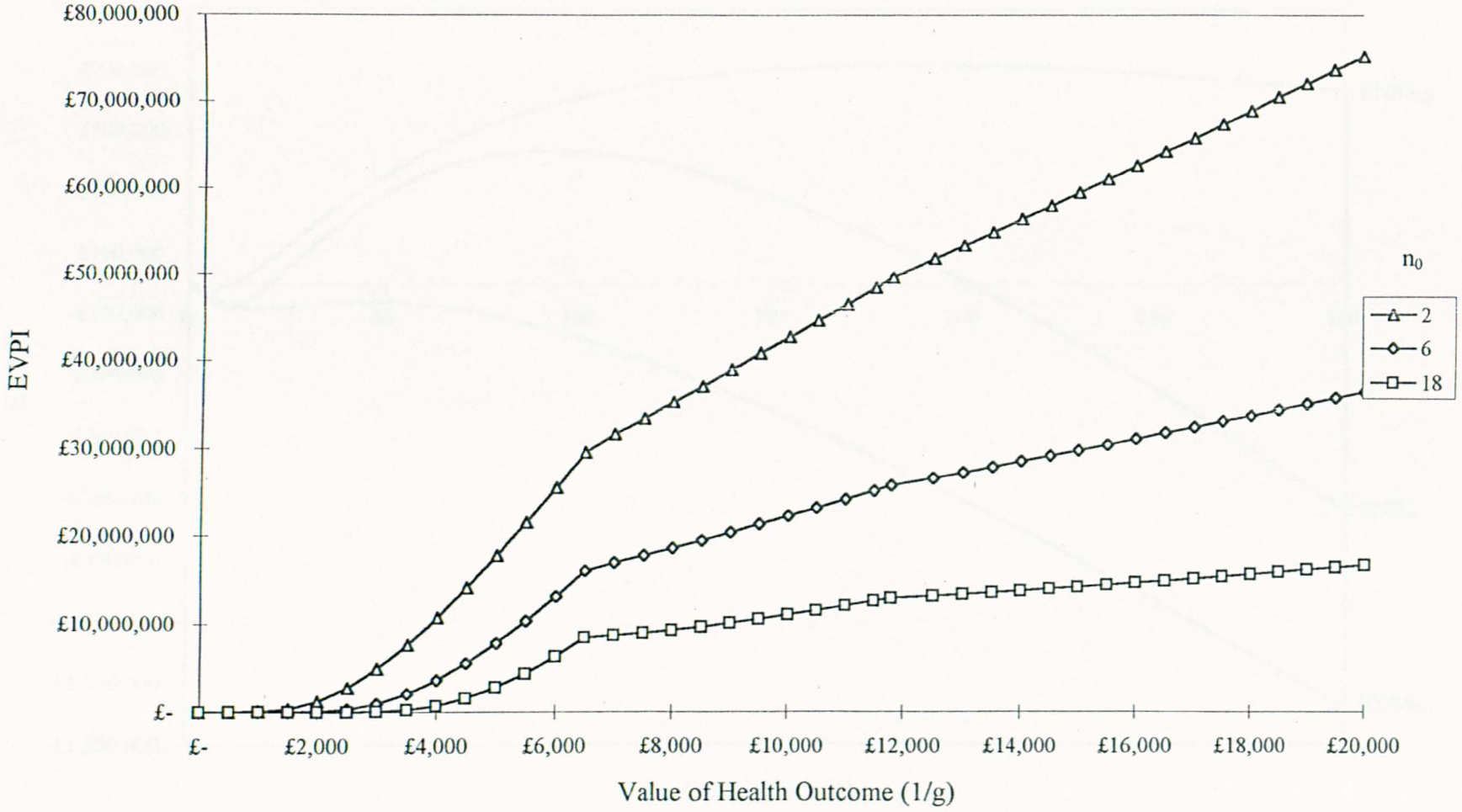


Figure 5.2.5a ENBS for the Two Stage Decision Problem (1/g=£4,000)

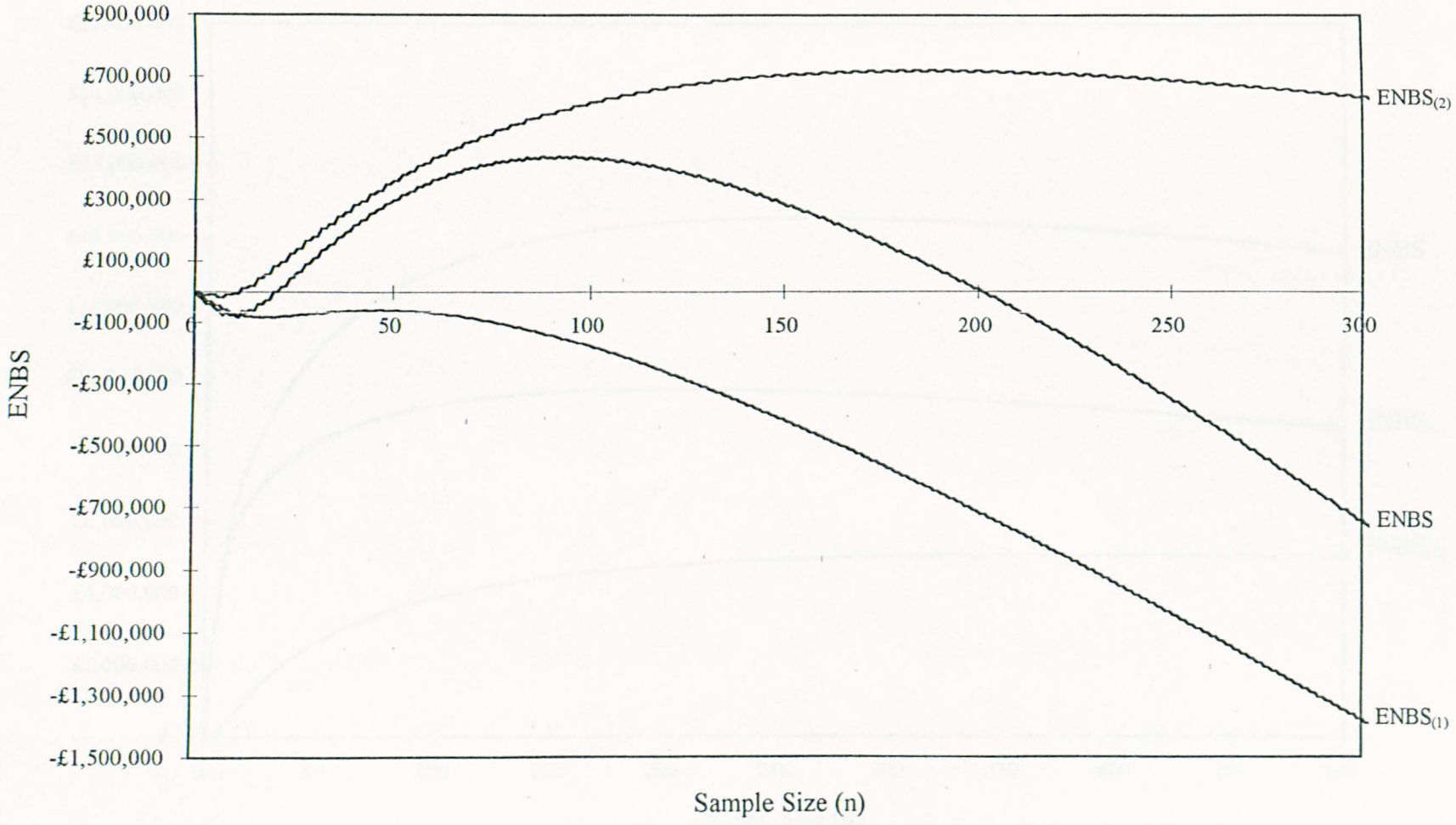


Figure 5.2.5b ENBS for the Two Stage Decision Problem (1/g=£10,000)

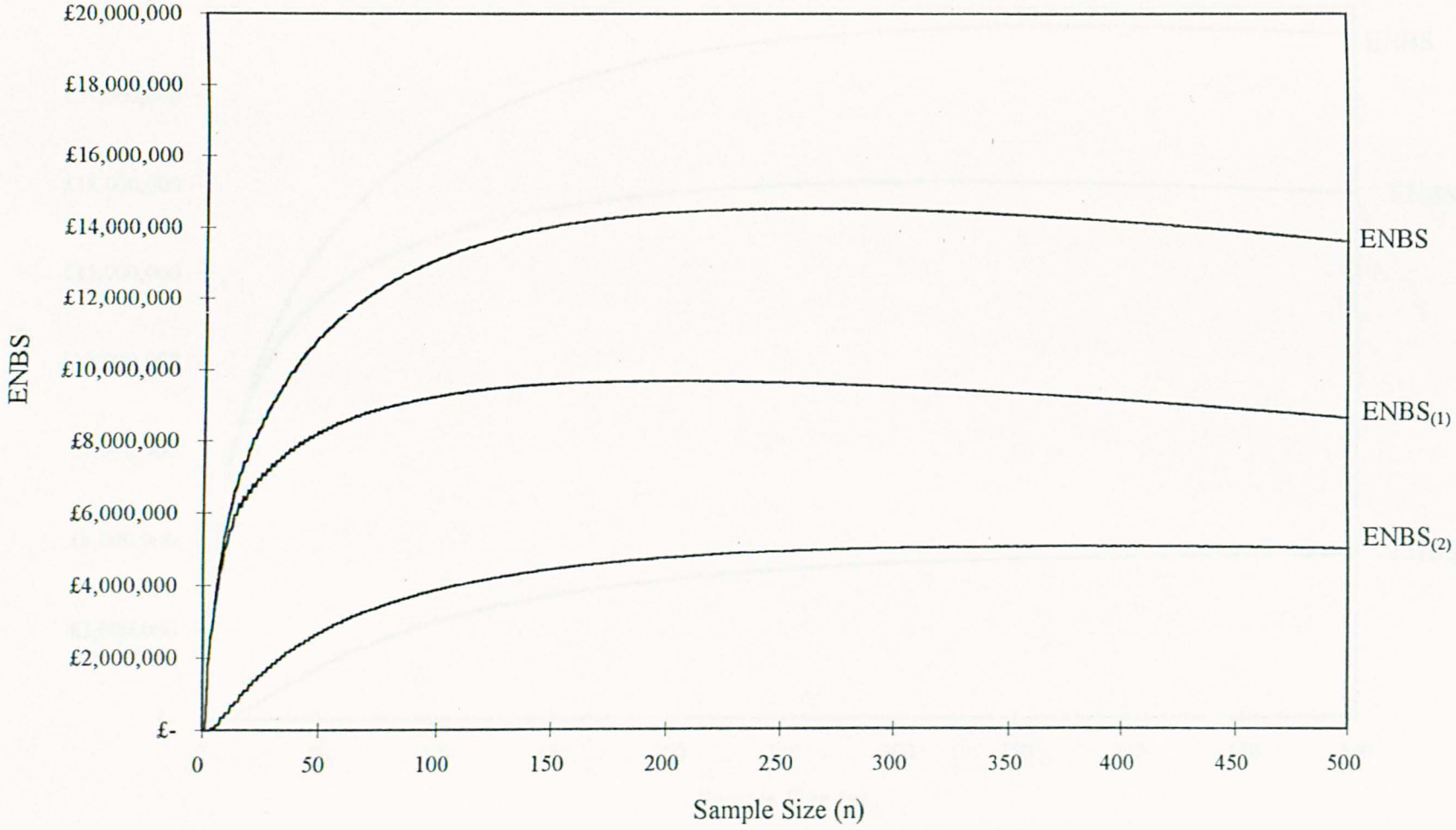


Figure 5.2.5c ENBS for the Two Stage Decision Problem (1/g=£20,000)

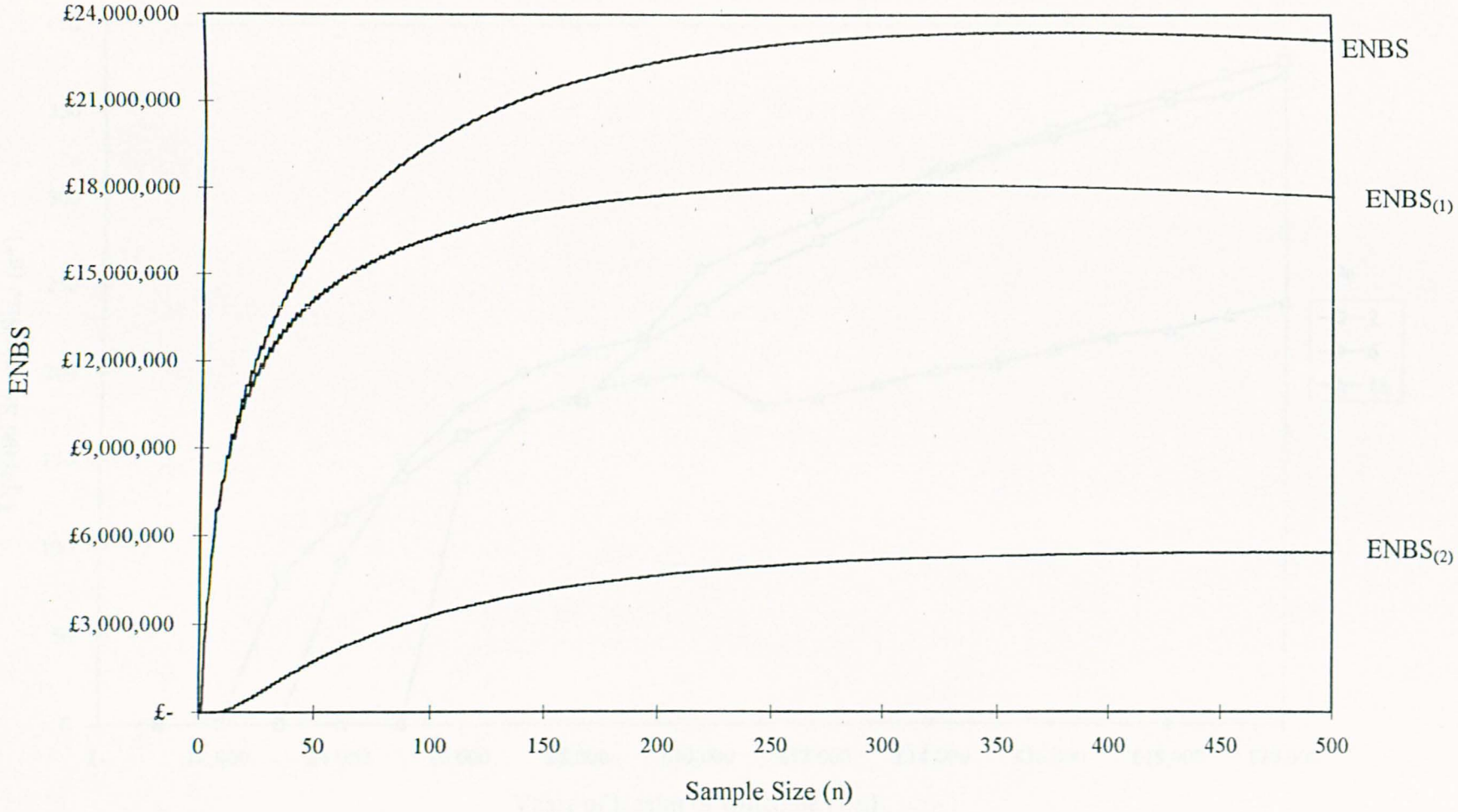


Figure 5.2.6 Optimal Sample Size for the Two Stage Decision Problem (n*)

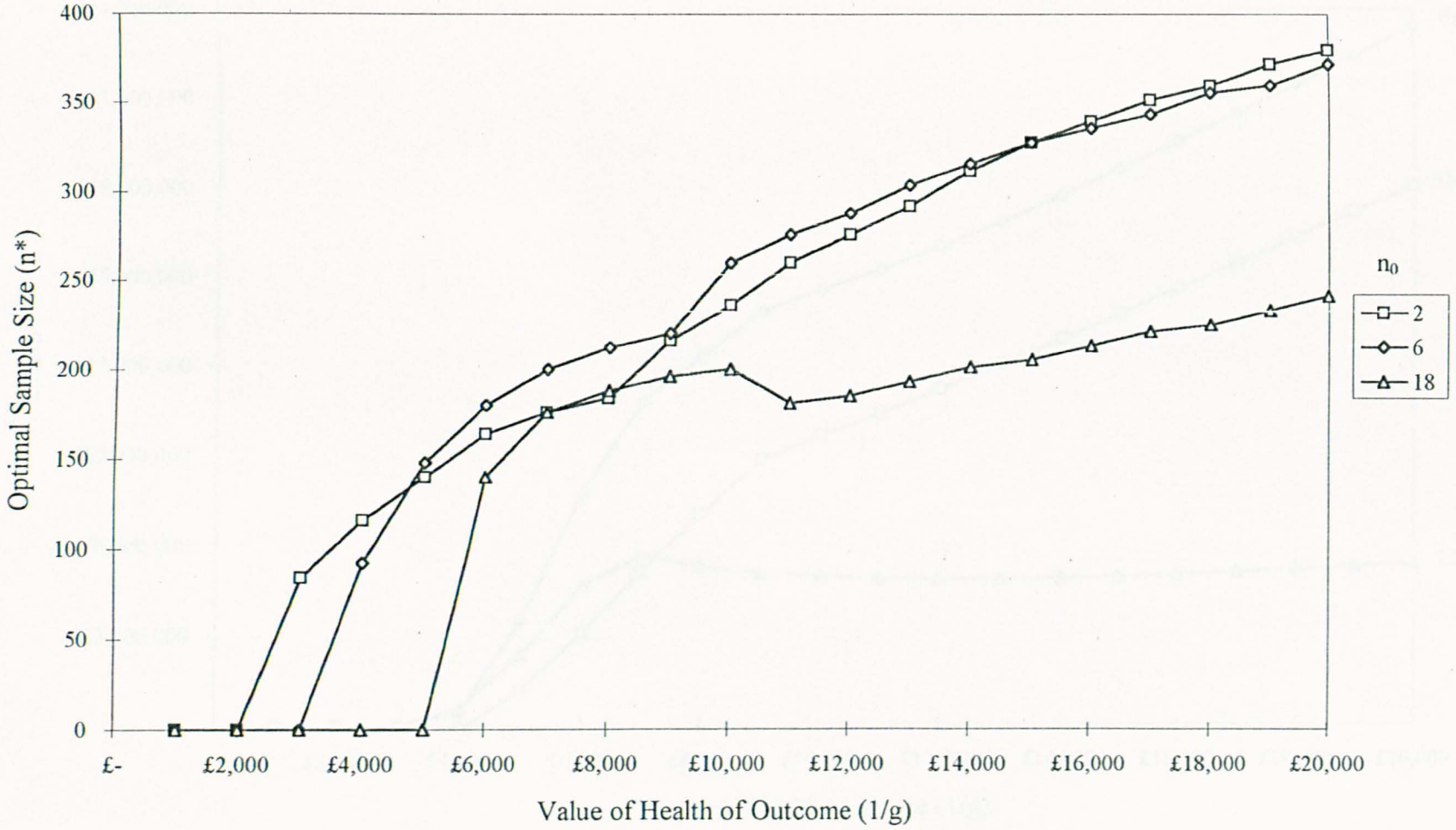


Figure 5.2.7a ENBS at Optimal Sample Size (n*) for the Two Stage Decision Problem

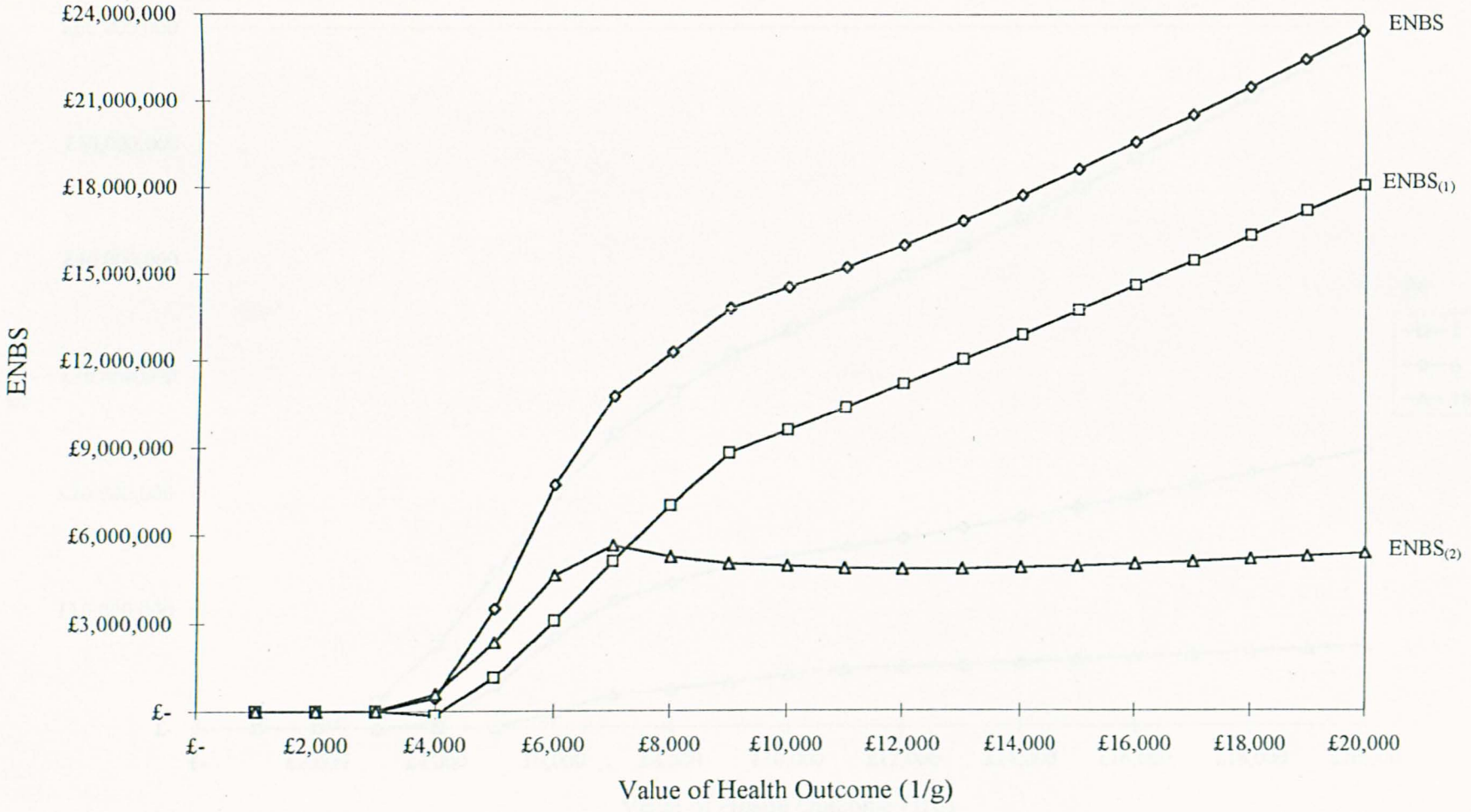


Figure 5.2.7b ENBS and the Strength of Prior Information

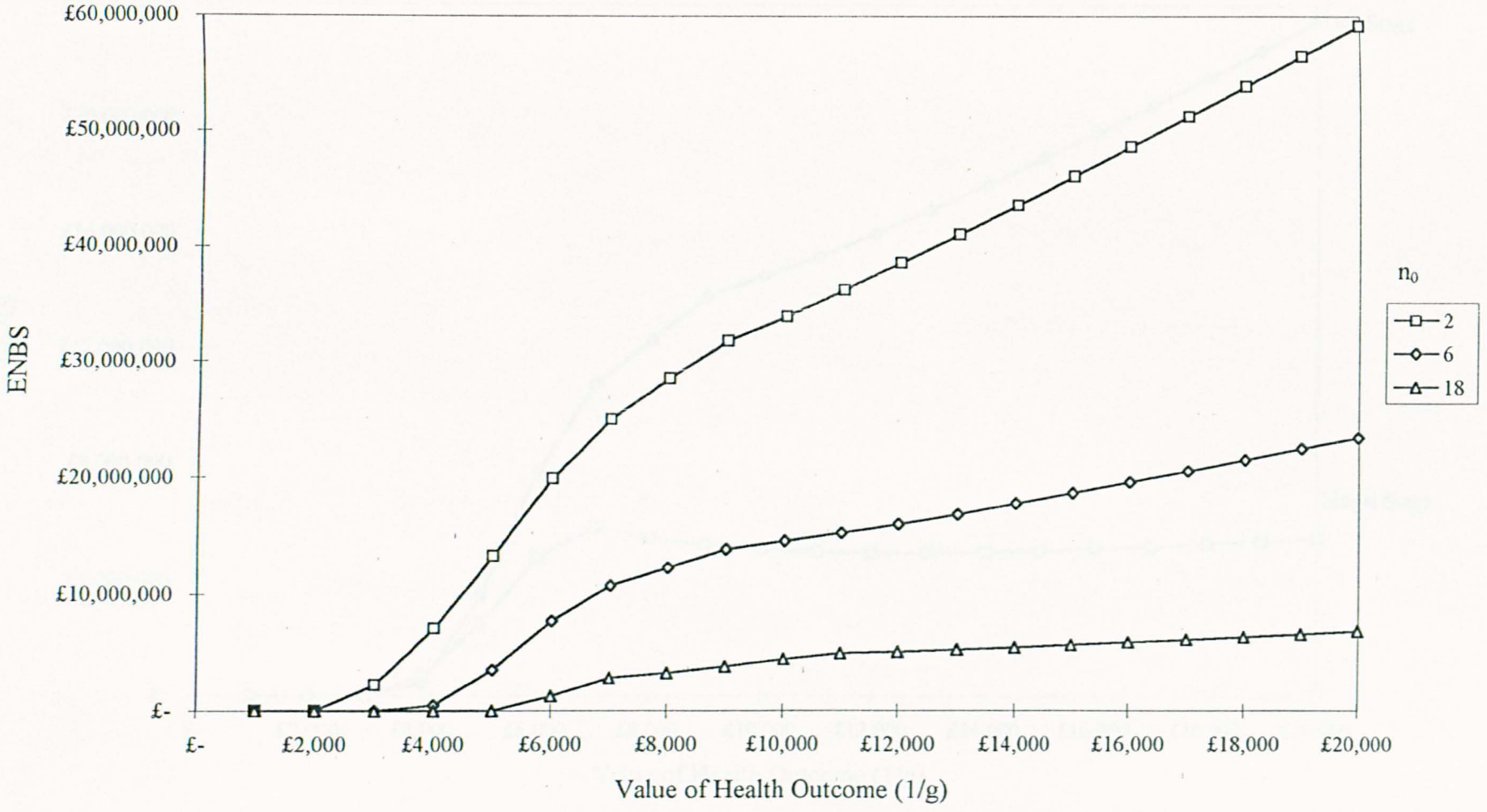


Figure 5.2.8a ENBS for the Two Stage and Single Stage Decision Problem

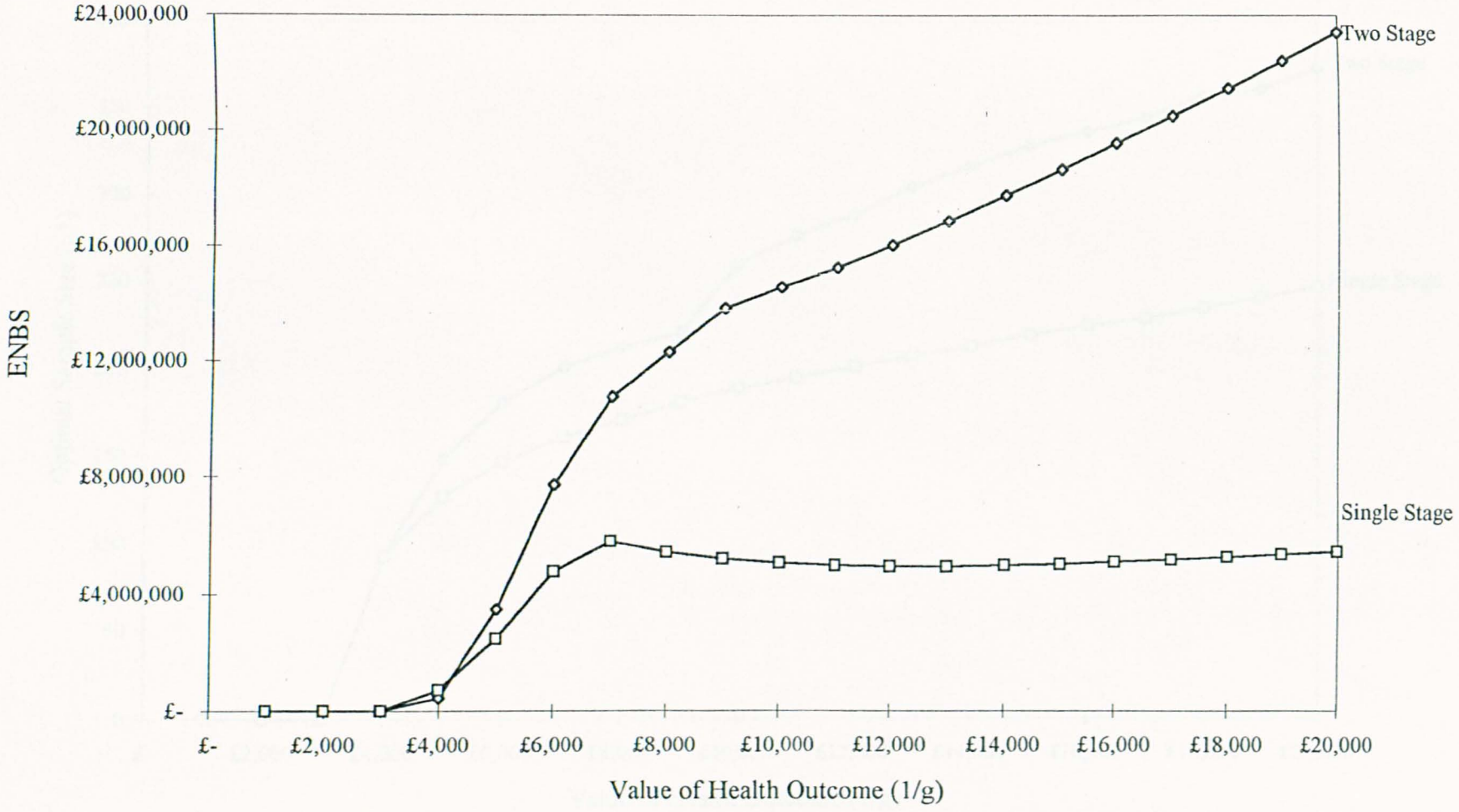


Figure 5.2.8b Optimal Sample Size for the Single Stage and Two Stage Decision Problem

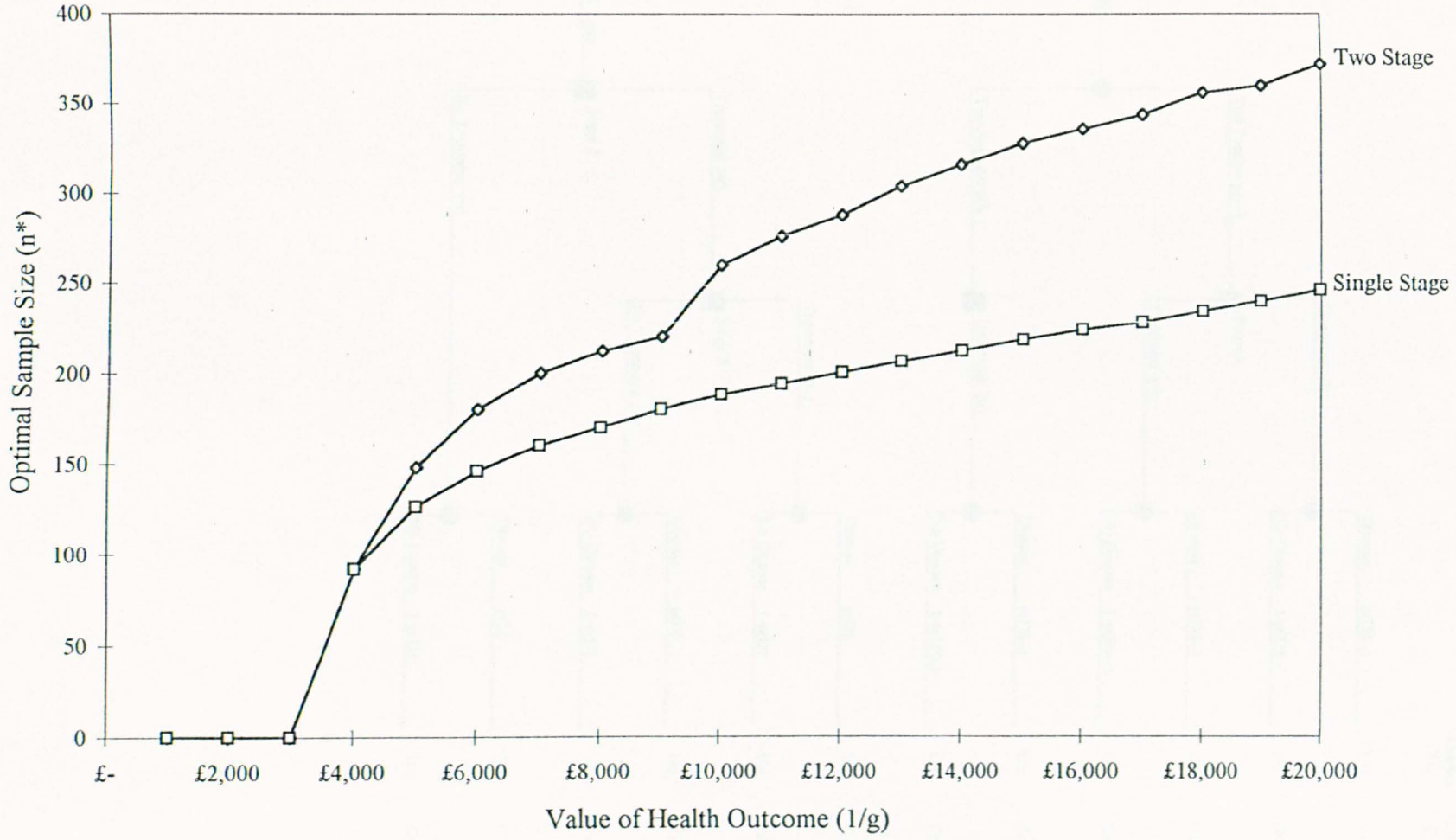


Figure 5.3.1 Decision Tree for the Four Stage Decision Problem

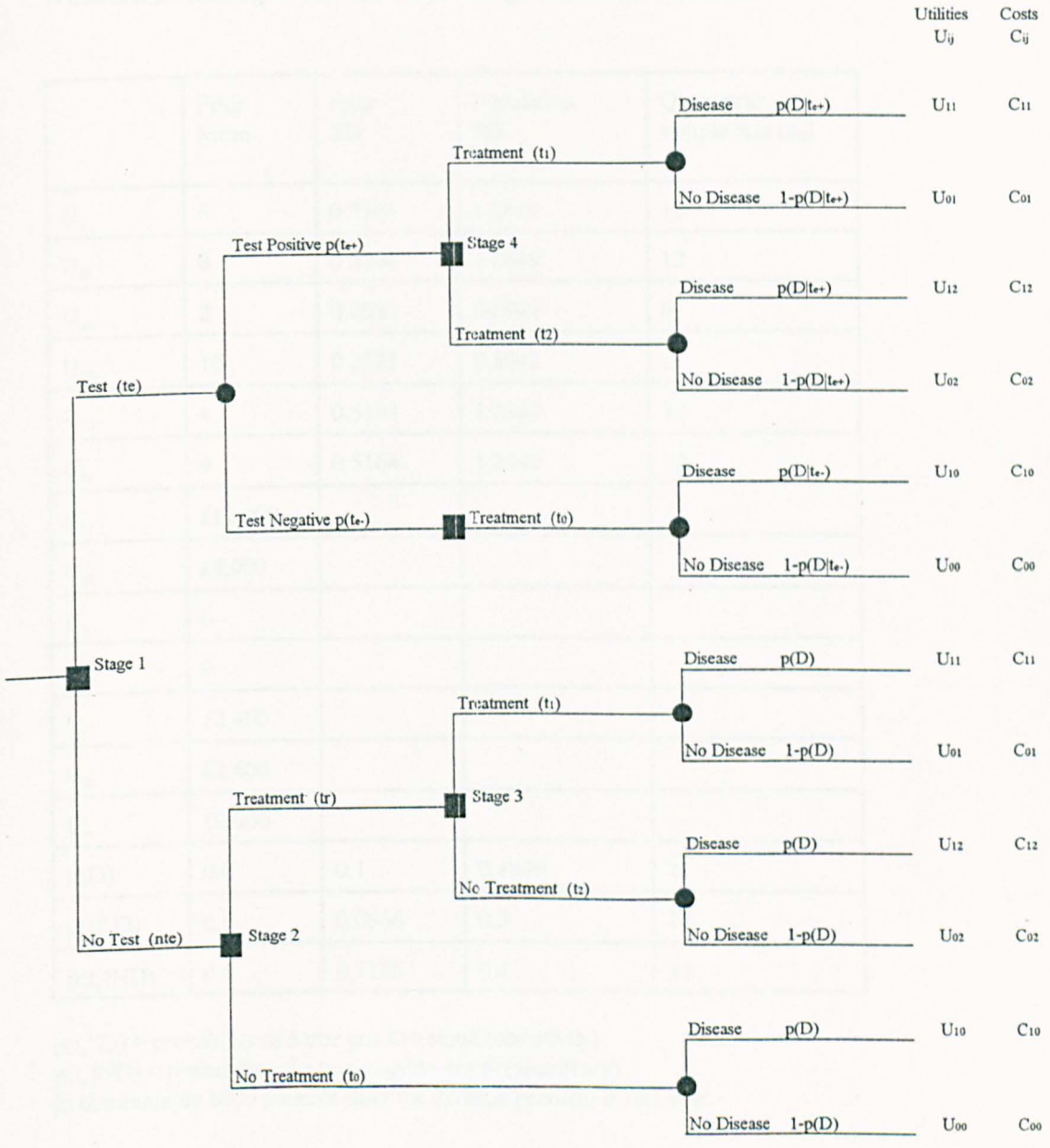


Table 5.3

Numerical Example for the Four-Stage Decision Problem

	Prior Mean	Prior SD	Population SD	Quasi prior sample size (n_0)
U_{11}	6	0.5164	1.2649	12
U_{01}	8	0.5164	1.2649	12
U_{10}	2	0.2582	0.8942	6
U_{00}	10	0.2582	0.8942	6
U_{12}	4	0.5164	1.2649	12
U_{02}	4	0.5164	1.2649	12
C_{11}	£12,000			
C_{01}	£8,000			
C_{10}	0			
C_{00}	0			
C_{12}	£2,400			
C_{02}	£2,400			
C_{te}	£8,000			
p(D)	0.6	0.1	0.4899	24
$p(t_e^+\|D)$	0.9	0.0866	0.3	12
$p(t_e^-\|ND)$	0.8	0.1155	0.4	12

$p(t_e^+|D)$ = probability of a true positive result (sensitivity)
$p(t_e^-|ND)$ = probability of a true negative result (specificity)
In this example 1000 patients enter the decision problem in one year.

Figure 5.3.2a EVPI for the Four Stage Decision Problem

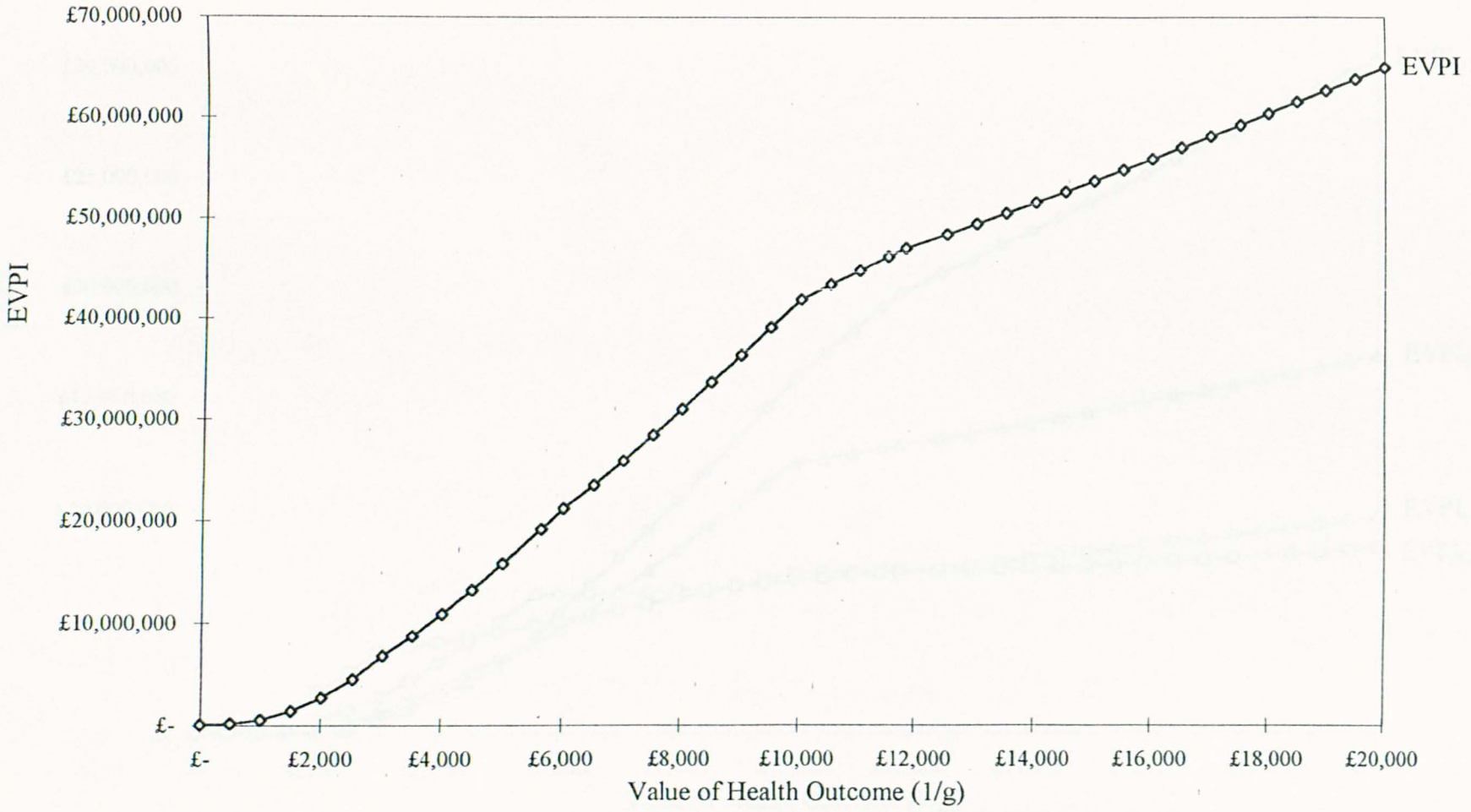


Figure 5.3.2b EVPI at Each Stage

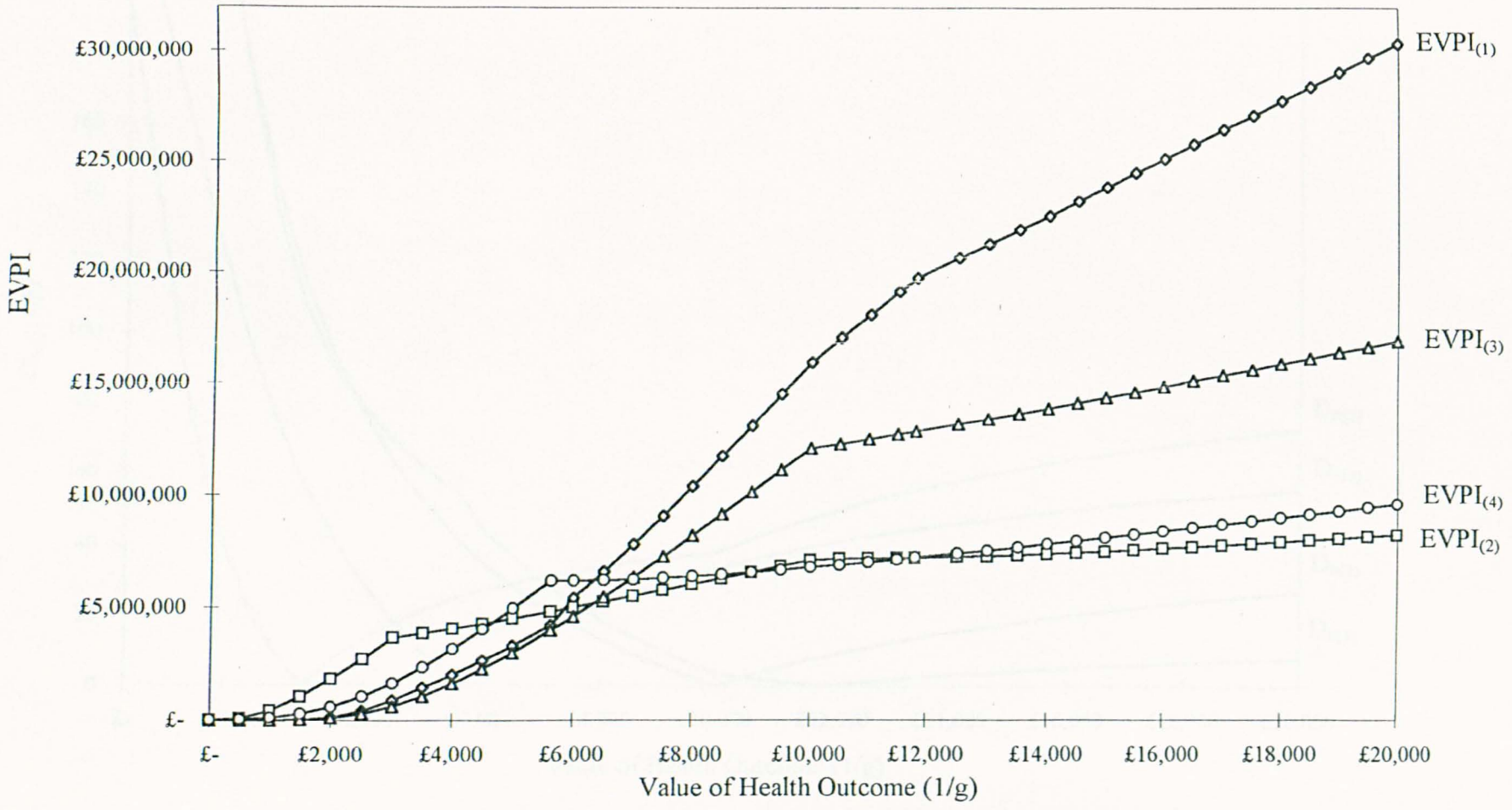


Figure 5.3.3a Standardised Distance at Stage (s) ($D_{0(s)}$)

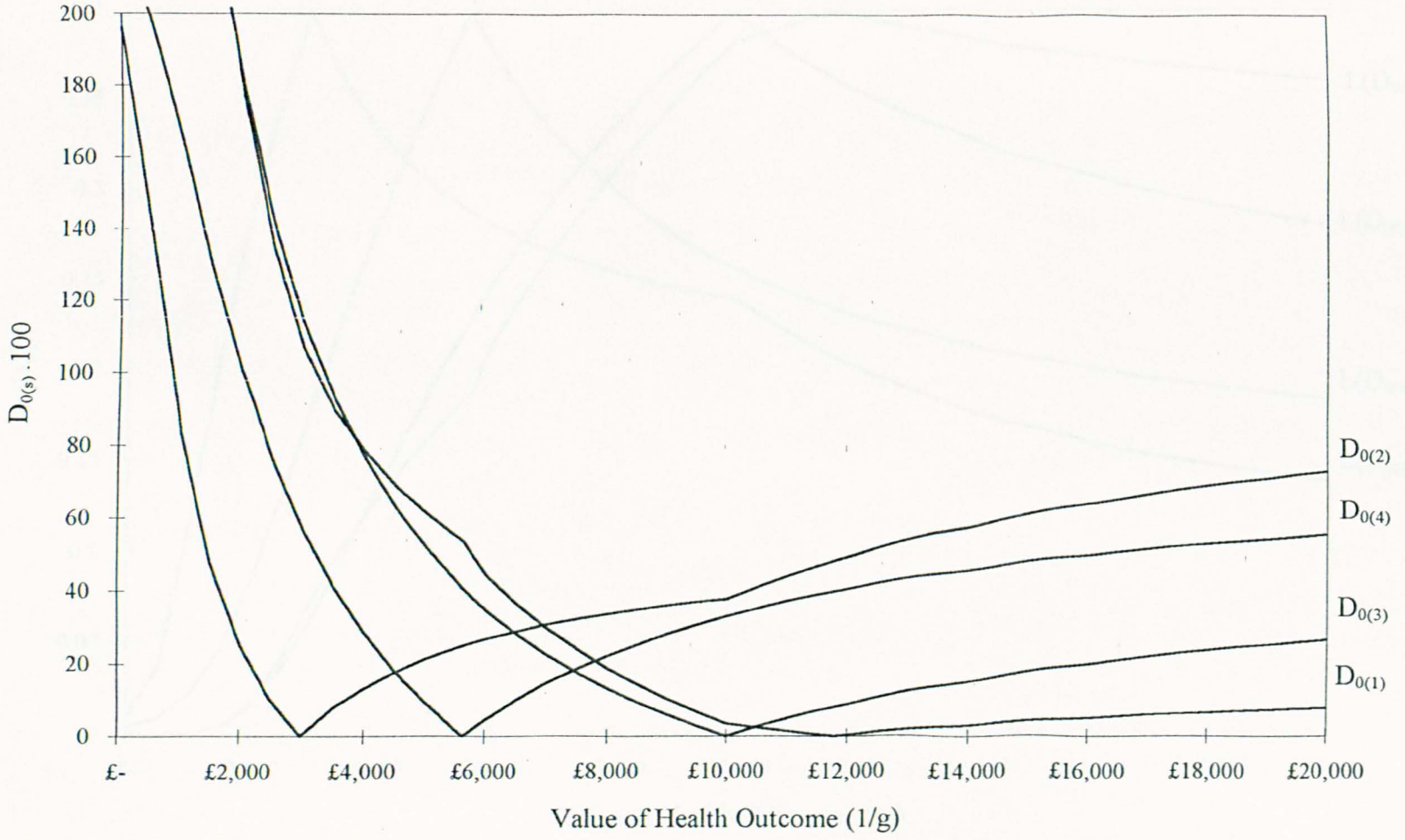


Figure 5.3.3b Loss Integral at each Stage (s) ($L(D_{0(s)})$

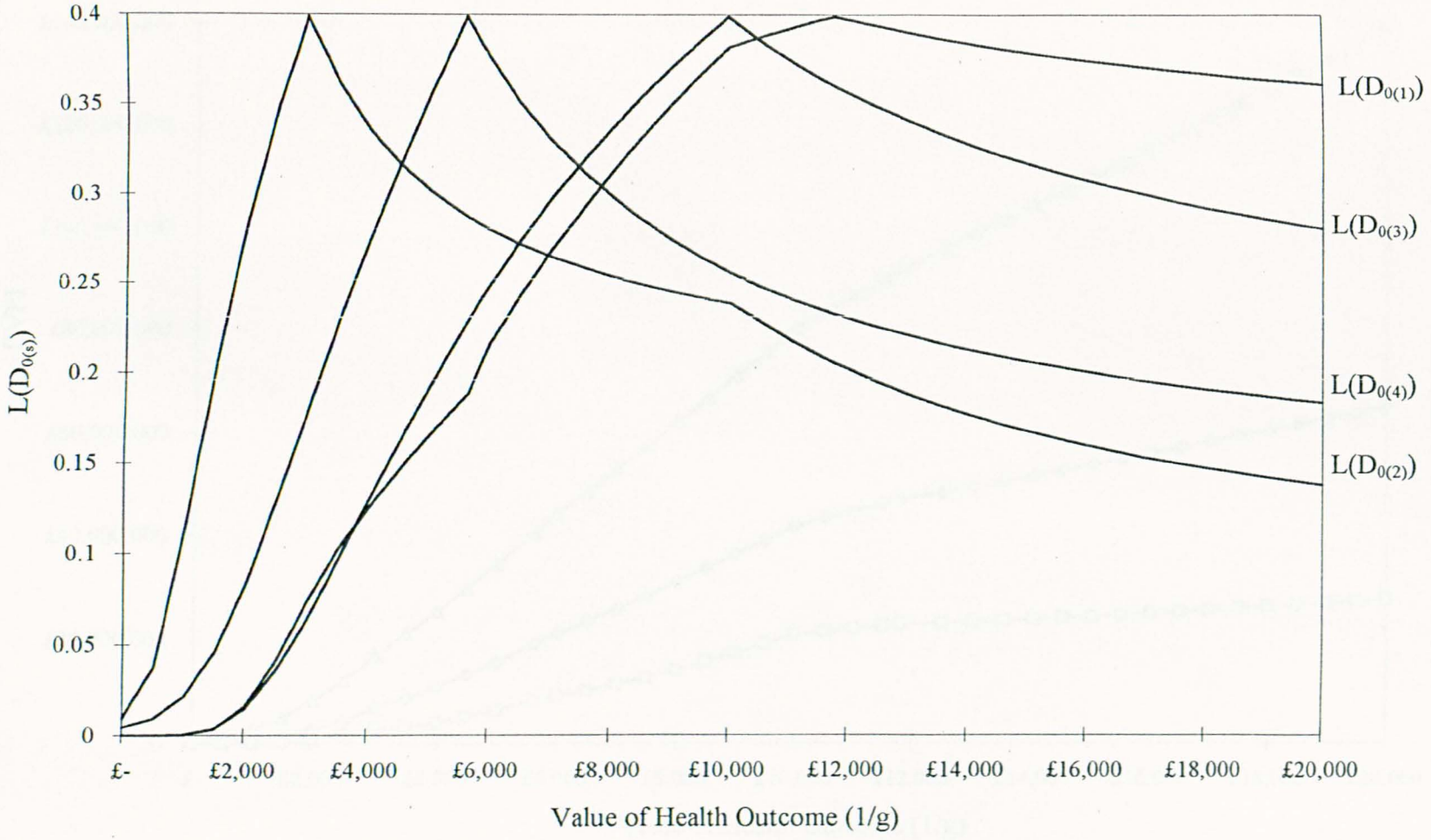


Figure 5.3.4 EVPI and the Strength of Prior Information

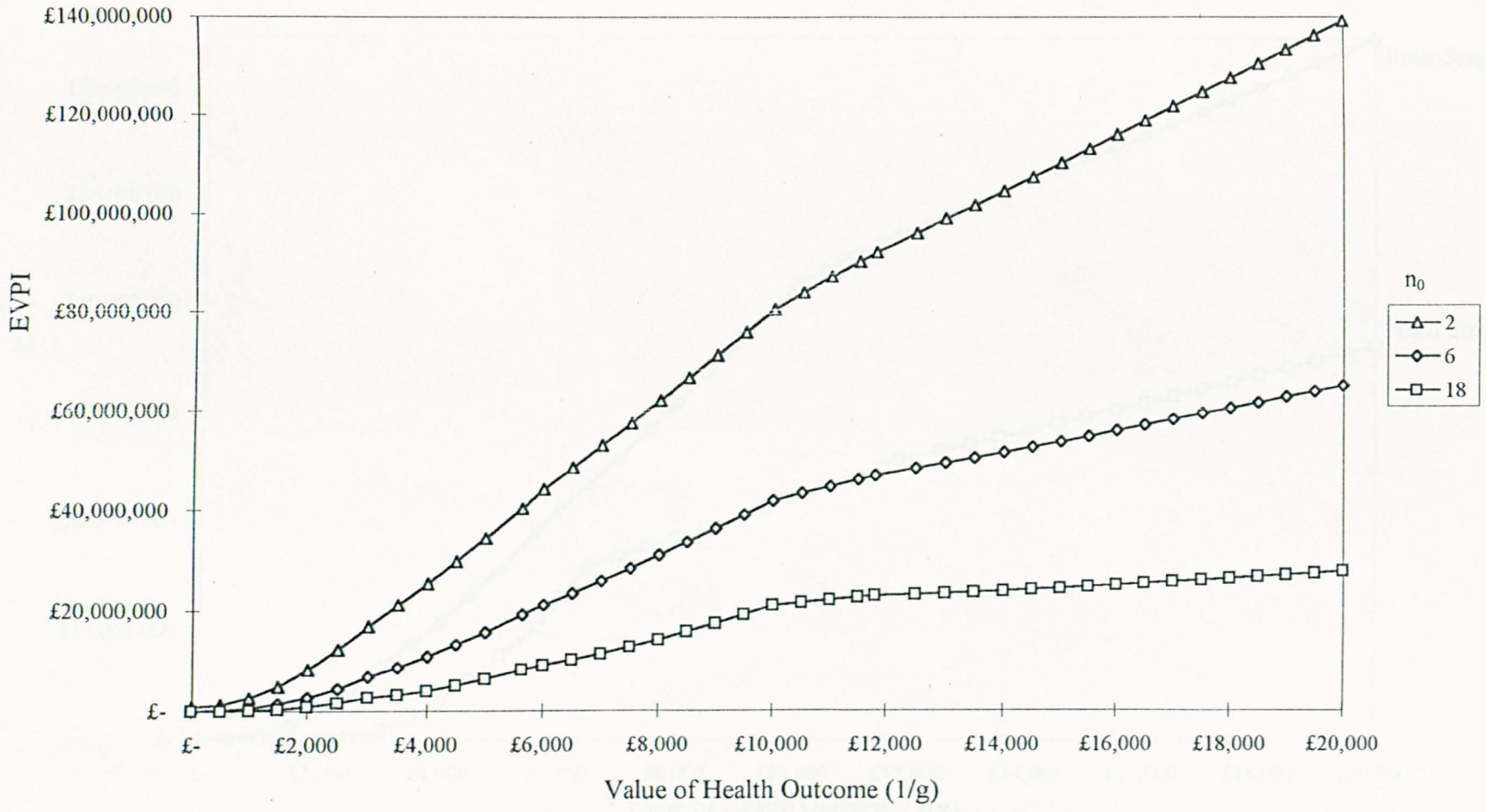


Figure 5.3.5 EVPI for the Four and Two Stage Decision Problems

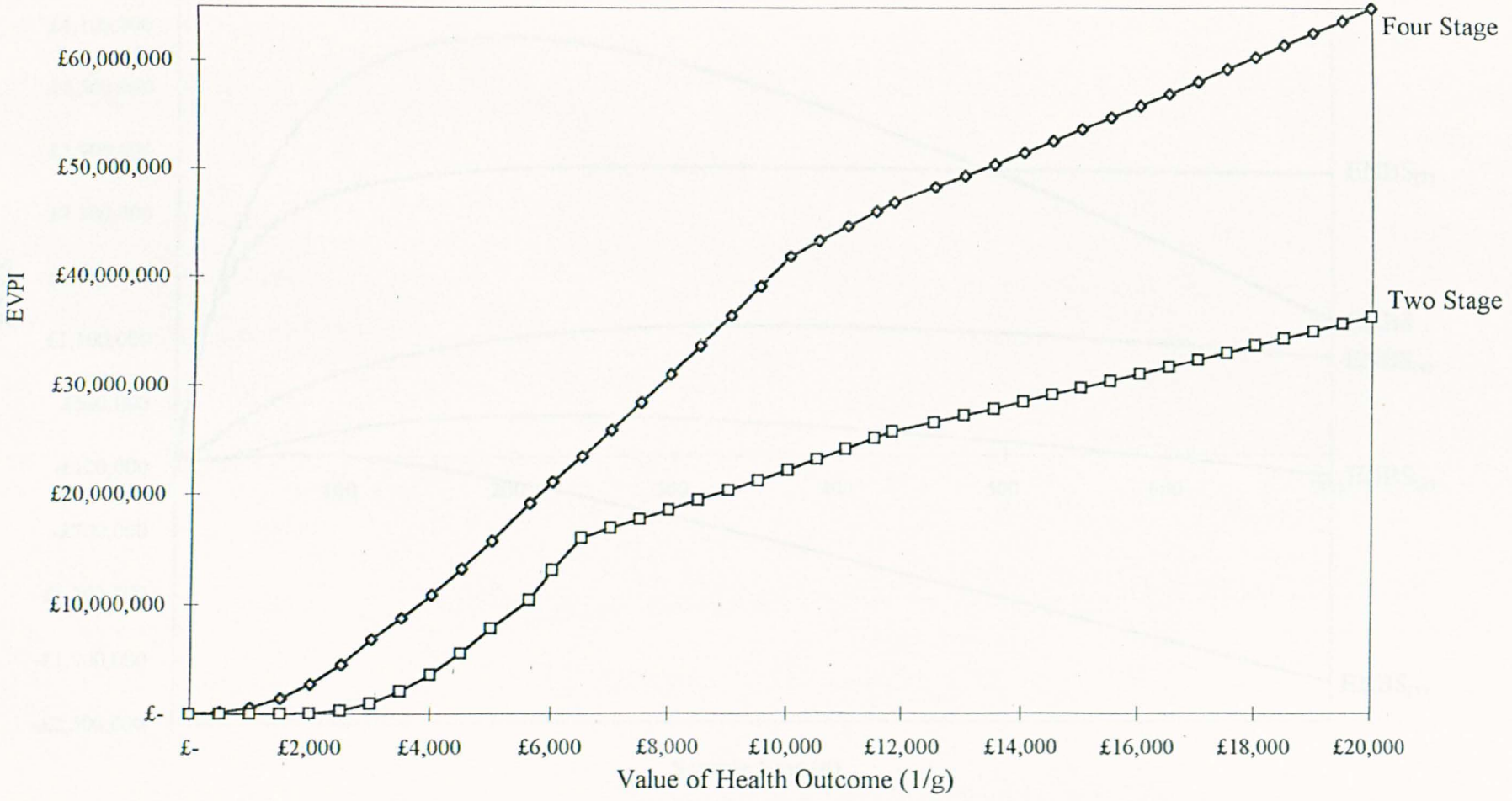


Figure 5.3.6a ENBS for the Four Stage Decision Problem (1/g=£4,000)

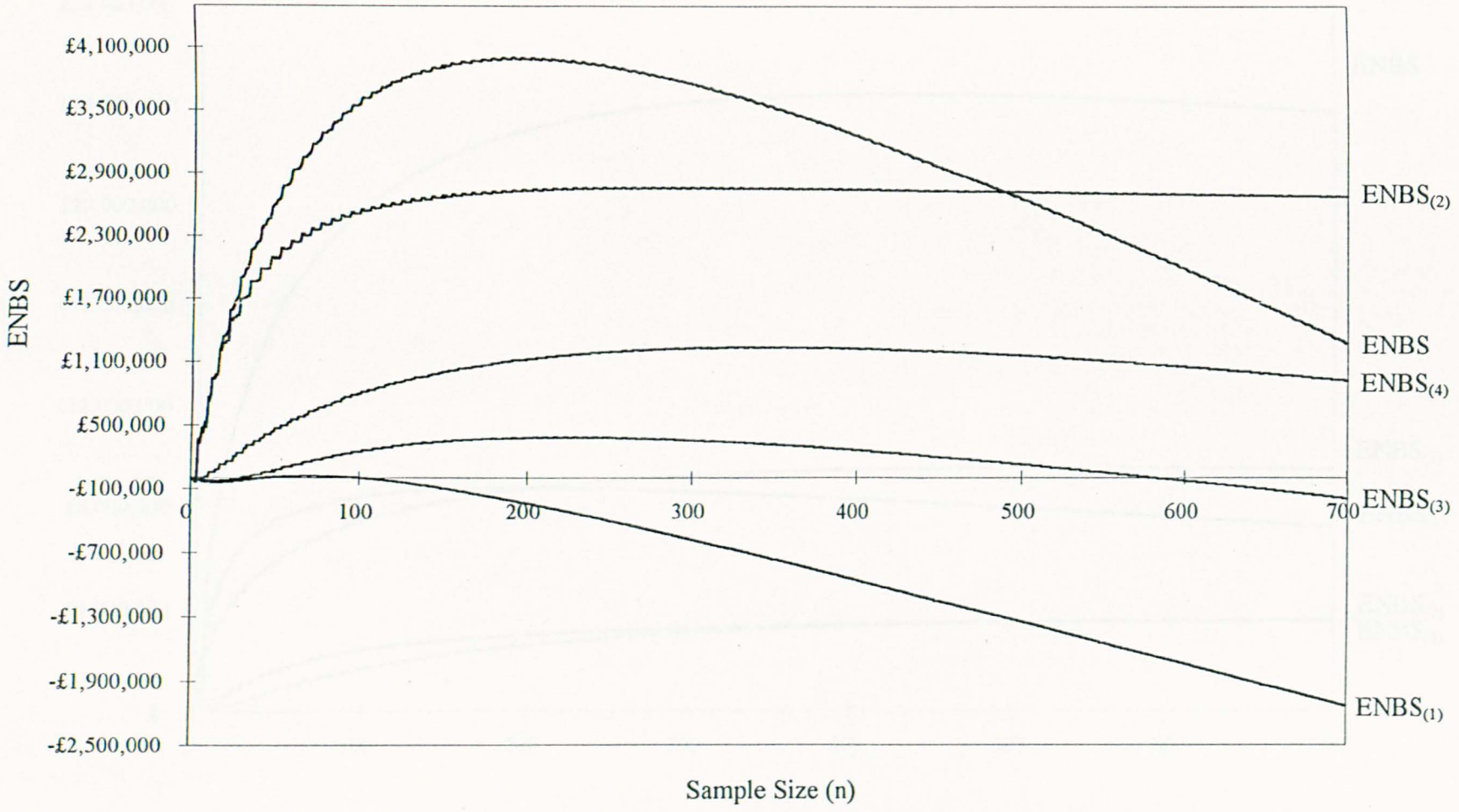


Figure 5.3.6b ENBS for the Four Stage Decision Problem (1/g=£10,000)

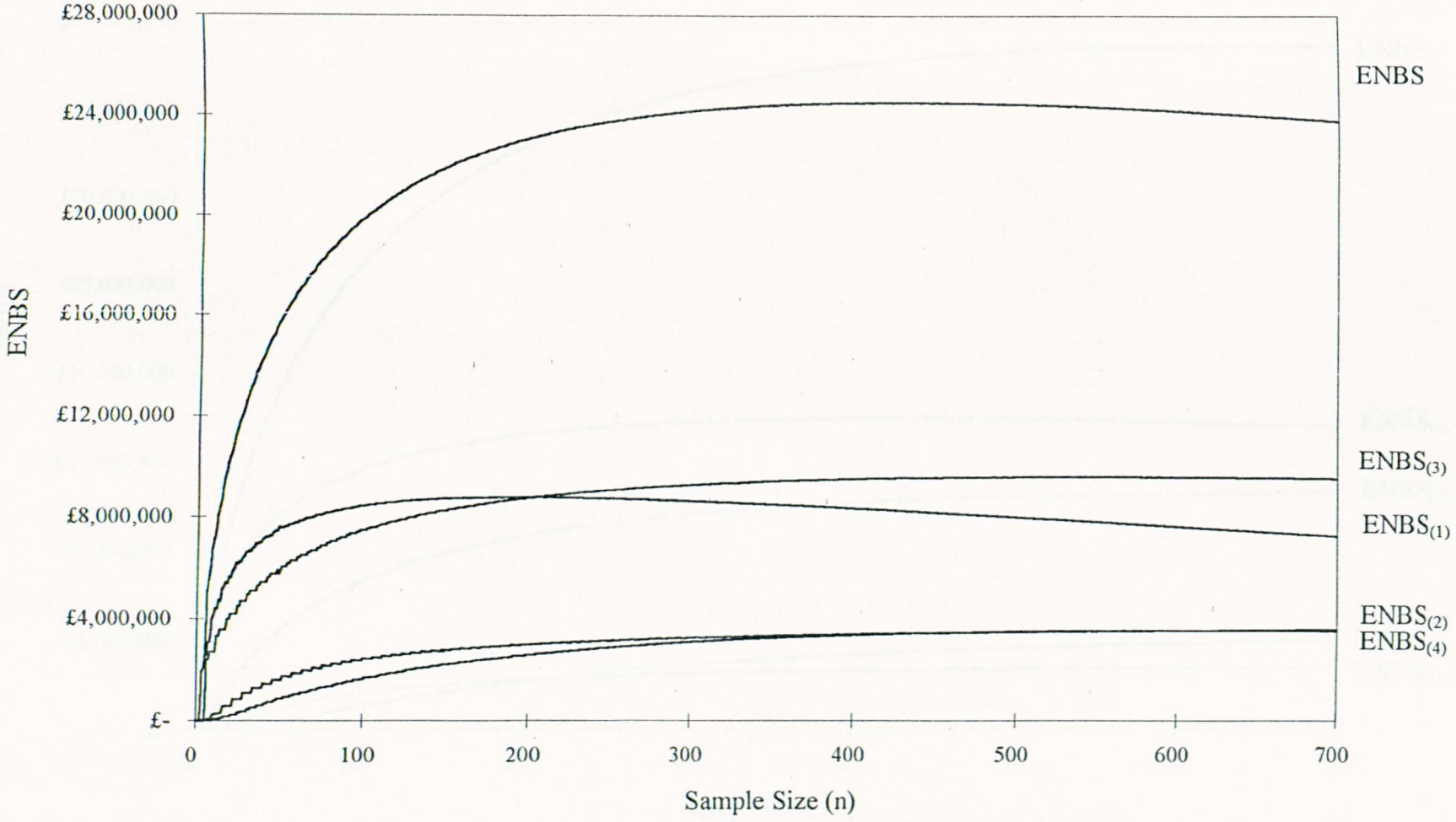


Figure 5.3.6c ENBS for the Four Stage Decision Problem (1/g=£20,000)

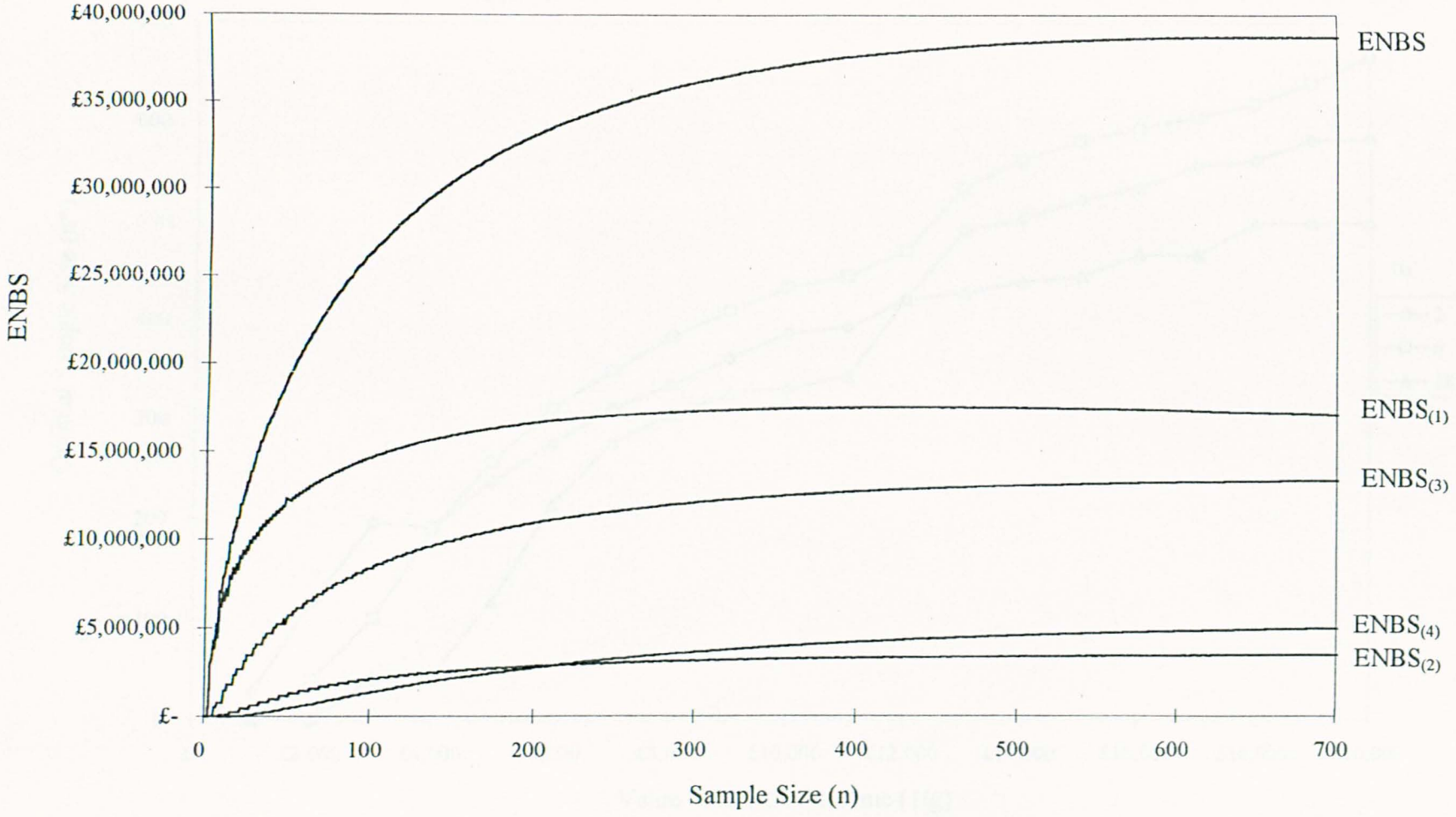


Figure 5.3.7 Optimal Sample Size (n*)

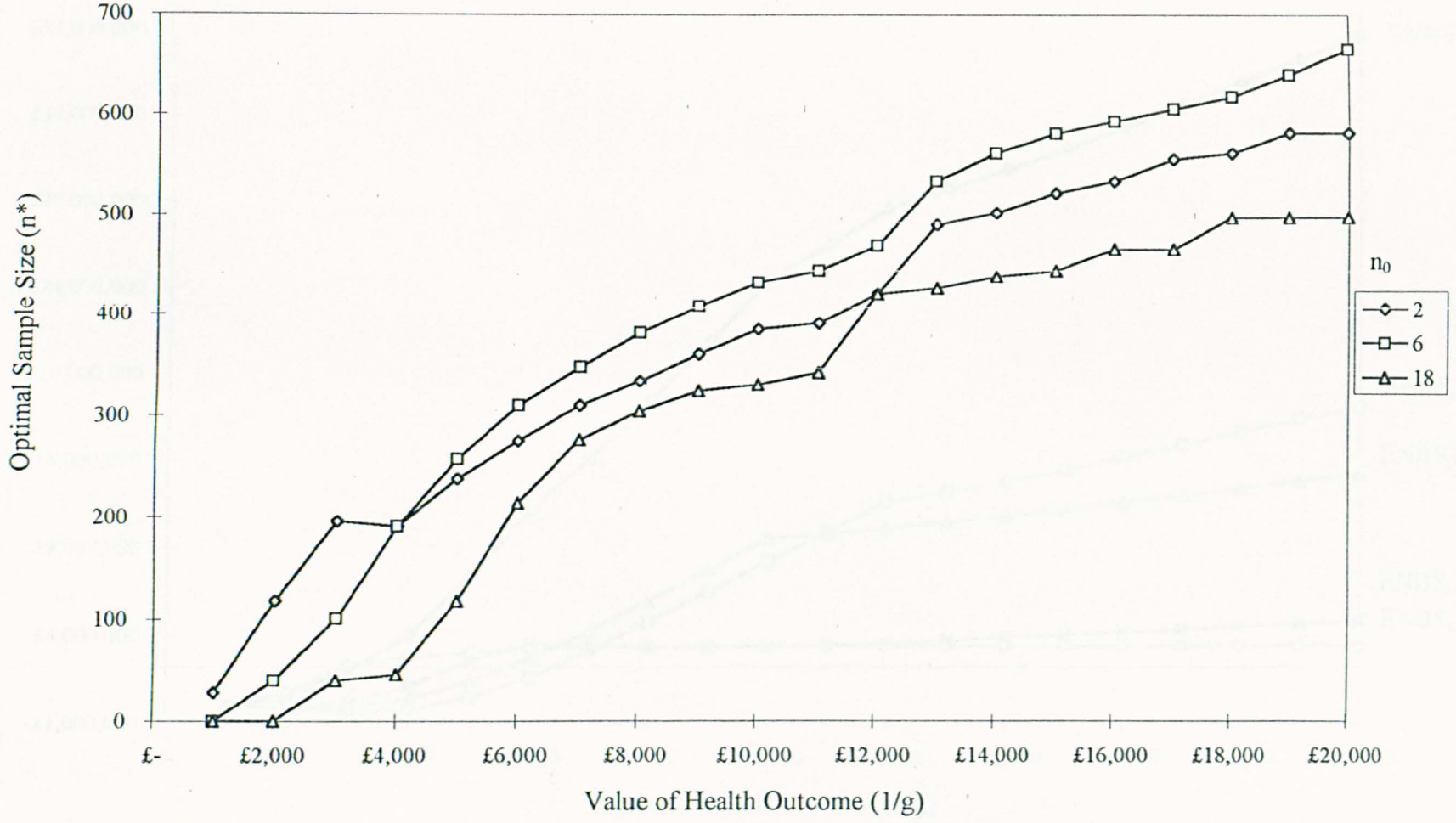


Figure 5.3.8a ENBS at Optimal Sample Size for the Four Stage Decision Problem

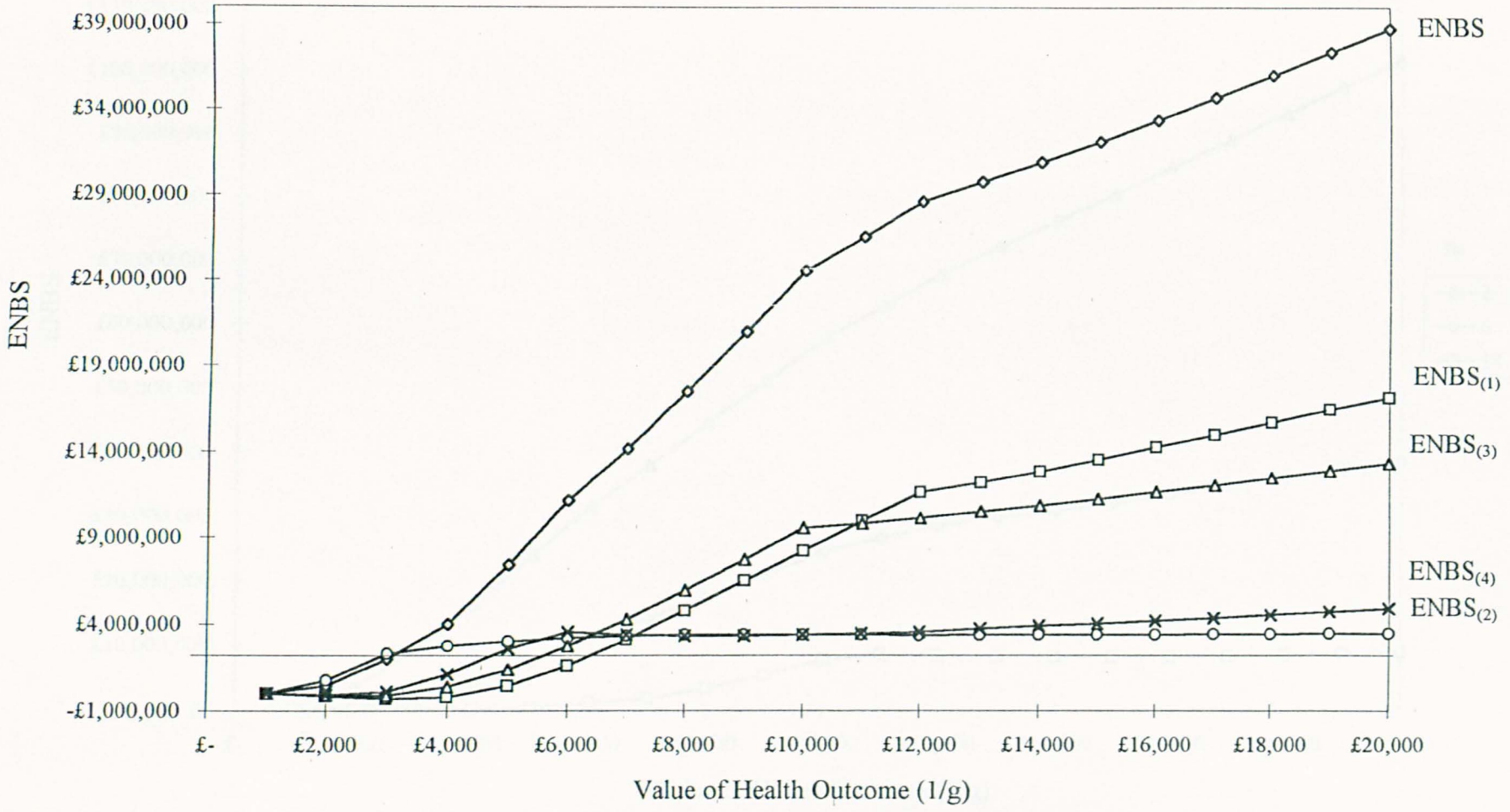


Figure 5.3.8b ENBS and the Strength of Prior Information

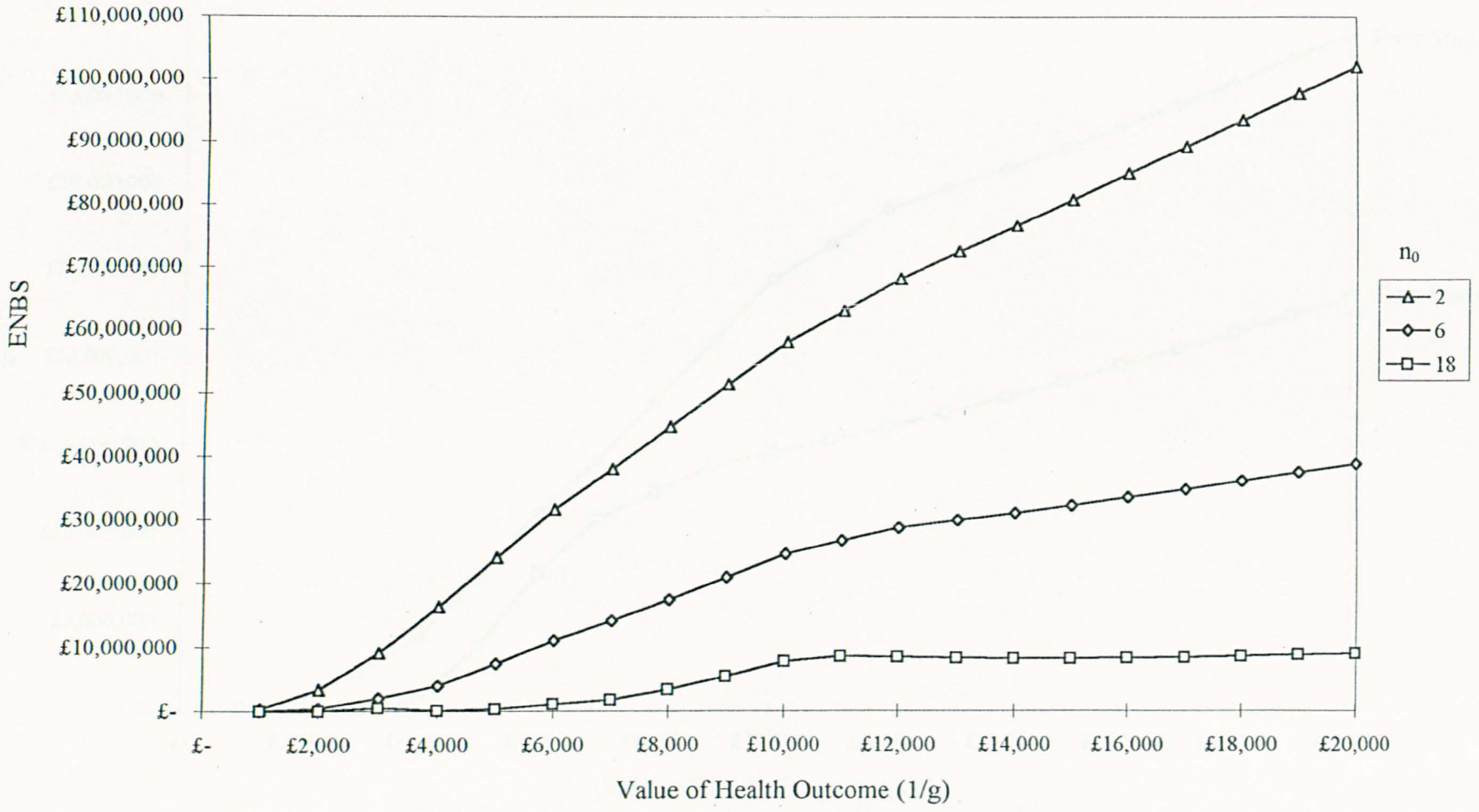


Figure 5.3.9a ENBS at Optimal Sample Size for Two and Four Stage Decision Problems

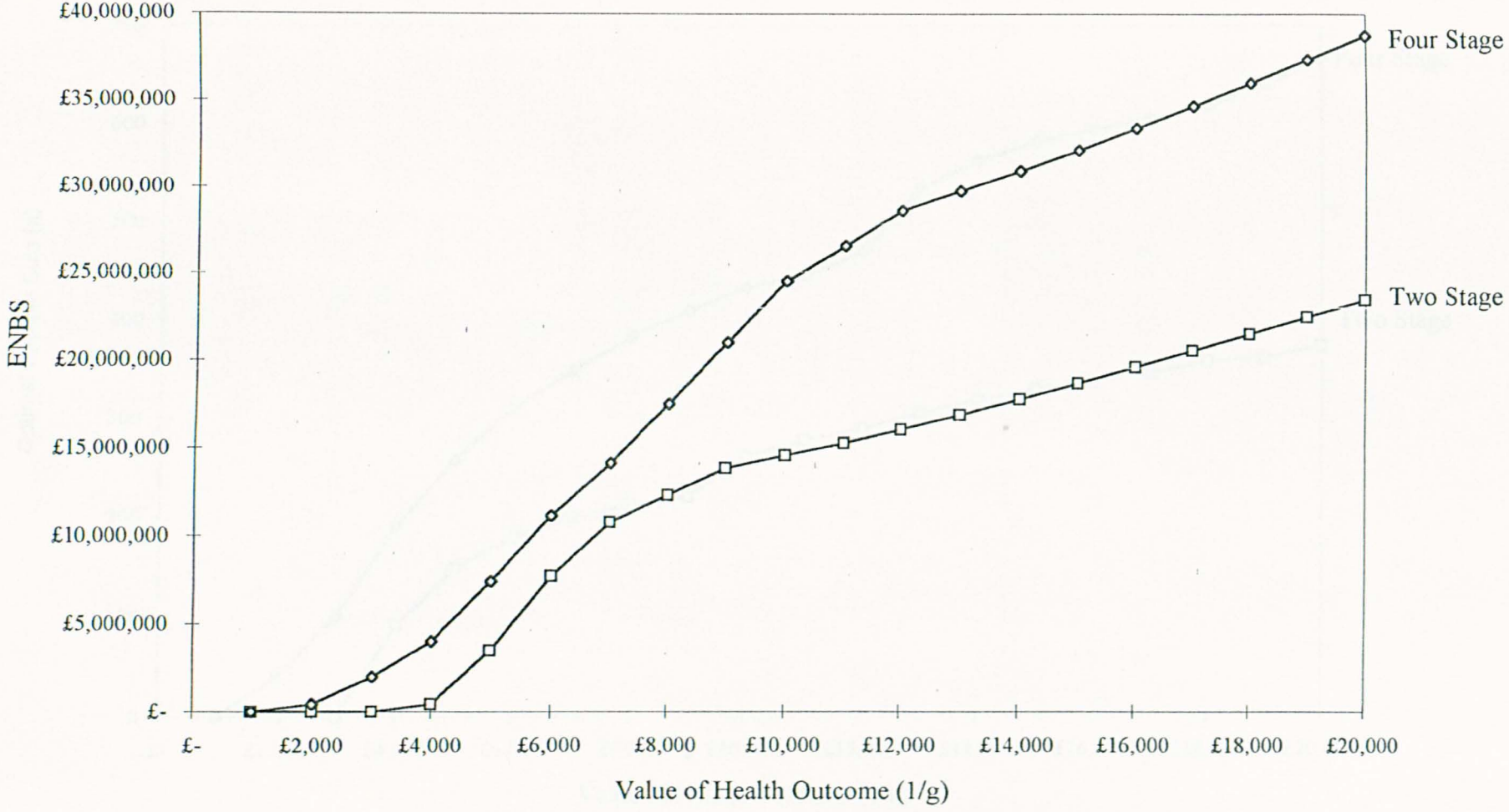


Figure 5.3.9b Optimal Sample Size for the Two and Four Stage Decision Problems

Chapter 6 The Value of Sample Information with Optimal Patient Allocation

Contents: Page

6.1 Introduction

The analysis in chapters 4 and 5 made a number of assumptions and simplifications when establishing the value of sample information. In particular a fixed and equal allocation of trial entrants to the alternative arms of the trial at each stage has been assumed. In this chapter this assumption is relaxed and a dynamic programming approach to optimal patient allocation is proposed. This is applied to the same numerical examples which were considered in previous chapters and demonstrates that optimal patient allocation increases the value of sample information and can be used to identify and rule-out irrelevant alternatives which should not be included in an efficient trial design. Indeed if arbitrary fixed allocation rules are used then the research design will be technically inefficient, the expected net benefits of sampling will be underestimated and there will be a danger that cost-effective research proposals may be rejected at the second hurdle.

The equal allocation of patients between experimental and control arms of a trial is often used and is implicitly justified by assuming that the variance of the outcome of interest for the control arm of the trial is the same as the experimental arm, so that the benefits (reduction in sample variance) of assigning an additional trial entrant to either arm of the trial will be the same [84, 123]. However there is little justification for this rule of precedent when the costs and benefits of allocating a trial entrant to the alternative arms of the trial are explicitly considered. In principle, whether an additional trial entrant should be allocated to a particular arm of a trial should be determined by the marginal benefits of assigning the patient to that arm (which will be determined by the variance of the net benefits of that arm) and the marginal costs of assigning the patient to that arm (which is determined by the additional treatment costs). The only circumstances in which an equal allocation could be justified would be when the variance and the marginal sampling costs of both arms of the trial are the same. In these circumstances the marginal net benefits of assigning the patient to the experimental arm will equal the marginal net benefits of assigning the patient to the control arm.

In the examples considered in the previous chapters the variance of the net benefits of the alternatives at each stage are not assumed to be the same but they are calculated based on the variance of the health outcomes, path probabilities, the values of costs and 1/g. Also the marginal cost of assigning a trial entrant to either of the alternatives at each stage will not be the same. In the single-stage decision problem which was considered in chapter 4 the marginal cost of assigning a patient to the treatment arm will be the additional treatment cost, but the marginal cost of assigning a patient to no treatment will be zero. In these examples the marginal reporting costs are assumed to be negligible but even if reporting costs are substantial it would not alter the key argument that the marginal sampling cost of assigning entrants to different arms of the trial will not be the same and should be taken into account when establishing optimal patient allocation at each stage.

There is a body of literature which considers the optimal allocation of trial entrants in sequential clinical trials were the results of the trial accumulate over time and can be used to assign entrants to the different arms[7]. An example of this type of approach is Bather's "play the winner rule" where patients are assigned to the arm of the trial which appears to be most effective given the accumulated trial results[8, 9]. This approach and others addressing the same problem[73, 75, 151] do not consider the marginal cost of sampling and tend to focus on minimising the potential health cost to individuals enrolled in the trial by establishing allocation rules so that the minimum number of individuals need to be enrolled in the less effective arm of the trial to achieve the specified power and statistical significance. They are primarily concerned with individual medical ethics rather than the collective ethical concerns for the costs of acquiring sample information and the future patients who will benefit from the information generated by the research[133]. These approaches are also primarily concerned with sequential clinical trials were the accumulated results from earlier participants in the trial are available and are used to allocate those entering the trial. This chapter addresses a more fundamental problem of optimal allocation in a fixed sample design where sample information is only

available at the end of the trial. The value of sample information, optimal sample size and the allocation of patients at each stage must be established before any sample information is available.

The analysis in chapter 4 and 5 assumed that the population of future patients who will benefit from the information generated by the proposed research is independent of the numbers enrolled in the trial. This assumption means that those entering the trial are regarded as separate and not part of the population which will ultimately benefit from the results of the research. Clearly this is not the case and those enrolled in a clinical trial are drawn from the same population of patients who will benefit from the sample information. This assumption is relaxed in this chapter and the incidence of patients entering the decision problem is endogenous and depends on the sample size. This means that the population of patients who could benefit from the results of the research are "used up" as the size of the sample considered is increased. This will reduce the benefits of sample information because an additional entrant will provide sample information but there will be one less patient available to benefit from it at the end of the trial. This can be regarded as an additional opportunity cost of sampling because an additional trial entrant will impose an opportunity cost equal to the EVSI for that individual. The results of optimal allocation in section 6.2, 6.3, and 6.4 are compared to the fixed allocation rule using the same example but with the population also dependent on sample size in the same way.

6.2 A Single-Stage Clinical Decision Problem

The analysis of the single-stage treatment decision problem in chapter 4 (the selection of the fallback strategy in chapter 2) assumed a fixed and equal allocation of trial entrants between the two alternative arms of the trial when the expected net benefits of sample information were estimated. The fixed and equal allocation rule used in chapter 4 implicitly assumes that the marginal benefits and marginal costs of assigning a trial entrant to either alternative are equal. In the traditional approach the marginal costs of sampling are not considered and it is often assumed that the variance of the control and experimental arms of the trial will be the same. In these circumstances the marginal benefits of assigning a trial entrant to either alternative would be equal and the fixed and equal allocation rule may be justified if the marginal cost of sampling is ignored.

However, when the marginal costs of sampling are explicitly considered and the variance of expected net benefits of t_1 and t_0 are derived from the variance of each of its components, the fixed and equal allocation rule will not be optimal because the marginal benefits and the marginal costs of assigning an entrant to the alternative arms of the trial will not be equal. The fixed allocation rule which was used in chapter 4 can be relaxed and an optimal patient allocation can be established by making contingent allocation decisions for a given sample size before the optimal sample size is selected. The expected net benefits of sampling using optimal patient allocation will be higher than the fixed allocation rules of chapter 4 and in this example the optimal sample size will be higher and the total cost of sampling will be lower. Proposed research which uses an arbitrary fixed allocation rule will be technically inefficient, the value of the proposed research will be underestimated and cost-effective research my be rejected at the second hurdle.

The same numerical example that was used in chapter 4 is used to illustrate the approach to optimal patient allocation. The decision problem is illustrated in

figure 6.2.1 and this is identical to figure 4.2.1 except that the expected net benefits of sampling and the optimal sample size is estimated in a two-stage process. At stage 2 an optimal contingent allocation between t_1 and t_0 is made based on the expected net benefits of every feasible allocation of each sample entering stage 2. The optimal sample size which maximises the expected net benefits of sampling is selected at stage 1 given that it will be allocated optimally at the second stage. This approach is consistent with the principles of backward induction where the contingent allocation decisions are solved at stage 2 before the expected net benefits of sampling are established and optimal sample size is selected at stage 1.

Figure 6.2.1

The first hurdle for this decision problem is identical to the analysis in chapter 4, because the EVPI depends only on prior information and is unaffected by how the sample is allocated between the alternatives. If the proposed research passes the first hurdle described in chapter 4 then the expected net benefits of sample information must be estimated to construct the second hurdle. In chapter 4, where a fixed allocation rule was used, only one estimate of the expected net benefit of sampling was required for each sample size considered and the maximum ENBS at the optimal sample size provided the value of the second hurdle. Once the fixed allocation rule is relaxed there will be a number of alternative estimates of the ENBS for each sample size.

This problem can be solved in two stages. Following the principles of backward induction the second and final stage is solved first where contingent optimal allocation decisions for a given sample size are established. The allocation of a given sample between t_1 and t_0 will determine both the benefits and the costs of sampling so the EVSI, the cost of sampling, and the ENBS, must be established for every feasible allocation between t_0 and t_1 of each sample size entering stage 2. There are a number of estimates of the ENBS for each sample entering stage 2

because it can be allocated between t_0 and t_1 in a number of different ways. For example two patients entering stage 2 can be allocated in three ways: with two allocated to t_1 and zero to t_0; or one allocated to t_1 and one to t_0; or zero to t_1 and two allocated to t_0. In general there will be n+1 ways to allocate each sample of n entering stage 2.

In this numerical example the sample sizes considered range from zero to 500. This generates 125,500 possible alternative combinations of sample size and allocation between t_1 and t_0 which must be considered and 125,500 estimates of the EVSI, the cost of sampling, and the ENBS, will be required. This is considerably more than the 501 estimates of the ENBS which were required for the same range of sample size when using the fixed allocation rule in chapter 4.

6.2.1 Optimal Allocation at Stage 2

The expected benefits of sampling will depend on both the size of the sample entering stage 2 and the way it is allocated between t_1 and t_0. The benefit of additional sample information is the reduction in the sample variance of the incremental net benefits of t_1 ($\sigma_{n(2)}^2$). The variance of the expected net benefits of t_1 and t_0 are not assumed to be the same and if the variance of the net benefit of t_1 (σ_{t1}^2) is greater than t_0 (σ_{t0}^2) then the marginal benefit (reduction in $\sigma_{n(2)}^2$) of assigning a trial entrant to t_1 will also be greater than t_0. In these circumstances assigning a trial entrant to t_1 would lead to a greater reduction in the uncertainty surrounding the treatment decision (greater reduction in $\sigma_{n(2)}^2$).

EVSI at Stage 2

To establish the expected benefits of sampling an estimate of the EVSI for every feasible allocation of each sample entering stage 2 is required. The $EVSI_{(2)}|n_{(2)},n_{t1}$ is a measure of the expected benefits of sample information given a sample of $n_{(2)}$

entering stage 2 where n_{t1} is allocated to t_1 and $(n_{(2)}-n_{t1})$ is allocated to t_0.

$$EVSI_{(2)}|n_{(2)},n_{t1} = K_t \cdot \sqrt{V}|n_{(2)},n_{t1} \cdot \sigma_{\delta(2)} \cdot L(D_{(2)}|n_{(2)},n_{t1}) \qquad 6.2.1a$$

$$D_{(2)}|n_{(2)},n_{t1} = (\delta_{\delta(2)}-\delta_b)/\sqrt{V}|n_{(2)},n_{t1}$$

$$\sqrt{V}|n_{(2)},n_{t1} = \sigma_{\delta(2)}^2/(\sigma_{\delta(2)}^2 + \sigma_{n(2)}^2)$$

$\delta_{\delta(2)}$ = prior incremental net benefit of t_1

$\sigma_{\delta(2)}^2$ = prior variance of $\delta_{\delta(2)}$

$\sigma_{n(2)}^2 = (\sigma_{t1}^2/n_{t1}) + (\sigma_{t0}^2/(n_{(2)}-n_{t1}))$

The $EVSI_{(2)}|n_{(2)}, nt_1$ takes account of the alternative ways to allocate a given sample entering stage 2 because the sample variance of $\delta_{\delta(2)}$ ($\sigma_{n(2)}^2$) is the sum of the sample variance of the net benefits of t_1 (σ_{t1}^2/n_{t1}) and the sample variance of the net benefits of t_0 ($\sigma_{t0}^2/(n_{(2)}-n_{t1})$). Sample variance and therefore the EVSI not only depends on the size of the sample entering stage 2 ($n_{(2)}$) but also how it is allocated between t_1 (n_{t1}) and t_0 ($n_{(2)}-n_{t1}$). The marginal benefits of allocating an additional entrant to t_1 or t_0 will only be equal if $\sigma_{t1}^2 = \sigma_{t0}^2$ and allocating an equal number of trial entrants to t_1 and t_0 would minimise $\sigma_{n(2)}^2$. In this numerical example σ_{t1}^2 is greater than σ_{t1}^2 (when 1/g=£4,000) and the marginal benefits of assigning a trial entrant to t_1 will be greater than t_0.

However σ_{t1}^2 and σ_{t0}^2 are dependent on the value of 1/g because g is regarded as a constant in the calculation of the variance of expected costs, and as g falls the variance of the expected costs also falls. In this numerical example the variance of the expected costs is a large component of σ_{t1}^2 and as the value of 1/g increases the difference between σ_{t1}^2 and σ_{t0}^2 falls (when 1/g=£20,000 $\sigma_{t1}^2 < \sigma_{t0}^2$). If no sample is allocated to one of the alternatives (either $n_{t1}=0$ or $n_{(2)}=n_{t1}$) then no comparison can be made between the alternatives using sample information and the EVSI will be zero. This may be the optimal allocation in some circumstances and indicates that one alternative can be ruled-out as not relevant because comparing the alternatives using sample information is not efficient.

Cost of Sampling at Stage 2

The optimal patient allocation will also depend on the cost of sampling ($Cs_{(2)}|n_{(2)},n_{t1}$) when trial entrants are allocated to the treatment alternatives. The marginal cost of assigning a trial entrant to t_0 will not be the same as assigning the entrant to t_1 at stage 2, because the marginal cost of sampling includes the additional treatment costs compared to current practice, which in this example is assumed to be t_0 (it is possible that the marginal sampling cost could be negative if the experimental treatment is less costly than current practice). The marginal cost of assigning an entrant to t_1 will include the additional cost of t_1 compared to t_0 ($E(C|t_1)$ -$E(C|t_0)$) but the additional treatment cost of assigning an entrant to t_0 will be zero. It is assumed that the marginal reporting costs are negligible, but this assumption does not weaken the argument that the marginal sampling cost will differ between the alternative arms of the trial as long as the additional treatment cost is one component.

$$Cs_{(2)}|n_{(2)},n_{t1} = (E(C|t_1) - E(C|t_0)).n_{t1} \qquad 6.2.1b$$

Payoff at Stage 2

In this example the cost of assigning trial entrants to t_1 at stage 2 is higher than t_0 but the marginal benefits of assigning entrants t_1 are higher than t_0. The optimal allocation between t_1 and t_0 will involve a trade-off between the additional benefits of assigning entrants to t_1 and the lower costs of assigning entrants to t_0. The $ENBS_{(2)}|n_{(2)},n_{t1}$ is the difference between the expected benefits and the cost of a sample of $n_{(2)}$ entering stage 2 with n_{t1} allocated to t_1 and $n_{(2)}$-n_{t1} allocated to t_0. The payoff at stage 2 from a sample of $n_{(2)}$ entering stage 2 ($\Pi_{(2)}|n_{(2)},n_{t1}$) is the expected net benefits of sample information and can be estimated for every feasible allocation between t_1 and t_0.

$$\Pi_{(2)}|n_{(2)},n_{t1} = ENBS_{(2)}|n_{(2)},n_{t1} \qquad 6.2.1c$$

$$= EVSI_{(2)}|n_{(2)},n_{t1} - (E(C|t_1)-E(C|t_0)).n_{t1}$$

The optimal allocation of any sample entering stage 2 (n_{t1}^*) will be where the payoff (ENBS) reaches a maximum. The $\Pi_{(2)}|n_{(2)},n_{t1}^*$ for each sample entering stage 2 can now be established and optimal contingent allocation decisions can be made. This approach of solving contingent allocative decisions at stage 2 illustrated in table 6.2.1a where the possible sample sizes entering stage 2 are represented by each row, the feasible allocations to t_1 are represented by the columns, and the payoffs are illustrated in the body of the table. The optimal allocation of each sample is the row maximum, and the optimal contingent allocations and associated payoffs are illustrated in the right hand columns.

Table 6.2.1a

Optimal Allocation at Stage 2

The optimal contingent allocation of a sample ($n_{(2)}$) entering stage 2 to t_1 (n_{t1}) for three values of 1/g is illustrated in figure 6.2.2. The fixed and equal allocation of the sample entering stage 2 is represented by the rising diagonal where half the sample is assigned to t_1 and half to t_0. A greater proportion of trial entrants are allocated to t_0 than with the fixed allocation rule because the marginal cost of assigning patients to t_1 is higher than t_0 where the cost of sampling is zero. This difference in marginal sampling cost offsets the higher marginal benefits of assigning entrants to t_1 ($\sigma_{t1}^2>\sigma_{t0}^2$ when 1/g <£20,000). So in this numerical example the optimal allocation of trial entrants will reduce the expected benefits of sample information, but this will be more than offset by the reduction in sampling cost. As the sample size entering stage 2 increases, the marginal benefits of additional sample information fall and the differences in the marginal benefits of assigning trial entrants to t_1 and t_0 will become less significant but the difference in the marginal cost of sampling remains constant. This means that the optimal allocation will change with sample size and in this example because the costs of assigning a patient to t_1 is higher than t_0 the proportion of the sample allocated to

t_1 falls as the sample size increases.

Figure 6.2.2

The optimal allocation is also dependent on the value placed on health outcome for two reasons: firstly because this determines the weight placed on the differences in the benefits and costs of assigning a trial entrant to each alternative, so that less weight is placed on the additional cost of t_1; and secondly the value of g partly determines the variance of the expected costs of each alternative and therefore the marginal benefits of assignment to t_1 and t_0. In this example when the value of 1/g is increased the difference between σ_{t1}^2 and σ_{t0}^2 falls (because the variance of the expected cost is a larger component of σ_{t1}^2), consequently the effect on the marginal benefits and on the weight attached to differences in marginal costs both work in the same direction, and a greater proportion of the sample is assigned to t_1.

6.2.2 Optimal Sample Size at Stage 1

The first stage is simply to select the optimal sample size given that each sample will be allocated optimally at stage 2. The payoff given a sample of $n_{(1)}$ at stage 1 ($\Pi_{(1)}|n_{(1)}$) is the payoff from stage 2 given an optimal allocation between t_1 and t_0 ($\Pi_{(2)}|n_{(2)},n_{t1}^*$). The optimal sample size at stage 1 ($n_{(1)}^*$) can now be identified in table 6.2.1b, where $\Pi_{(1)}|n_{(1)}$ (or $ENBS_{(1)}|n_{(1)}$) reaches a maximum.

$$\Pi_{(1)}|n_{(1)} = \Pi_{(2)}|n_{(2)},n_{t1}^* \qquad 6.2.2$$
$$ENBS_{(1)}|n_{(1)} = ENBS_{(2)}|n_{(2)}, n_{t1}^*$$
$$n_{(1)} = n_{(2)}$$

Once the optimal sample size has been selected at stage 1 the allocation of this sample at stage 2 is given by the contingent allocative decisions that have been

made at stage 2, and which were illustrated in table 6.2.1a and figure 6.2.2. It is the payoff (or the ENBS) at the optimal sample size which is the second hurdle that the proposed research must overcome. This will be greater than the ENBS when the fixed and equal allocation rule is used.

Table 6.2.1b

The two-stage approach to this single-stage decision problem is in fact the full enumeration of all possible allocative decisions for a range of sample size and it does provide an optimal solution to this problem. It is convenient to separate the contingent allocative decision at stage 2 from the selection of optimal sample size at stage 1, because it helps the explanation of the dynamic programming approach which is used to solve the two and four-stage decision problems in section 6.3 and 6.4. In these more complex problems there is a recursive relationship between the payoffs at each stage and the full enumeration of these more complex sequential problems quickly become intractable.

In general the solution to this single-stage problem using a fixed allocation rule will require n+1 estimates of the ENBS where n is the maximum sample size considered, but the solution to the same problem using an optimal allocation of the sample will require $((n+1)^2+n+1)/2$ estimates of the ENBS. Solving this problem using optimal allocation rule requires considerable additional computation. The solution using a fixed allocation rule requires just over one minute of computing time (6 estimates of the ENBS per second) but the solution to the optimal allocation problem requires almost 6 hours of computing time when considering a maximum sample size of 500 (125,751 estimates of the ENBS). The computational requirements become even more extreme in the two and four stage decision problems considered in section 6.3 and 6.4.

6.2.3 Expected Net Benefits of Sample Information

The expected net benefit of sampling at stage 1 given an optimal allocation at stage 2 is illustrated in figure 6.2.3a when the value of 1/g is £4,000. The expected net benefit when a fixed and equal allocation rule is used is illustrated in figure 6.2.3b. The expected net benefit reaches a maximum of £724,970 at a sample size of 116 (n_{t1}=29) when an optimal allocation is used. This is higher than when the fixed allocation rule is used, where the maximum ENBS is £626,920 with an optimal sample size of 76 (n_{t1}=38). The optimal allocation of trial entrants increases the ENBS and the value of the second hurdle. If arbitrary allocation rules are used then the ENBS will be underestimated, the value of the second hurdle will be biased, and there is a danger that proposed research which would be cost-effective will be rejected.

Figure 6.2.3a

Figure 6.2.3b

The expected value of sample information and the costs of sampling are also illustrated in figure 6.2.3a and 6.2.3b. The EVSI increases at a decreasing rate and actually declines in figure 6.2.3a as sample size is increased above 121 (and 185 in figure 6.2.3b). This is because the population of patients who will benefit from sample information is no longer independent of sample size and as the sample size increases this population is "used up" in the trial. This was not the case in chapter 4 where the estimates of the EVSI assumed that the population who would benefit from the information generated by the trial was independent of sample size. Now an additional trial entrant will increase the sample information and increase the EVSI for each individual but will reduce the number of patients who will benefit from this additional information. This can be regarded as an additional opportunity cost of sampling because an additional trial entrant will impose an opportunity cost equal to the EVSI for that individual.

Figure 6.2.3a also shows that the total variable cost of sampling will not increase at a constant rate as in figure 6.2.3b but will increase at a decreasing rate as the optimal allocation of trial entrants changes as sample size increases. In this case the marginal cost of sampling declines because the proportion of the sample allocated to t_1 declines with sample size in figure 6.2.2.

The ENBS in figure 6.2.3b is negative when the sample size is low ($n_{(1)}<6$), because the benefit of small amounts of sample information is less than the costs. However when an optimal allocation is used the ENBS will not be negative at low sample sizes because all the sample can be allocated to t_0 at zero cost. If all the sample is allocated to t_0 no comparison between t_0 and t_1 using sample information is possible and the EVSI, the cost of sampling, and the ENBS is zero (when $n_{(1)}<5$ in figure 6.2.3a).

Value of Health Outcome

The value of sample information depends on the value of 1/g and the budgetary restrictions on service provision. The ENBS when 1/g is increased to £10,000 is illustrated in figure 6.2.4a when optimal allocation at stage 2 is used, and this can be compared to the ENBS in figure 6.2.4b when a fixed and equal allocation rule is used. The optimal allocation of trial entrants generates a maximum ENBS of £4,308,900 at an optimal sample size of $n_{(1)}^{*}=123$ ($n_{t1}^{*}=46$). This is higher than when the fixed allocation rule is used where the maximum ENBS =£4,234,305 and $n_{(1)}^{*}=110$ ($n_{t1}=55$). The value of information increases with the value of 1/g and in this example the optimal sample size also increases. Figure 6.2.4a and 6.2.4b also shows that the EVSI not only increases at a decreasing rate but will ultimately decline because the population benefits are no longer independent of sample size. Figure 6.2.4a also illustrates that the marginal cost of sampling at stage 1 is no longer constant but declines because as sample size increases the proportion of trial entrants allocated to t_1 at stage 2 declines.

Figure 6.2.4a

Figure 6.2.4b

The expected net benefit of sampling when the value of 1/g is increased to £20,000 is illustrated in figure 6.2.5a when an optimal allocation is used and in figure 6.2.5b when a fixed and equal allocation rule is used. With optimal allocation the ENBS reaches a maximum of £4,390,086 at an optimal sample size of 153 (n_{t1}^{*}=57). This is greater than when a fixed allocation rule is used where the maximum ENBS is £4,297,885 at an optimal sample size of 140 (nt_1=70). The EVSI will decline when the $n_{(1)}$>172 in figure 6.2.5a and when $n_{(1)}$>191 in figure 6.2.5b. The marginal costs of sampling in figure 6.2.5a also declines because a smaller proportion of the sample is allocated to t_1 at stage 2 as the sample size is increased.

Figure 6.2.5a

Figure 6.2.5b

Optimal Patient Allocation

The optimal patient allocation increases the value of information and it will also change the optimal scale of research. The optimal sample size when optimal patient allocation is used can be compared to optimal sample size with the fixed allocation rule in figure 6.2.6. In this numerical example the optimal sample size is higher when optimal allocation is used. Indeed figure 6.2.6 shows the circumstances in which optimal allocation will result in a positive sample size but the fixed rule indicates that no sample should be taken (when 1/g=£3,200). So it is possible that by using optimal patient allocation sampling may become efficient whereas with a fixed rule a trial would not be cost-effective and the optimal sample size would be zero.

Figure 6.2.6

The allocation of the sample is also illustrated in figure 6.2.6 and this demonstrates that optimal patient allocation assigns fewer entrants to t_1, because the marginal cost of assigning trial entrants to t_1 is higher. Consequently the total variable cost of the research is lower when optimal patient allocation is used despite the fact that the total sample size is higher. The optimal sample size using fixed allocation rules is not the same as in chapter 4 because the population benefits are not independent of sample size and in this example the discontinuity in the relationship between sample size and 1/g occurs at 1/g=£6,500 where the clinician would be indifferent between t_1 and t_0. At this point the opportunity cost of enrolling an additional trial entrant (EVSI) reaches a peak.

Figure 6.2.7

The difference between the ENBS, the EVSI, and the cost of sampling when optimal and fixed allocation rules are used is illustrated in figure 6.2.7. This demonstrates that the optimal patient allocation will increase the maximum expected net benefits of sampling and the value of the second hurdle. If an arbitrary allocation rule is used to design a clinical trial then the potential benefits of this research may be underestimated, the costs may be overestimated, and there will be a danger that proposed research which is cost-effective may be rejected at the second hurdle. The design and scale of proposed research will also be biased and by using arbitrary rules those proposals which pass the second hurdle will be conducted at less than the technically efficient scale with possibly higher cost and lower net benefits. In this numerical example the optimal allocation of trial entrants means that fewer entrants are allocated to t_1 and this reduces the benefits of sample information (when 1/g>£4,000), but this is more than offset by the reduction in the costs of sampling.

Implications for Research Design

This approach to optimal allocation of trial entrants explicitly considers the marginal benefit and marginal cost of assigning trial entrants to the alternative arms of the trial. The marginal costs and benefits of allocating trial entrants to alternative arms may differ and in these circumstances an optimal allocation should be applied when designing efficient clinical research. This numerical example demonstrates that the optimal allocation of trial entrants is not simply an issue for sequential clinical trials but is a more fundamental issue which is also relevant to the fixed sample design considered here, where all the results are available at the same time at the end of the trial. There appears to be little justification for using arbitrary fixed rules of precedent and these arbitrary rules will lead to inefficient research design and errors at the second hurdle. The approach taken to this single-stage decision problem involves the full enumeration of all feasible allocations of each sample size considered and therefore provides the optimal solution to this problem. This requires a considerable increase in computation compared to the fixed allocation rule used in chapter 4, however the additional computation becomes extreme when more complex decision problems are considered. Full enumeration of more complex decision problems is simply not feasible and a dynamic programming solution to the two and four-stage decision problems is presented in section 6.3 and 6.4.

6.3 A Two-Stage Sequential Clinical Decision Problem

The approach to optimal patient allocation can also be applied to sequential clinical decision problems. In this section it is applied to the two-stage decision problem which was considered in chapter 5.2. This is the same problem which was also considered in chapter 2, where the clinician must decide whether to use a diagnostic test and treat according to the test results, or use the fallback treatment strategy. This decision problem is illustrated in figure 6.3.1 and it is identical to the problem illustrated in figure 5.2.1 except that it is solved in three stages when optimal patient allocation is required.

Figure 6.3.1

At stage 3 an optimal contingent allocation between t_1 and t_0 for each sample entering stage 3 is made based on the expected net benefit of every feasible allocation of each sample. At stage 2 an optimal contingent allocation between the test and no test arm of the trial is made for each sample entering stage 2. This is based on the expected net benefit of sampling at stage 2 and the payoff from the optimal allocation of the sample assigned to the no test arm which enters stage 3. Finally the optimal sample size which will maximise expected net benefits can be selected at stage 1. Solving contingent allocation decisions before the optimal sample size is selected at stage 1 is consistent with the principle of backward induction and can be characterised as a simple three-stage dynamic programme, where the payoff for each sample size considered at stage 1 is the expected net benefit given that an optimal patient allocation policy will be followed from stage 1 to the end. This approach of making contingent allocative decisions avoids the full enumeration of all possible allocations of each sample and dramatically reduces the computation required. Indeed the full enumeration of all the possible alternatives in a sequential decision problem for the range of possible sample sizes considered here is not feasible. The approach taken in this chapter provides a practical solution to the problem.

The same numerical example that was used in chapter 5.2 is used to illustrate this approach to optimal patient allocation. The expected net benefits using optimal patient allocation are compared to the fixed and equal allocation that was used in chapter 5.2 but in this section (as in section 6.2) the population of patients who will benefit from the information generated by the research is no longer assumed to be independent of the sample size selected. Consequently the results with optimal patient allocation are compared to the fixed allocation rule when both have endogenous population benefits.

The estimates of the EVPI and the value of the first hurdle is identical to the estimates in chapter 5.2 because the EVPI is based only on prior information and is not affected by the allocation rules. If the proposed research passes the first hurdle the second hurdle must be constructed to demonstrate that it will be cost-effective and to establish the technically efficient scale of the research. Just as in section 6.2 the second hurdle is dependent on the allocation rules used to assign trial entrants to the alternative arms of the trial at each stage because this will determine both the expected benefits and the costs of a given sample.

6.3.1 Optimal Allocation at Stage 3

At stage 3 the optimal contingent allocation between t_1 and t_0 for each sample entering stage 3 must be made based on estimates of the expected net benefits of sampling for every feasible allocation of each sample entering stage 3. The optimal contingent allocation will be where the expected net benefit reaches a maximum for a given sample size. This is the same problem that was considered in the previous section of this chapter except that it is now a contingent treatment decision which follows the initial diagnostic decision.

Once the fixed allocation rule is relaxed the expected benefits of sampling depend not only on the sample size entering stage 3 but also on the way this is allocated to

t_1 and t_0. An estimate of the expected benefits of sampling for every feasible allocation of each sample entering stage 3 is required. The $EVSI_{(3)}|n_{(3)},n_{t1}$ is a measure of the expected benefit of sampling given a sample of $n_{(3)}$ entering stage 3 with n_{t1}allocated t_1 and $(n_{(3)}-n_{t1})$ allocated to t_0 and from 6.2.1a:

$$EVSI_{(3)}|n_{(3)},nt_1 = K_t \cdot \sqrt{Vn_{(3)},n_{t1}} \cdot \sigma_{0(3)} \cdot L(D_{(3)}|n_{(3)},nt_1) \qquad 6.3.1a$$

$$D_{(3)}|n_{(3)},nt_1 = (\delta_{0(3)}-\delta_b)/\sqrt{Vn_{(3)},n_{t1}}$$

$$\delta_{0(3)} = (E(U|t_1)-g.E(C|t_1)) - (E(U|t_0)-g.E(C|t_0))$$

$$\sqrt{Vn_{(3)},n_{t1}} = \sigma_{0(3)}^2/(\sigma_{0(3)}^2+\sigma_{n(3)}^2)$$

$$\sigma_{n(3)}^2 = (\sigma_{t1}^2/n_{t1})+(\sigma_{t0}^2/(n_{(2)}-n_{t1}))$$

In this numerical example σ_{t1}^2 is greater than σ_{t0}^2 and the marginal benefits of assigning a trial entrant to t_1 will be greater than t_0. However σ_{t1}^2 and σ_{t0}^2 are partly determined by the value of 1/g, and the difference between σ_{t1}^2 and σ_{t0}^2 declines as 1/g is increased.

The total variable cost of sampling $(Cs_{(3)}|n_{(3)},n_{t1})$ is also determined by how the sample is allocated between t_1 and t_0 because the additional treatment cost of assigning a patient to t_0 will be zero but the marginal cost of assigning a trial entrant to t_1 will be the additional treatment cost of t_1.

$$Cs_{(3)}|n_{(3)},n_{t1} = (E(C|t_1) - E(C|t_0)).n_{t1} \qquad 6.3.1b$$

The payoff at stage 3 given a sample of $n_{(3)}$ entering stage 3 $(\Pi_{(3)}|n_{(3)},n_{t1})$ is the expected net benefits of sampling given a sample of $n_{(3)}$ entering stage 3, with n_{t1} allocated to t_1 and $n_{(3)}-n_{t1}$ allocated to t_0.

$$\Pi_{(3)}|n_{(3)},n_{t1} = ENBS_{(3)}|n_{(3)},n_{t1} = EVSI_{(3)}|n_{(3)},n_{t1} - Cs_{(3)}|n_{(3)},n_{t1} \qquad 6.3.1c$$

The payoff can now be established for every feasible allocation of each sample

entering stage 3. The optimal contingent allocation is the allocation where the ENBS reaches a maximum for each sample size. This is illustrated in table 6.3.1a. The payoffs in the body of the table and the optimal contingent decision are the same as at stage 2 from table 6.2.1a in the previous section of this chapter.

Table 6.3.1a

The optimal contingent allocation of a sample entering stage 3 for this numerical example is illustrated in figure 6.3.2a. These contingent allocations for three values of 1/g are the same as the contingent optimal allocation at stage 2 in section 6.2. The marginal cost of allocating entrants to t_1 results in a smaller proportion of the sample allocated to t_1 than when using the fixed and equal allocation rule in chapter 5.

Figure 6.3.2a

6.3.2 Optimal Allocation at Stage 2

At stage 2 the optimal contingent allocation between the test (t_e) and the no test (nt_e) arm of the trial for each sample entering stage 2 must be made based on the expected net benefit of sampling at stage 2 and the payoff given the optimal allocation of the sample which will enter stage 3.

Payoff at Stage 2

A sample entering stage 2 ($n_{(2)}$) which is allocated between test (n_{te}) and no test alternatives ($n_{(2)}$-n_{te}) will generate benefits and costs of sampling at stage 2, but when part of the sample entering stage 2 is allocated to the no test arm of the trial it will enter stage 3 ($n_{(3)}$=$n_{(2)}$-n_{te}) and will be allocated optimally, generating payoffs of $\Pi_{(3)}|n_{(3)},n_{t1}^{*}$ which where illustrated in table 6.3.1a. This is a recursive

relationship [10, 11] where the payoff at stage 2 is partly determined by the payoffs associated with the optimal contingent allocation at stage 3. So the payoffs $(\Pi_{(2)}|n_{(2)},n_{te})$ and the optimal contingent allocation at stage 2 will be determined by the ENBS at stage 2 and the payoff from allocating part of the sample to the no test arm given that an optimal allocation policy will be followed at stage 3.

$$\Pi_{(2)}|n_{(2)},n_{te} = ENBS_{(2)}|n_{(2)},n_{te} + \Pi_{(3)}|n_{(3)},n_{t1}^{*} \qquad 6.3.2a$$

$$n_{(3)} = n_{(2)}-n_{te}$$

The payoffs for every feasible allocation of each sample entering stage 2 are illustrated in table 6.3.1b (when 1/g=£4,000). The possible samples entering stage 2 are represented by each row. The feasible allocations to the testing arm of the trial (n_{te}) are represented by the columns, and the sample allocated to the no test arm $n_{(2)}-n_{te}$ will enter stage 3 and will be allocated optimally in table 6.3.1b. The optimal contingent allocation at stage 2 and the associated payoffs are the row maximums and are illustrated on the right hand columns of the table.

Table 6.3.1b

EVSI at Stage 2

The payoffs in the main body of table 6.3.1b require estimates of the $ENBS_{(2)}|n_{(2)},n_{te}$ and both the expected benefit and cost of sampling will be determined by the sample size and how it is allocated. The expected benefit of sampling (reduction in sample variance $(\sigma_{n(2)}^{2})$) depends on $n_{(2)}$ and how it is allocated between the alternatives. The marginal benefits of assigning a trial entrant to either the test or no test arms will be determined by the variance of the net benefit of testing (σ_{te}^{2}) and the variance of the net benefit of not testing (σ_{nte}^{2}). In this numerical example σ_{te}^{2} is greater than σ_{nte}^{2} and the marginal benefits of assigning an entrant to the test arm will be greater than the no test arm of the trial. However the difference between σ_{te}^{2} and σ_{nte}^{2} declines as 1/g increases. The

$EVSI_{(2)}|n_{(2)},n_{te}$ is a measure of the expected benefit of an sample of $n_{(2)}$ entering stage 2 with n_{te} allocated to the test arm and $n_{(2)}-n_{te}$ allocated to the no test arm of the trial.

$$EVSI_{(2)}|n_{(2)}, n_{te} = K_t \cdot \sqrt{Vn_{(2)},n_{te}} \cdot \sigma_{0(2)} \cdot L(D_{(2)}|n_{(2)}, n_{te}) \qquad 6.3.2b$$

$$D_{(2)}|n_{(2)}, n_{te} = (\delta_{0(2)}-\delta_b)/ \sqrt{Vn_{(2)},n_{te}}$$

$$\delta_{0(2)} = (E(U|t_e)-g.E(C|t_e)) - (E(U|nt_e)-g.E(C|nt_e))$$

$$\sqrt{Vn_{(2)},n_{te}} = \sigma_{0(2)}^2/(\sigma_{0(2)}^2 + \sigma_{n(2)}^2)$$

$$\sigma_{n(2)}^2 = (\sigma_{te}^2/n_{te}) + (\sigma_{nte}^2/(n_{(2)}-n_{te}))$$

The prior incremental net benefit ($\delta_{0(2)}$), the prior variance of $\delta_{0(2)}$ ($\sigma_{0(2)}^2$), and the variance of the net benefits of not testing (σ_{nte}^2) depend on the sample assigned to nt_e which will enter stage 3 ($n_{(3)}=n_{(2)}-n_{te}$). This is the same problem of establishing the net benefits of sampling at stage 1 in chapter 5.2 when a fixed allocation rule is used. But now the prior mean and the variance of the net benefits of not testing depend not only on the sample allocated to the no test arm, but also on the way this sample will be allocated between t_1 and t_0 at stage 3.

The prior net benefits of not testing $(E(U|nt_e)-g.E(C|nt_e))$ and therefore $\delta_{0(2)}$ is dependent on the sample allocated to the no test alternative at stage 2 which will enter stage 3, and from 5.2.8a and 5.2.8b:

$$E(U|nt_e)-g.E(C|nt_e) = p(\delta_{x(3)}>\delta_{x(3)}^*).(E(U|t_1)-g.E(C|t_1)) \qquad 6.3.2c$$
$$+ 1-p(\delta_{x(3)}>\delta_{x(3)}^*).(E(U|t_0)-g.E(C|t_0))$$

$$p(\delta_{x(3)}>\delta_{x(3)}^*) = p(Z>((\delta_{0(3)}>\delta_{x(3)}^*)/\sigma_{n(3)}))$$

$$\sigma_{n(3)}^2 = (\sigma_{t1}^2/n_{t1}^*) + (\sigma_{t0}^2/(n_{(3)}-n_{t1}^*))$$

The value of $p(\delta_{x(3)}>\delta_{x(3)}^*)$ is determined by the sample assigned to the no test arm ($n_{(2)}-n_{te}=n_{(3)}$) which will enter stage 3 and the way it is allocated between t_1 and t_0. The value of $E(U|nt_e)-g.E(C|nt_e)$ is calculated assuming an optimal

allocation policy is followed at stage 3, and this is the prior information used to establish the EVSI at stage 2.

The prior variance of $(E(U|nt_e)\text{-}g.E(C|nt_e))$ and therefore $\sigma_{0(2)}^2$ are also dependent on the sample allocated to nt_e which enters stage 3. This is because it determines the value of $p(\delta_{x(3)}>\delta_{x(3)}^*)$ and because the prior variance of $(E(U|nt_e)\text{-}g.E(C|nt_e))$ at stage 2 is based on the posterior variance of $(E(U|t_1)\text{-}g.E(C|t_1))$ and $(E(U|t_0)\text{-}g.E(C|t_0))$ from stage 3. The posterior variance of $(E(U|t_1)\text{-}g.E(C|t_1))$ and $(E(U|t_0)\text{-}g.E(C|t_0))$ is a combination of the prior and sample variance of each and is calculated for each sample assigned to the no test arm assuming that an optimal allocation policy is followed at stage 3. The population variance of the net benefit of not testing (σ_{nte}^2) is also partly determined by the sample entering and allocated at stage 3, because this determines the value of $p(\delta_{x(3)}>\bar{\delta}_{x(3)}^*)$. The expected benefit of every feasible allocation of each sample entering stage 2 can now be established given that a sample allocated to the no test arm will be allocated optimally at stage 3.

Cost of Sampling at Stage 2

The cost of sampling at stage 2 $(Cs_{(2)}|n_{(2)},n_{te})$ is the additional treatment cost. The marginal cost of assigning an entrant to the no test arm of the trial at stage 2 will be zero because the additional cost of assigning a trial entrant to t_1 has already been taken into account in the estimates of the ENBS at stage 3. However the additional treatment cost of assigning an entrant to the testing arm of the trial will be the additional expected cost of the testing strategy.

$$Cs_{(2)}|n_{(2)},n_{te} = (E(C|t_e)\text{-}E(C|t_0)).n_{te} \qquad 6.3.2d$$

The $ENBS_{(2)}|n_{(2)},n_{te}$ can now be established and is the difference between the expected benefit and the expected cost of a sample entering stage 2 with n_{te} allocated to the test arm of the trial and $n_{(2)}\text{-}n_{te}$ allocated to the no test arm.

$$ENBS_{(2)}|n_{(2)},n_{te} = EVSI_{(2)}|n_{(2)},\ n_{te} - Cs_{(s)}|n_{(2)},\ n_{te} \qquad 6.3.2e$$

Optimal Allocation at Stage 2

The optimal allocation will be where the marginal payoff of assigning a trial entrant to test arm is equal to the marginal payoff of assigning the entrant to the no test arm, or where the $\Pi_{(2)}|n_{(2)},n_{te}$ reaches a maximum for each sample entering stage 2 (row maximum in table 6.3.1b). The optimal contingent allocation of the sample entering stage 2 for this numerical example is illustrated in figure 6.3.2b for three values of 1/g. The optimal allocation to the testing arm of the trial can be compared to the fixed allocation rule used in section 5.2 where half the sample is allocated to the test and half to the no test arm.

Figure 6.3.2b

Figure 6.3.2b illustrates that the optimal allocation depends on both the sample size entering stage 2 and the value of 1/g because this will determine the weight placed on the differences in the marginal cost and benefit of assigning entrants to either t_e or to nt_e. In this example the variance of the test arm is greater than the no test arm so the marginal benefits of allocating an entrant to t_e will be greater than nt_e. However the marginal costs of assigning an entrant to te is greater than nt_e. When the value of 1/g=£4,000 n_{te}=0 and the optimal allocation is to assign all the sample to the no test arm. In these circumstances it is not efficient to compare the test and no test strategies using sample information and the testing arm of the trial can be ruled out as an irrelevant alternative. However if the value of 1/g is increased to £10,000 then less weight is placed on the additional costs of assigning entrants to the testing arm of the trial and more weight is placed on the additional benefits. Now it is efficient to assign entrants to t_e and the test arm becomes a relevant alternative.

6.3.3 Optimal Sample Size at Stage 1

The first stage is simply to select the optimal sample size given that a sample considered at stage 1 will enter stage 2 and will be allocated optimally between test and no test arms, and the sample entering stage 3 will be allocated optimally between t_1 and t_0. The payoff from a sample of $n_{(1)}$ at stage 1 ($\Pi_{(1)}|n_{(1)}$) is simply the payoff given that an optimal patient allocation policy is followed at each subsequent stage.

$$\Pi_{(1)}|n_{(1)} = \Pi_{(2)}|n_{(2)},n_{te}^{*} \qquad 6.3.3$$

$$ENBS_{(1)}|n_{(1)} = ENBS_{(2)}|n_{(2)},n_{te}^{*} + ENBS_{(3)}|n_{(3)},n_{t1}^{*}$$

$$n_{(1)} = n_{(2)}$$

$$n_{(3)} = n_{(2)}-n_{te}$$

The optimal sample size at stage 1 ($n_{(1)}^{*}$) can now be selected in table 6.3.1c and will be where $\Pi_{(1)}|n_{(1)}$ ($ENBS_{(1)}|n_{(1)}$) reaches a maximum. This is the value of the second hurdle for this two-stage decision problem and will be higher than when a fixed and equal allocation rule is used. Once the optimal sample size has been selected at stage 1 the allocation of this sample between each alternative at each stage is given by the contingent allocative decisions that have already been made at stage 2 and 3 and which were illustrated in tables 6.3.1b and 6.3.1a, and figures in 6.3.2b and 6.3.2a.

Table 6.3.1c

The Dynamic Programming Approach

This approach to optimal patient allocation dramatically reduces the computation required to establish the optimal sample size and the value of the second hurdle. The optimal contingent allocation which has been made at stage 3 reduces the number of alternatives which must be considered at stage 2 because only the

payoffs from stage 3 given optimal contingent allocation need to be considered. Only one estimate of $\Pi_{(3)}|n_{(3)},n_{t1}^{*}$ is required for each sample allocated to the no test arm, rather than estimates for all the feasible allocations of each sample. This is a considerable reduction in the computation required because using this approach the optimal allocation of a sample at stage 2 requires $((n+1)^2+(n+1))/2$ estimates of $\Pi_{(3)}|n_{(3)},n_{t1}$ at stage 3 and $((n+1)^2+(n+1))/2$ estimates of $\Pi_{(2)}|n_{(2)},n_{te}$ at stage 2, where n is the maximum sample size considered.

In this numerical example the maximum sample considered is 500 and a total of 251,502 estimates of payoff are required to solve the optimal contingent allocation at stage 2 and stage 3 (12 hours computer time at 6 estimates per second). However the full enumeration of all feasible allocations of the sample at stage 2 and stage 3 would require $125,751^2$ estimates which would take almost 84 years of computing time. Even if a 100-fold increase in computing speed was possible the solution would still require more than 10 months of computing time. Clearly the full enumeration of all possible alternatives is not feasible and this surprising result is a consequence of the problem of dimensionality[10, 11]. When A is the number of alternative combinations of sample size and allocation at each stage, and S is the number of stages in the decision problem then total number of alternatives will be A^S. The simple dynamic programming approach[99, 124] to the problem taken in this chapter provides a feasible and practical solution by reducing the number of alternatives which must be considered to A*S for an S stage decision problem[10].

6.3.4 Expected Net Benefit of Sample Information

The expected net benefits of sampling at each stage with optimal patient allocation at stage 3 and stage 2 is illustrated in figure 6.3.3a and this can be compared to the expected net benefits when a fixed and equal allocation rule is used in figure 6.3.3b, when 1/g=£4,000. The $ENBS_{(1)}$ when a optimal allocation rule is used

reaches a maximum of £724,970 at an optimal sample size of 116 (n_{te}^{*}=0, and n_{t1}^{*}=29) with all the sample allocated to the no test arm and the treatment decision at stage 2, consequently $ENBS_{(2)}$=0 and $ENBS_{(1)}$=$ENBS_{(3)}$. This is greater than when the fixed allocation rule is used where the maximum $ENBS_{(1)}$ is £362,366 at an optimal sample size of $n_{(1)}^{*}$=80. This is because the negative $ENBS_{(2)}$ from allocating half the sample to test alternative at stage 2 in figure 6.3.3b can be avoided when the optimal allocation rule is used. In this example the optimal allocation reduces the benefits from sampling (because no sample is allocated to te which has a higher variance of net benefits), but this is more than off-set by the reduction in the cost of sampling.

Figure 6.3.3a

Figure 6.3.3b

The optimal allocation of the sample at each stage allows irrelevant or inefficient alternatives to be identified and ruled-out explicitly and consistently because it is possible to allocate none of the sample to that arm of the trial. This is not the case when a fixed allocation rule is used in chapter 5 because half the sample is always allocated to each alternative at each stage irrespective of the costs and benefits. In this example the optimal allocation enables no sample to be assigned to the test arm of the trial, and the test arm can be ruled-out as an irrelevant alternative because the comparison of the test and no test alternative using sample information is not cost-effective. Once the test alternative has been ruled-out as irrelevant the optimal trial design is identical to single-stage problem in section 6.2 with the same optimal sample size and the same expected net benefits.

Value of Health Outcome

The ENBS, the optimal sample size, and optimal allocation is dependent on the value of 1/g and the budgetary restrictions on service provision. The ENBS when the value of 1/g is increased to £10,000 is illustrated in figure 6.3.4a when the

sample is allocated optimally and in figure 6.3.4b when the fixed allocation rule is used. In figure 6.3.4a the ENBS reaches a maximum of £12, 155,102 at an optimal sample size of 136 (n_{te}^{*}=68 and n_{t1}^{*}=31). The testing arm at stage 2 is now a relevant alternative because less weight is placed on the marginal costs of assigning entrants to the test arm and the additional benefits are valued more highly. Half the sample is allocated to the testing arm and this demonstrates that what can be regarded as a relevant alternative is dependent on the value placed on health outcome. Once again the maximum ENBS and optimal sample size with fixed allocation is lower (ENBS=£12,098,145 and $n_{(1)}^{*}$=133). In this case the expected benefits of sampling are higher and the costs of sampling are lower with optimal allocation despite the fact that the optimal sample size is higher.

Figure 6.3.4a

Figure 6.3.4b

The ENBS when the value of 1/g is increased to £20,000 is illustrated in figure 6.2.5a when optimal patient allocation is used and in figure 6.2.5b when a fixed allocation rule is used. In figure 6.3.5a the ENBS reaches a maximum of £18,499,910 at an optimal sample size of 147 (n_{te}*=86, and n_{t1}*=28), and now more than half of the sample is assigned to the testing arm of the trial at stage 2. This is because more weight is placed on additional benefits of assigning the sample to the testing arm and less weight is placed on the additional costs (this was also illustrated in figure 6.3.2b). Clearly the testing arm is a relevant alternative and in figure 6.3.5a the greatest share of the $ENBS_{(1)}$ is accounted for by the net benefits of sampling at stage 2.

Figure 6.3.5a

Figure 6.3.5b

The ENBS when the fixed allocation rule is used in figure 6.3.5b is lower (£18,313,023) but the optimal sample size is higher (153). So in this case the

optimal patient allocation leads to a smaller optimal sample size, but this smaller sample generates a higher expected benefit which off-sets the higher sampling cost. This is because more entrants are assigned to the testing arm of the trial where the benefits and costs are highest. This example demonstrates that the optimal patient allocation at each stage of the decision problem may increase or reduce optimal sample size and may also increase or reduce the costs of sampling depending on the particular example and the value of 1/g. However what is clear is that optimal allocation will increase the expected net benefits of sampling and the value of the second hurdle.

Optimal Patient Allocation

The optimal sample size and allocation to the testing arm of the trial for both fixed and optimal allocation rules can be compared in figure 6.3.6 for a range of values of 1/g. As already noted the optimal allocation of the sample can lead to greater or smaller optimal sample size and more or less assigned to the testing arm of the trial depending on the value of 1/g. However figure 6.3.6 also illustrates circumstances where optimal allocation leads to a positive sample size when no sample would be taken if a fixed rule is used. When 1/g<£4,000 no sample is taken when a fixed allocation rule is used because it forces trial entrants to be assigned to the costly testing arm of the trial. However when 1/g≥£3,000 the optimal allocation of trial entrants leads to positive sample size, demonstrating that optimal patient allocation can lead to proposed research being cost-effective when an arbitrary fixed rule would make it inefficient.

Figure 6.3.6

The difference in the $ENBS_{(1)}$, $EVSI_{(1)}$ and the cost of sampling at the optimal sample size between the optimal and fixed allocation is illustrated in figure 6.3.7 for a range of values of 1/g. This demonstrates that the optimal allocation of trial entrants will increase the expected net benefits of sampling. This may be achieved

by a reduction in the cost of sampling (possibly with a reduction in the expected benefits of sampling) or alternatively it may be achieved by an increase in the expected benefits of sampling (possibly with an increase in the cost of sampling). These differences are determined by the value of 1/g because this determines the relative weight placed on the additional benefits and costs of assigning entrants to each arm of the trial.

Figure 6.3.7

The way that the optimal allocation of trial entrants can be used to identify relevant alternatives explicitly and consistently can be illustrated in figure 6.3.8a and 6.3.8b where the maximum $ENBS_{(1)}$ and the optimal sample size for the single-stage decision problem considered in 6.2 can be compared with the two-stage decision problem considered here. When the value of 1/g≤£4,000 the $ENBS_{(1)}$ for the two-stage decision problem is the same as the $ENBS_{(1)}$ for the single-stage decision problem and the optimal sample size is also the same for this range of value of 1/g. The test alternative at stage 2 is not a relevant alternative and should not be included in the design of prospective research. This approach to optimal patient allocation allows no sample to be allocated to an alterative at each stage and if it is optimal not to allocate a sample to an alternative then it can be regarded as irrelevant and can be excluded from prospective research.

Figure 6.3.8a

Figure 6.3.8b

In this example when 1/g≤£4,000 the two-stage decision problem is identical to the single-stage decision problem because the testing arm is an irrelevant alternative and no sample is assigned to testing at stage 2. However when the value of 1/g>£4,000 less weight is placed on the higher marginal cost of assigning entrants to the testing arm of the trial and testing is a relevant alternative. An analysis which simplified this two-stage decision problem to a single-stage

treatment decision would underestimated the value of the research and seriously bias the trial design. This demonstrates that what are relevant alternatives depends on the value of 1/g and alternatives cannot be ruled-out as irrelevant before the shadow price of the budget constraint (the value of health outcome) has been established.

6.4 A Four-Stage Sequential Clinical Decision Problem

The dynamic programming approach to optimal patient allocation can also be applied to more complex sequential clinical decision problems. In this section the approach is applied to the four-stage decision problem that was considered in chapter 5.3. This is the problem that was used to generalise the Phelps Mushlin strategy in chapter 3 and is identical to the two-stage problem considered in the previous section except that an alternative treatment t_2 is available. This is illustrated in figure 6.4.1 and is identical to figure 5.3.1 except that this four-stage decision problem is solved in five stages when optimal patient allocation is required.

Figure 6.4.1

At stages 5 and 4 the optimal contingent allocation between t_1 and t_2 is made based on the expected net benefits of sampling for every feasible allocation of each sample entering the stage. At stage 3 contingent allocation decisions must be made between treatment (with either t_1 or t_2) and no treatment. The optimal allocation is based on the expected net benefits of sampling at stage 3 and the payoff from stage 4 given an optimal allocation of the sample allocated to the treatment arm of the trial. At stage 2 an optimal allocation between the test and no test arms of the trial is made for each sample entering stage 2. This contingent allocation is based on the expected net benefits of sampling at stage 2, the payoff from the optimal allocation of the sample which will enter stage 3, and the payoff from the optimal allocation of the sample which will enter stage 5. Finally the optimal sample size can be selected at stage 1 given that an optimal allocation policy will be followed at each subsequent stage.

The value of the first hurdle and the EVPI for this decision problem is identical to chapter 5.3 and if the proposed research passes this first hurdle the second hurdle must be constructed to demonstrate that the research will be cost-effective at the

optimal scale. The value of the second hurdle is dependent on the allocation rules used at each stage because this will determine the expected benefit and cost of a given sample. Once again this example demonstrates that arbitrary rules will underestimate the value of proposed research and will lead to technically inefficient research design. The same numerical example that was used in chapter 5.3 is used to illustrate this approach to optimal patient allocation. The results are compared to the fixed an equal allocation rule used in chapter 5.3 but assuming that the population benefits of the proposed research are endogenous.

6.4.1 Optimal Allocation at Stage 5

At stage 5 the optimal contingent allocation between t_1 and t_2 for each sample entering stage 5 must be made. The expected benefit of sampling not only depends on the sample entering stage 5 but also on the way it is allocated to t_1 and t_2. The $EVSI_{(5)}|n_{(5)},n_{t1}$ is a measure of the expected benefit of sample information at stage 5 given a sample of $n_{(5)}$ entering stage 5 with n_{t1} allocated to t_1, and $n_{(5)}$-n_{t1} allocated to t_2.

$$EVSI_{(5)}|n_{(5)},n_{t1} = K_t \cdot \sqrt{Vn_{(5)},n_{t1}} \cdot \sigma_{0(5)} \cdot L(D_{(5)}|n_{(5)},n_{t1}) \qquad 6.4.1a$$

$$D_{(5)}|n_{(5)},n_{t1} = (\delta_{0(5)}-\delta_b)/\sqrt{Vn_{(5)},n_{t1}}$$

$$\delta_{0(5)} = (E(U|t_e^+, t_1)-g.E(C|t_e^+, t_1)) - (E(U|t_e^+, t_2)-g.E(C|t_e^+, t_2))$$

$$\sqrt{Vn_{(5)},n_{t1}} = \sigma_{0(5)}^2/(\sigma_{0(5)}^2+\sigma_{n(5)}^2)$$

$$\sigma_{n(5)}^2 = (\sigma_{tet1}^2/n_{t1})+(\sigma_{tet2}^2/(n_{(5)}-n_{t1}))$$

The benefit of a sample entering stage 5 is the reduction in the sample variance ($\sigma_{n(5)}^2$) of the incremental net benefits of t_1. The benefit of sampling is determined by both the size of the sample entering stage 5 and the way it is allocated between the alternatives. In this example (when 1/g=£4,000) the variance of the net

benefits of t_1 given a positive test result (σ_{tet1}^2) is greater than the variance of the net benefits of t_2 given a positive test result (σ_{tet2}^2) so the marginal benefits of assigning an entrant to t_1 will be higher than t_2. However the variance is partly determined by the value of g and when 1/g is higher (1/g=£20,00) σ_{tet2}^2 is greater than σ_{tet1}^2 and the marginal benefits of allocating a trial entrant to t_2 will be greater than t_1.

The cost of sampling at stage 5 ($Cs_{(5)}|n_{(5)},n_{t1}$) is also determined by the way the sample is allocated because the marginal cost of assigning a patient to t_1 will be greater than the additional cost of assigning a patient to t_2.

$$Cs_{(5)}|n_{(5)},n_{t1} = (E(C|t_1)-E(C|t_0)).n_{t1} +(E(C|t_2)-E(C|t_0)).(n_{(5)}-n_{t1}) \qquad 6.4.1b$$

The payoff at stage 5 ($\Pi_{(5)}|n_{(5)},n_{t1}$) is simply the expected net benefits of sampling given a sample of $n_{(5)}$ entering stage 5 with n_{t1} allocated to t_1 and $n_{(5)}$-n_{t1} allocated to t_2.

$$\Pi_{(5)}|n_{(5)},n_{t1} = ENBS_{(5)}|n_{(5)},n_{t1} = EVSI_{(5)}|n_{(5)},n_{t1} - Cs_{(5)}|n_{(5)},n_{t1} \qquad 6.4.1c$$

The payoff can be established for every feasible allocation of each sample entering stage 5. This is illustrated in table 6.4.1a where the rows represent the sample entering stage 5, the feasible allocations are represented by the columns, and the optimal contingent allocation to t_1 are the row maximums ($\Pi_{(5)}|n_{(5)},n_{t1}^*$) on the right of the table.

Table 6.4.1a

The optimal contingent allocation of a sample entering stage 5 to t_1 is illustrated in figure 6.4.2a for three values of 1/g. These optimal allocations can be compared with the fixed allocation rule used in chapter 5.3 where half the sample was allocated to t_1 and half to t_2. In this example more trial entrants are allocated to t_2

because the marginal cost of assigning a trial entrant to t_1 is greater than t_2. The proportion of the sample allocated to t_1 falls as the sample size entering stage 5 increases because differences in the marginal benefit of assigning trial entrants to t_1 and t_2 become less significant as the sample size increases, but the difference in the marginal cost of sampling remains constant. When the value of 1/g increases a greater proportion of the sample is allocated to t_1 because less weight is placed on the additional cost of assigning the entrant to t_1.

Figure 6.4.2a

6.4.2 Optimal Allocation at Stage 4

At stage 4 the optimal contingent allocation between t_1 and t_2 must be made based on the expected net benefits of sampling for every feasible allocation of each sample entering stage 4. The $EVSI_{(4)}|n_{(4)},n_{t1}$ is a measure of the expected benefits of a sample of $n_{(4)}$ entering stage 4 with n_{t1} allocated to t_1 and $n_{(4)}$-n_{t1} allocated to t_2.

$$EVSI_{(4)}|n_{(4)},n_{t1} = K_t \cdot \sqrt{Vn_{(4)},n_{t1}} \cdot \sigma_{0(4)} \cdot L(D_{(4)}|n_{(4)},n_{t1}) \qquad 6.4.2a$$

$$D_{(4)}|n_{(4)},n_{t1} = (\delta_{0(4)}-\delta_b)/\sqrt{Vn_{(4)},n_{t1}}$$

$$\delta_{0(4)} = (E(U|t_1)-g.E(C|t_1)) - (E(U|t_2)-g.E(C|t_2))$$

$$\sqrt{Vn_{(4)},n_{t1}} = \sigma_{0(4)}^2/(\sigma_{0(4)}^2+\sigma_{n(4)}^2)$$

$$\sigma_{n(4)}^2 = (\sigma_{t1}^2/n_{t1})+(\sigma_{t2}^2/(n_{(4)}-n_{t1}))$$

The benefit of sample information at stage 4 is the reduction in the sample variance ($\sigma_{n(4)}^2$) of the incremental net benefit of t_1 and the marginal benefit of assigning a trial entrant to t_1 or t_2 is determined by the variance of the net benefit of t_1 (σ_{t1}^2) and the variance of the net benefit of t_2 (σ_{t2}^2). In this example σ_{t2}^2 is greater than σ_{t1}^2 and the difference between σ_{t2}^2 and σ_{t1}^2 increases with the value of 1/g, so the marginal benefits of assigning an entrant to t_2 will be greater than t_1.

The cost of sampling at stage 4 ($Cs_{(4)}|n_{(4)},n_{t1}$) is also determined by the additional treatment cost of assigning an entrant to t_1 which is greater than the additional treatment cost of t_2.

$$Cs_{(4)}|n_{(4)},n_{t1} = (E(C|t_1)-E(C|t_0)).n_{t1} +(E(C|t_2)-E(C|t_0)).(n_{(4)}-n_{t1}) \qquad 6.4.2b$$

The payoff at stage 4 ($\Pi_{(4)}|n_{(4)},n_{t1}$) is simply the expected net benefit of sampling given a sample of $n_{(4)}$ entering stage 4 with n_{t1} allocated to t_1 and $n_{(4)}$-n_{t1} allocated to t_2.

$$\Pi_{(4)}|n_{(4)},n_{t1} = ENBS_{(4)}|n_{(4)},n_{t1} = EVSI_{(4)}|n_{(4)},n_{t1} - Cs_{(4)}|n_{(4)},n_{t1} \qquad 6.4.2c$$

The optimal contingent allocation can now be established by calculating the payoff for every feasible allocation of each sample entering stage 4. This is illustrated in table 6.4.1b where the possible samples entering stage 4 are represented by each row, and the feasible allocations are represented by each column. The optimal allocation for a given sample size entering stage 4 is the row maximum ($\Pi_{(4)}|n_{(4)},n_{t1}^*$) on the right of the table.

Table 6.4.1b

The optimal contingent allocation of each sample size entering stage 4 for three values of 1/g is illustrated in figure 6.4.2b. A smaller proportion of each sample is allocated to t_2 than with the fixed and equal allocation rule because the marginal cost of assigning entrants to t_1 is higher than t_2. This optimal allocation is not the same as at stage 5 because although the marginal cost of sampling is the same, the marginal benefits of assigning entrants to either t_2 or t_1 differ. In general a greater proportion of the sample is allocated to t_2 because the marginal benefits of allocating an entrant to t_2 is higher than at stage 5. The proportion of the sample allocated to t_1 falls as the sample size is increased because the difference in marginal benefit of assignment to t_2 or t_1 declines with sample size. The proportion

of each sample allocated to t_1 increases with 1/g because less weight is placed on the additional cost of assigning an entrant to t_1, and this does not offset the additional benefits of allocating entrants to t_2.

Figure 6.4.2b

6.4.3 Optimal Allocation at Stage 3

At stage 3 the optimal contingent allocation between treatment (t_r) (with either t_1 or t_2 at stage 4) and no treatment (t_0) must be made based on estimates of the expected net benefits of sampling at stage 3 and the payoff from stage 4 given that the sample allocated to the treatment arm will be allocated optimally between t_1 and t_2 at stage 4.

Payoff at Stage 3

A sample entering stage 3 ($n_{(3)}$) which is allocated between the treatment and no treatment arms will generate an expected net benefit of sampling at stage 3, but the sample allocated to the treatment arm will enter stage 4 and will be allocated optimally between t_1 and t_2 generating a payoff of $\Pi_{(4)}|n_{(4)},n_{t1}^*$. So there is a recursive relationship between the payoff at stage 3 and stage 4, where the payoff and the optimal allocation at stage 3 is determined by the ENBS at stage 3 and the payoff given an optimal allocation policy at stage 4.

$$\Pi_{(3)}|n_{(3)},n_{tr} = ENBS_{(3)}|n_{(3)},n_{tr} + \Pi_{(4)}|n_{(4)},n_{t1}^* \qquad 6.4.3a$$

$$n_{(4)} = n_{tr}$$

The payoffs given optimal contingent allocation at stage 4 have already been determined at stage 4 and were illustrated in table 6.4.1b and figure 6.4.2b. But to establish the optimal contingent allocation at stage 3 estimates of the

$ENBS_{(3)}|n_{(3)},n_{te}$ are required for every feasible allocation of each sample entering stage 3.

EVSI at Stage 3

The $EVSI_{(3)}|n_{(3)},n_{tr}$ is a measure of the expected benefits of a sample of $n_{(3)}$ entering stage 3 with n_{tr} allocated to the treatment arm and $n_{(3)}-n_{tr}$ allocated to t_0.

$$EVSI_{(3)}|n_{(3)},n_{tr} = K_t \cdot \sqrt{Vn_{(3)},n_{tr}} \cdot \sigma_{0(3)} \cdot L(D_{(3)}|n_{(3)},n_{tr}) \qquad 6.4.3b$$

$$D_{(3)}|n_{(3)},n_{tr} = (\delta_{0(3)}-\delta_b)/\sqrt{Vn_{(3)},n_{tr}}$$

$$\delta_{0(3)} = (E(U|t_r)-g.E(C|t_r)) - (E(U|t_0)-g.E(C|t_0))$$

$$\sqrt{Vn_{(3)},n_{tr}} = \sigma_{0(3)}^2/(\sigma_{0(3)}^2+\sigma_{n(3)}^2)$$

$$\sigma_{n(3)}^2 = (\sigma_{tr}^2/n_{tr})+(\sigma_{t0}^2/(n_{(3)}-n_{tr}))$$

Following the analysis of stage 2 in chapter 5.3 the prior incremental net benefits of treatment ($\delta_{0(3)}$), the prior variance of $\delta_{0(3)}$ ($\sigma_{0(3)}^2$), and the variance of the net benefit of treatment (σ_{tr}^2) all depend on the sample allocated to the treatment arm.

The prior net benefits of treatment $(E(U|t_r)-g.E(C|t_r))$ and therefore $\delta_{0(3)}$ are dependent on the sample allocated to the treatment arm, because this sample will be allocated optimally between t_1 and t_2 at stage 4, and from 5.3.7:

$$(E(U|t_r)-g.E(C|t_r)) = p(\delta_{x(4)}>\delta_{x(4)}^*).(E(U|t_1)-g.E(C|t_1)) + 1-p(\delta_{x(4)}>\delta_{x(4)}^*).(E(U|t_2)-g.E(C|t_2)) \qquad 6.4.3c$$

$$p(\delta_{x(4)}>\delta_{x(4)}^*) = p(Z>((\delta_{0(4)}-\delta_{x(4)}^*)/\sigma_{n(4)}^2))$$

$$\sigma_{n(4)}^2 = (\sigma_{t1}^2/n_{t1}^*)+(\sigma_{t2}^2/(n_{(4)}-n_{t1}^*))$$

$$n_{(4)} = n_{tr}$$

The value of $p(\delta_{x(4)}>\delta_{x(4)}^*)$ is determined by the both the size of the sample allocated to the treatment arm of the trial (n_{tr}) and the allocation of this sample at

stage 4. The value of $(E(U|t_r)\text{-}g.E(C|t_r))$ is calculated for each sample assigned to the treatment arm given an optimal allocation between t_1 and t_2 at stage 4, and these values are the prior information used to establish the EVSI stage 3.

The prior variance of $E(U|t_r)\text{-}g.E(C|t_r)$ at stage 3 is a combination of the posterior variance of $E(U|t_1)\text{-}g.E(C|t_1)$ and $E(U|t_2)\text{-}g.E(C|t_2)$ from stage 4. The sample allocated to t_r ($n_{tr}=n_{(4)}$) determines the value of $p(\delta_{x(4)}>\delta_{x(4)}^{*})$ and the sample variance of $E(U|t_1)\text{-}g.E(C|t_1)$ and $E(U|t_2)\text{-}g.E(C|t_2)$ (σ_{t1}^2/n_{t1}^{*}, and $\sigma_{t2}^2/(n_{(4)}\text{-}n_{t1}^{*})$ respectively). Since posterior variance is a combination of sample and prior variance, the prior variance of $(E(U|t_r)\text{-}g.E(C|t_r))$ at stage 3 is calculated for every sample assigned to t_r given an optimal allocation policy at stage 4. The population variance of the net benefits of treatment (σ_{tr}^2) is also partly determined by the sample allocated to t_r, because this determines the value of $p(\delta_{x(4)}>\delta_{x(4)}^{*})$, so σ_{tr}^2 is also calculated for every sample assigned to t_r given an optimal allocation at stage 4.

Cost of Sampling at Stage 3

The marginal cost of sampling at stage 3 will be zero because the additional treatment cost of assigning an entrant to t_0 will be zero and the cost of assigning an entrant to the treatment arm will also be zero because the additional treatment cost of t_1 and t_2 has been included in the estimates of the ENBS at stage 4. The expected net benefit of sampling for every feasible allocation of each sample entering stage 3 can now be established:

$$ENBS_{(3)}|n_{(3)},n_{tr} = EVSI_{(3)}|n_{(3)},n_{tr} \qquad 6.4.3d$$

The payoffs at stage 3 for every feasible allocation of each sample entering stage 3 can now be calculated in 6.4.3a and the optimal contingent allocation can be established. This is illustrated in table 6.4.1c where the payoffs at stage 3 in the main body of the table include the ENBS at stage 3 and the payoff given an

optimal allocation of the sample assigned to t_r which enters stage 4. The sample entering stage 3 is represented by each row of the table and the feasible allocations to the treatment arm are represented by each column. The optimal allocations at stage 3 are the row maximums and the optimal allocation (n_{tr}^*) and associated payoff ($\Pi_{(4)}|n_{(4)},n_{t1}^*$) for each sample entering stage 3 are illustrated on the right of the table. This approach of using the contingent allocation established at stage 4 reduces the computation required at stage 3 because only one estimate of $\Pi_{(4)}|n_{(4)},n_{t1}^*$ is required for each sample allocated to the test arm, rather than estimates for every feasible allocation.

Table 6.4.1c

Optimal Allocation at Stage 3

The optimal allocation at stage 3 is illustrated in figure 6.4.2c for three values of 1/g. The optimal allocation depends on the sample entering stage 3 and as $n_{(3)}$ increases the marginal benefits of sampling decline and a greater proportion of the sample is allocated to t_0 where the marginal cost of sampling is zero. The optimal allocation also depends on the value of 1/g, and as 1/g is increased a greater proportion of the sample will be allocated to the treatment arm because the payoff at stage 4 (the net benefits of assigning an entrant to tr) will be greater as less weight is placed in the costs of allocating entrants to t_1 and t_2. Indeed in this example when 1/g=£20,000 and when $n_{(3)}<33$ all the sample is allocated to the treatment arm of the trial because the payoff at stage 4 is greater than the expected net benefits of allocating the entrant to t_0. Over this range of sample sizes t_0 is not a relevant alternative.

Figure 6.4.2c

6.4.4 Optimal Allocation at Stage 2

At stage 2 the optimal contingent allocation between the test and no test arms of the trial must be established for every feasible allocation of each sample entering stage 2. The payoff at stage 2 is based on the ENBS at stage 2, the payoff from the optimal allocation of the sample assigned to the no test arm which enters stage 3 and the payoff from the optimal allocation of the sample assigned to the test arm which enters stage 5.

Payoff at Stage 2

A sample entering stage 2 ($n_{(2)}$) which is allocated to the test (n_{te}) and no test alternatives ($n_{(2)}$-n_{te}) will generate an expected benefit and cost of sampling at stage 2. Also $n_{(2)}$-n_{te}=$n_{(3)}$ will be allocated optimally at stage 3 generating a payoff of $\Pi_{(3)}|n_{(3)},n_{tr}^{*}$, and $p(t_e^{+}).n_{te}$=$n_{(5)}$ will be allocated optimally at stage 5 generating a payoff of $\Pi_{(5)}|n_{(5)},n_{t1}^{*}$. This is a recursive relationship where the payoff and the optimal contingent allocation at stage 2 is determined by the expected net benefits of sampling at stage 2 but also by the payoffs given an optimal allocation policy at subsequent stages.

$$\Pi_{(2)}|n_{(2)},n_{te} = ENBS_{(2)}|n_{(2)},n_{te} + \Pi_{(3)}|n_{(3)},n_{tr}^{*} + \Pi_{(5)}|n_{(5)},n_{t1}^{*} \qquad 6.4.4a$$

$$n_{(3)} = n_{(2)}-n_{te}$$

$$n_{(5)} = p(t_e^{+}).n_{te}$$

The payoffs from the sample assigned to the no test arm, which will be allocated optimally at stage 3, have already been established ($\Pi_{(3)}|n_{(3)},n_{tr}^{*}$) and were illustrated in table 6.4.1c. The payoff from the sample assigned to the test arm, which is allocated optimally at stage 5, have also been established ($\Pi_{(5)}|n_{(5)},n_{t1}^{*}$) and were illustrated in table 6.4.1a. To determine the optimal allocation at stage 2 the $ENBS_{(2)}|n_{(2)},n_{te}$ must be estimated for every feasible allocation of each sample entering stage 2.

EVSI at Stage 2

The $EVSI_{(2)}|n_{(2)},n_{te}$ is a measure of the expected benefit of a sample of $n_{(2)}$ entering stage 2 with n_{te} allocated to the test alternative and $n_{(2)}-n_{te}$ allocated to the no test alternative.

$$EVSI_{(2)}|n_{(2)},n_{te} = K_t \sqrt{Vn_{(2)},n_{te}} .\sigma_{0(2)} .L(D_{(2)}|n_{(2)},n_{te}) \qquad 6.4.4b$$

$$D_{(2)}|n_{(2)},n_{te} = (\delta_{0(2)}-\delta_b)/ \sqrt{Vn_{(2)},n_{te}}$$

$$\delta_{0(2)} = (E(U|t_e)-g.E(C|t_e) - (E(U|nt_e)-g.E(C|nt_e))$$

$$\sqrt{Vn_{(2)},n_{te}} = \sigma_{0(2)}^2/(\sigma_{0(2)}^2+\sigma_{n(2)}^2)$$

$$\sigma_{n(2)}^2 = (\sigma_{te}^2/n_{te})+(\sigma_{nte}^2/(n_{(2)}-n_{te}))$$

Following the analysis of stage 1 in chapter 5.3 the prior incremental net benefit of testing ($\delta_{0(2)}$), the prior variance of $\delta_{0(2)}$ ($\sigma_{0(2)}^2$), and the population variance of the net benefit of test and no test alternatives (σ_{te}^2, and σ_{nte}^2) are all partly determined by the sample allocated to the test and the no test arms of the trial. From 5.3.8a the net benefit of not testing will be the expected net benefit from stage 3 given the optimal allocation of a sample of $n_{(2)}-n_{te}=n_{(3)}$ at stage 3.

$$E(U|nt_e)-g.E(C|nt_e) = p(\delta_{x(3)}>\delta_{x(3)}^*).(E(U|t_r)-g.E(C|t_r)) \qquad 6.4.4c$$
$$+1-p(\delta_{x(2)}>\delta_{x(2)}^*).(E(U|t_0)-g.E(C|t_0))$$

$$p(\delta_{x(3)}>\delta_{x(3)}^*) = p(Z>((\delta_{0(3)}-\delta_{x(3)}^*)/\sigma_{n(3)}^2))$$

$$\sigma_{n(3)}^2 = (\sigma_{tr}^2/n_{tr}^*)+(\sigma_{t0}^2/(n_{(3)}-n_{t0}^*))$$

$$n_{(3)} = n_{(2)}-n_{te}$$

The value of $p(\delta_{x(3)}>\delta_{x(3)}^*)$ and therefore $E(U|nt_e)-g.E(C|nt_e)$ are determined by the sample allocated to the no test arm. The value of $E(U|nt_e)-g.E(C|nt_e)$ is calculated for each sample allocated to the no test arm given that this sample will be allocated optimally at stage 3 (and subsequently at stage 4), and these values are used as the prior information used to establish the EVSI at stage 2. The prior

variance of $E(U|nt_e)$-g.$E(C|nt_e)$ at stage 2 is a combination of the prior variance of $(E(U|t_0)$-g.$E(C|t_0))$ and the posterior variance of $(E(U|t_r)$-g.$E(C|t_r))$ at stage 3. The prior variance of $E(U|nt_e)$-g.$E(C|nt_e)$ at stage 2 is calculated for each sample allocated to the no test arm given that it will be allocated optimally at stage 3. The population variance of $E(U|nt_e)$-g.$E(C|nt_e)$ is partly determined by $p(\delta_{x(3)}>\delta_{x(3)}^{*})$ and is also calculated for each sample assigned to the no test arm given an optimal allocation policy at stage 3 (and subsequently at stage 4).

The prior net benefit of treatment following a positive test result and therefore the net benefit of testing at stage 2 $(E(U|t_e)$-g.$E(C|t_e))$ are determined by the sample allocated to the test arm, and from 5.3.8b:

$$\begin{aligned} E(U|t_e)\text{-g.}E(C|t_e) \;=\; & p(t_e^{-}).(E(U|t_e^{-}, t_0)\text{-g.}E(C|t_e^{-}, t_0)) \qquad 6.4.4d \\ & + p(t_e^{+}).(p(\delta_{x(5)}>\delta_{x^{*}(5)}).(E(U|t_e^{+}, t_1)\text{-g.}E(C|t_e^{+}, t_1)) \\ & + 1\text{-}p(\delta_{x(5)}>\delta_{x^{*}(5)}).(E(U|t_e^{+}, t_2)\text{-g.}E(U|t_e^{+}, t_2))) \end{aligned}$$

$$p(\delta_{x(5)}>\delta_{x(5)}^{*}) = p(Z>((\delta_{0(5)}-\delta_{x(5)}^{*})/\sigma_{n(5)}^{2}))$$

$$\sigma_{n(5)}^{2} = (\sigma_{tet1}^{2}/n_{t1}^{*})+(\sigma_{tet2}^{2}/(n_{(5)}-n_{t1}^{*}))$$

$$n_{(5)} = p(t_e^{+}).n_{te}$$

The value of $p(\delta_{x(5)}>\delta_{x(5)}^{*})$ is determined by the sample allocated to the test arm of the trial which enters stage 5 and $E(U|t_e)$-g.$E(C|t_e)$ is calculated for each sample allocated to the test arm at stage 2 given that the sample entering stage 5 will be allocated optimally. These values are the prior information used to establish the EVSI at stage 2.

The prior variance of $E(U|t_e)$-g.$E(C|t_e)$ at stage 2 is a combination of the prior variance of $(E(U|t_e^{-}, t_0)$-g.$E(C|t_e^{-}, t_0)$, and the posterior variance of $(E(U|t_e^{+}, t_1)$-g.$E(C|t_e^{+}, t_1))$, and $(E(U|t_e^{+}, t_2)$-g.$E(U|t_e^{+}, t_2)))$ given the optimal allocation at stage 5 of each sample allocated to the test alternative at stage 2. The population variance of $E(U|t_e)$-g.$E(C|t_e)$ can also be calculated for each sample allocated to

the test arm which enters stage 5 because this determines the value of $p(\delta_{x(5)}>\delta_{x(5)}^{*})$. The expected benefit for every feasible allocation of each sample entering stage 2 can now be established given that the sample allocated to the no test arm will be allocated optimally at stage 3 (and subsequently stage 4) and given that the sample allocated to the test arm of the trial will be allocated optimally at stage 5.

Cost of Sampling at Stage 2

The cost of sampling for each sample entering stage 2 ($Cs_{(2)}|n_{(2)},n_{te}$) is the additional treatment cost of the test and no test alternatives. The marginal cost of assigning an entrant to the no test alternative will be zero because the additional treatment costs at stage 3 and stage 4 are included in the payoff from stage 3. The additional cost of assigning a trial entrant to the test arm will simply be the cost of the diagnostic test (C_{te}) because the additional cost of t_0 given a negative test result will be zero and the additional treatment cost given a positive test result is included in the payoff at stage 5.

$$Cs_{(2)}|n_{(2)},n_{te} = C_{te}.n_{te} \qquad 6.4.4e$$

The expected net benefit of sampling at stage 2 ($ENBS_{(2)}|n_{(2)},n_{te}$) can now be established for each sample of $n_{(2)}$ entering stage 2, with n_{te} allocated to the test arm, and $n_{(2)}$-n_{te} allocated to no test arm of the trial.

$$ENBS_{(2)}|n_{(2)},n_{te} = EVSI_{(2)}|n_{(2)},n_{te} - Cs_{(2)}|n_{(2)},n_{te} \qquad 6.4.4f$$

The optimal contingent allocation of each sample entering stage 2 can be determined based on the payoff at stage 2 ($\Pi_{(2)}|n_{(2)},n_{te}$) for every feasible allocation of each sample entering stage 2. This is illustrated in table 6.4.1d where the sample entering stage 2 is represented by the rows of the table and the feasible allocations by the columns. The optimal allocation of each sample is the row

maximum ($\Pi_{(2)}|n_{(2)},n_{te}^{*}$) and the optimal allocation (n_{te}^{*}) and the associated payoffs are illustrated in the right hand columns of the table.

Table 6.4.1d

Optimal Allocation at Stage 2

The optimal contingent allocation of a sample entering stage 2 to the test arm of the trial for three values of 1/g is illustrated in figure 6.4.2d. The optimal allocation depends on both the size of the sample entering stage 2 and the value of 1/g. In general a greater proportion of the sample is allocated to the more costly test arm of the trial as the value of 1/g is increased and less weight is placed on the cost of sampling. It is worth noting that when 1/g is low (£4,000) and $n_{(2)}< 50$ then no sample is allocated to the test arm of the trial and testing is not a relevant alternative, however when 1/g=£20,000 and $n_{(2)} <50$ then almost all the sample is allocated to the test arm of the trial. The proportion of the sample allocated to the test arm of the trial also tends to decline as the sample size entering stage 2 increases because any difference in the marginal benefit of allocation to the test and no test alternatives becomes less significant. The relationship between the optimal allocation, the value of 1/g, and $n_{(2)}$ is more complex at stage 2 because the optimal allocation is not simply determined by the ENBS at stage 2 but also by the payoffs and the optimal allocation at subsequent stages.

Figure 6.4.2d

6.4.5 Optimal Sample Size at Stage 1

The first stage is simply to select the optimal sample size given that an optimal allocation policy will be followed at each subsequent stage. The payoff given a sample of $n_{(1)}$ selected at stage 1 ($\Pi_{(1)}|n_{(1)}$) is simply the payoff from stage 2 given

an optimal allocation of $n_{(2)}$ between the test arm (n_{te}^*) and the no test arm $n_{(2)}-n_{te}^*$.

$$\Pi_{(1)}|n_{(1)} = \Pi_{(2)}|n_{(2)},n_{te}^* \qquad 6.4.5$$

$$ENBS_{(1)}|n_{(1)} = ENBS_{(2)}|n_{(2)},n_{te}^* + ENBS_{(3)}|n_{(3)},n_{tr}^* + ENBS_{(4)}|n_{(4)},n_{t1}^* + ENBS_{(5)}|n_{(5)},n_{t1}^*$$

$$n_{(2)} = n_{(1)}$$

$$n_{(3)} = n_{(2)}-n_{te}^*$$

$$n_{(4)} = n_{tr}^*$$

$$n_{(5)} = p(t_e^+).n_{te}^*$$

The optimal sample size at stage 1 ($n_{(1)}^*$) can be selected in table 6.4.1e and is where $\Pi_{(1)}|n_{(1)}$ or the $ENBS_{(1)}|n_{(1)}$ reaches a maximum. The $ENBS_{(1)}|n_{(1)}^*$ provides the value of the second hurdle for this 4 stage decision problem. Once the optimal sample size at stage 1 has been selected the allocation of this sample at each subsequent stage is given by the contingent allocative decisions which have already been made and which were illustrated in tables 6.4.1a, 6.4.1b, 6.4.1c, 6.4.1d, and figures 6.4.2a, 6.4.2b, 6.4.2c, 6.4.2d.

Table 6.4.1e

The Dynamic Programming Approach

The approach taken to this four-stage decision problem is a simple five-stage dynamic programme where contingent allocative decisions at each stage are solved before the optimal sample at stage 1 can be selected. The payoff for each sample considered at stage 1 is established given that an optimal allocation policy will be followed at each subsequent stage. This approach reduces the computation required to solve this problem even more dramatically than in the two-stage decision problem considered in section 6.3. Only one estimate of $ENBS_{(3)}|n_{(3)},n_{tr}^*$ and $ENBS_{(4)}|n_{(4)},n_{t1}^*$ is required for each sample allocated to the no

test arm at stage 2 and only one estimate of $ENBS_{(5)}|n_{(5)},n_{t1}^*$ is required for each sample allocated to the test arm, rather than estimates of the ENBS for every feasible allocation of each sample allocated to the test and no test arms. This reduces the number of estimates of the ENBS from $125,751^4$ if full enumeration of this four stage decision problem is required to 125,751*4 where a maximum sample of 500 is considered at each stage. The solution for this numerical example using this approach requires 503,004 estimates of the ENBS which takes approximately 24 hours of computing time (6 estimates per second). This compares very favourably to the full enumeration of all possible alternatives which (even with a 100 fold increase in computing speed) would require over 13 billion years of computing time, a task so enormous it can safely be regarded as impossible.

6.4.6 Expected Net Benefits of Sample Information

The expected net benefit of sampling given optimal patient allocation at each stage is illustrated in figure 6.4.3a when 1/g=£4,000. This can be compared to the expected net benefit when the fixed and equal allocation rule is used at each stage in figure 6.4.3b. The $ENBS_{(1)}$ with optimal patient allocation reaches a maximum of £3,865,420 at an optimal sample size of 165, with an allocation of n_{te}^*=67 at stage 2, n_{tr}^*=27 at stage 3, n_{t1}^*=12 at stage 4, and n_{t1}^*=19 at stage 5. This is greater than with fixed allocation where the $ENBS_{(1)}$ reaches a maximum of £3,529,502 at an optimal sample size of 148. In this example the optimal patient allocation increases the expected benefits and reduces the costs of sampling. An arbitrary allocation rule will lead an underestimate of the value and optimal scale of the proposed research.

Figure 6.4.3a

Figure 6.4.3b

Value of Health Outcome

The expected net benefit of sampling when the value of 1/g is increased to £10,000 is illustrated in figure 6.4.4a when an optimal allocation policy is followed at each stage. The expected net benefit of sampling reaches a maximum of £20,207,332 at an optimal sample size of 199. This sample is allocated optimally at each stage with n_{te}^{*}=96 at stage 2, n_{tr}^{*}=53 at stage 3, n_{t1}^{*}=24 at stage 4, and n_{t1}^{*}=27 at stage 5. A greater proportion of the sample is allocated to the more costly alternatives at each stage because less weight is placed on the additional sampling cost when the value of 1/g is increased. The expected net benefit when the fixed allocation rule is used is lower and reaches a maximum of £19,988,081 at an optimal sample size of 192 in figure 6.4.4b.

Figure 6.4.4a

Figure 6.4.4b

The expected net benefit of sampling when the value of 1/g is increased to £20,000 illustrated in figure 6.4.5a when the optimal allocation is used and in figure 6.4.5b when the fixed allocation is used. The ENBS reaches a maximum of £28,865,947 at an optimal sample size of 257 when an optimal allocation policy is followed. This sample is allocated optimally with n_{te}^{*}=116 at stage 2, n_{tr}^{*}=75 at stage 3, n_{t1}^{*}=34 at stage 4, and n_{t1}^{*}=33 at stage 5. The ENBS with optimal allocation is greater than with the fixed allocation rule where the ENBS reaches a maximum of £28,543,130 at an optimal sample size of 244 in figure 6.4.4b.

Figure 6.4.5a

Figure 6.4.5b

Optimal Patient Allocation

These examples demonstrate that the optimal patient allocation increases the ENBS and the value of the second hurdle, and in this numerical example optimal patient allocation also increases the optimal sample size. It also demonstrates that the difference in the ENBS, optimal sample size and the allocation at subsequent stages is determined by the value placed on health outcome. This is illustrated in figure 6.4.6 where the optimal sample size and the allocation to the test arm of the trial for optimal and fixed allocation rules can be compared. In this example optimal sample size with optimal patient allocation is greater than with fixed allocation and a smaller proportion of the sample is allocated to the test arm which has high sampling cost. Indeed when 1/g=£3,000 no sample is assigned to the test arm of the trial and the testing alternative can be excluded as an irrelevant alternative. The difference in the ENBS, the EVSI and the cost of sampling between optimal and fixed allocation is illustrated in figure 6.4.7. This demonstrates that using arbitrary allocation rules will underestimate the ENBS. In this example optimal patient allocation increases the expected benefit of sampling and it also reduces the cost of sampling despite larger sample size.

Figure 6.4.6

Figure 6.4.7

The optimal allocation of trial entrants in this example assigns a sample to every arm of the trial when 1/g≥£4,000. This means that it is efficient to compare each alternative at each stage using sample information and all the alternatives can be regarded as relevant. However when 1/g=£3,000 $n_{te}^{*}=0$ at stage 2 and $n_{t1}=0$ at stage 4, and the sample is allocated to t_0 and t_2 which in this case are the only relevant alternatives. The new treatment t_2 is a relevant alternative and this is also confirmed by comparing the ENBS and optimal sample size for the two stage and the four stage decision problem in figure 6.4.8a and 6.4.8b. The difference between the two and five-stage problem is that treatment t_2 was not available in

the two-stage problem. Since the ENBS and the optimal sample size is greater when t_2 is included in the decision problem it is clear that t_2 is a relevant alternative. An analysis which simplified what is a four-stage problem to a two-stage problem by excluding t_2 would underestimate the ENBS and the optimal scale of the research at the second hurdle.

Figure 6.4.8a

Figure 6.4.8b

The analysis of the two-stage problem in the previous section of this chapter found that testing was not a relevant alternative when 1/g=£4,000 and $n_{te}^*=0$. However when t_2 is included testing is a relevant alternative with $n_{te}^*=67$ when 1/g=£4,000. So excluding t_2 would also exclude the testing arm of the trial. This demonstrates a theme which has been discussed in previous chapters: that the selection of relevant alternatives and the exclusion of some alternatives based on judgements and implicit decision rules can seriously bias the analysis, lead to errors at the second hurdle, and inefficient research design. If this inconsistency arose (due to a higher implicit value of health outcome) the ENBS would be seriously underestimated (£3,865,420 compared to £724,970) and there would be a very real danger that this research could be rejected at the second hurdle when it would be cost-effective if t_2 was included. The design of the research would also be inefficient because testing would be excluded as an irrelevant alternative and rather than a trial that assigned patients to each alternative at each stage, it would simply compare t_1 and t_2 at a single-stage. If the value of 1/g is lower (£3,000) then the ENBS and the sample size is zero when t_2 is wrongly excluded and no research will be undertaken.

6.5 Conclusions

In this chapter the fixed and equal allocation rule which was used in chapter 4 and 5 is relaxed. The numerical examples in this chapter demonstrate that the optimal allocation of trial entrants is not simply an issue for sequential clinical trials but is a more fundamental problem which is also relevant to the fixed sample designs considered here.

The approach to optimal patient allocation taken in this chapter explicitly considers the marginal cost and the marginal benefit of assigning trial entrants to the alternatives arms of the trial at each stage. In section 6.2 the single stage decision problem was solved in two stages and required the full enumeration of all feasible allocations of each sample considered. However the full enumeration of all feasible allocations in more complex sequential decision problems is not tractable and in section 6.3 and 6.4 the two and four-stage decision problems were solved using three and five-stage dynamic programmes. This simple dynamic programming approach utilises the recursive relationship between the payoffs at each stage of the decision problem and provides a practical solution to the problem. This may not be the optimal solution (because although there is a recursive relationship between the payoffs at each stage the prior variance and incremental net benefit at earlier stages are partly determined by the contingent allocation at later stages), but the dynamic programming approach is the only feasible solution and clearly provides a better solution than the arbitrary fixed allocation rule used in chapter 5.

The optimal allocation of trial entrants will increase the expected net benefit of sampling and the value of the second hurdle. This may be achieved by a reduction in the cost of sampling (possibly with a reduction in the expected benefits of sampling) or alternatively it may be achieved by an increase in the expected benefits of sampling (possibly with an increase in the cost of sampling). The optimal allocation at each stage, optimal sample size, and the expected net benefits

of sampling are determined by the value of health outcome because this determines the relative weight placed on the additional benefits and costs of assigning entrants to each arm of the trial. These numerical examples demonstrate that arbitrary and fixed allocation rules are inefficient and will lead to an underestimate of the value of proposed research and there is a danger that research which should be accepted at the second hurdle will be rejected. Research which is accepted at the second hurdle despite an arbitrary allocation will be designed inefficiently, may include the comparison of irrelevant alternatives, and the value of the research will be underestimated.

The simple dynamic programming approach to optimal patient allocation provides and explicit and consistent method to identify relevant alternatives which should be compared in the trial. This is one of the problems which was posed by the analysis of the Phelps Mushlin strategy in chapter 3 and could not be adressed by using the fixed allocation rule in chapter 5. Optimal allocation provides a method to rule out irrelevant alternatives consistently based on an assessment of the expected benefit and cost of comparing alternatives, because it allows no sample to be allocated to an alterative at each stage. If it is optimal not to allocate a sample to an alternative then it can be regarded as irrelevant and can be excluded from prospective research.

These numerical examples demonstrate that what are relevant alternatives depends on the value of 1/g and alternatives cannot be ruled out as irrelevant before the shadow price of the budget constraint (the value of health outcome) has been established. They also illustrate the danger of ruling-out alternatives based on implicit and inconsistent decision rules. An analysis which simplified the two-stage decision problem in section 6.3 to the single-stage treatment decision in section 6.2 would underestimated the value of the research and seriously bias the trial design. Similarly an analysis which simplified the four-stage problem in section 6.3 to the two-stage problem in section 6.3 by excluding t_2 would underestimate the ENBS and the optimal scale of the research at the second hurdle. This

demonstrates an argument which was also made in previous chapters: that the selection of relevant alternatives and the exclusion of some alternatives based on judgements and implicit decision rules can seriously bias the analysis, lead to errors at the second hurdle, and lead to inefficient research design.

Appendix E Tables and Figures for Chapter 6

Contents:

Figure 6.2.1 Decision Tree for the Single-Stage Decision Problem

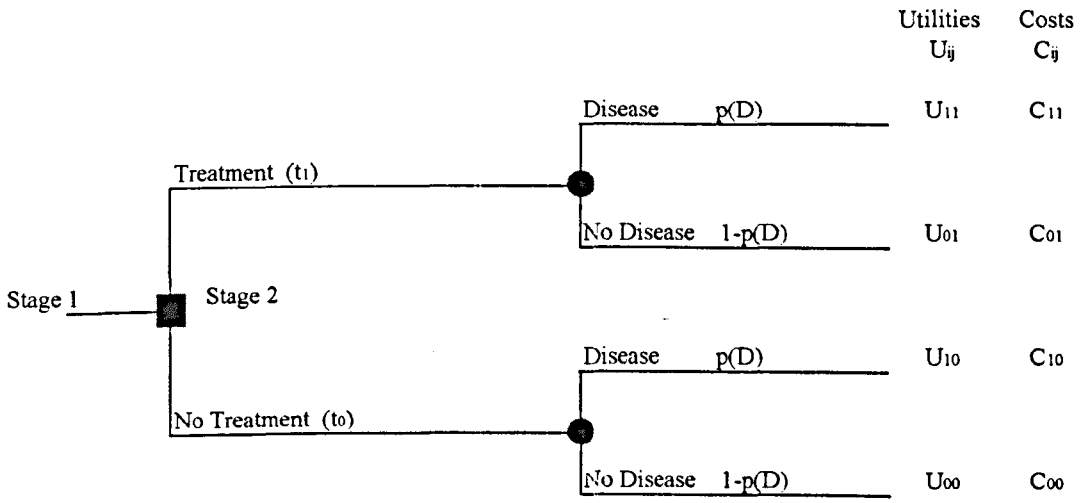


Table 6.2.1a Optimal Allocation of the Sample Entering Stage 2 (1/g=£4,000)

n_{t1} / $n_{(2)}$		Payoff at stage 2 $\Pi_{(2)}\|n_{(2)},n_{t1} = ENBS_{(2)}\|n_{(2)},n_{t1}$ Allocation to t_1 (n_{t1}) (with n_{t1} allocated to t_1, and $n_{(2)}$-n_{t1} allocated to t_0)														Maximum payoff and optimal allocation (n_{t1}^*)	
		0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(2)}\|n_{(2)},n_{t1}^*$	n_{t1}^*
Sample Entering Stage 2 $n_{(2)}=n_{(1)}$	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	1	0	-£10,400	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	2	0	-£9,952	-£20,800	-	-	-	-	-	-	-	-	-	-	-	0	0
	3	0	-£8,395	-£18,405	-£31,200	-	-	-	-	-	-	-	-	-	-	0	0
	4	0	-£7,031	-£8,728	-£26,905	-£41,600	-	-	-	-	-	-	-	-	-	0	0
	5	0	-£6,015	£712	-£8,166	-£35,812	-£52,000	-	-	-	-	-	-	-	-	£712	2
	6	0	-£5,257	£8,216	£11,289	-£9,355	-£45,065	-£62,400	-	-	-	-	-	-	-	£11,289	3
	7	0	-£4,677	£14,053	£27,397	£19,266	-£12,344	-£54,572	-£72,800	-	-	-	-	-	-	£27,397	3
	8	0	-£4,224	£18,645	£40,301	£43,646	£24,092	-£16,786	-£64,259	-£83,200	-	-	-	-	-	£43,646	4
	9	0	-£3,860	£22,314	£50,664	£63,540	£55,766	£26,223	-£22,300	-74,084	-£93,600	-	-	-	-	£63,540	4
	10	0	-£3,562	£25,304	£59,072	£79,752	£81,979	£64,182	£26,199	-£28,711	-£84,010	-£104,000	-	-	-	£81,979	5
	11	0	-£3,314..	£27,795..	£66,028..	£93,093..	£103,658..	£96,005..	£69,558..	£24,479..	-35,753£	-£94,013..	-£114,400..	-	-	£103,658	5
	12	0	-£3,108	£29,906	£71,840	£104,237	£121,602	£122,536	£106,255	£72,503	£21,399	-£43,304	-£104,078	-£124,800	-	£122,536	6
	..	0	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

Figure 6.2.2 Optimal Allocation at Stage 2

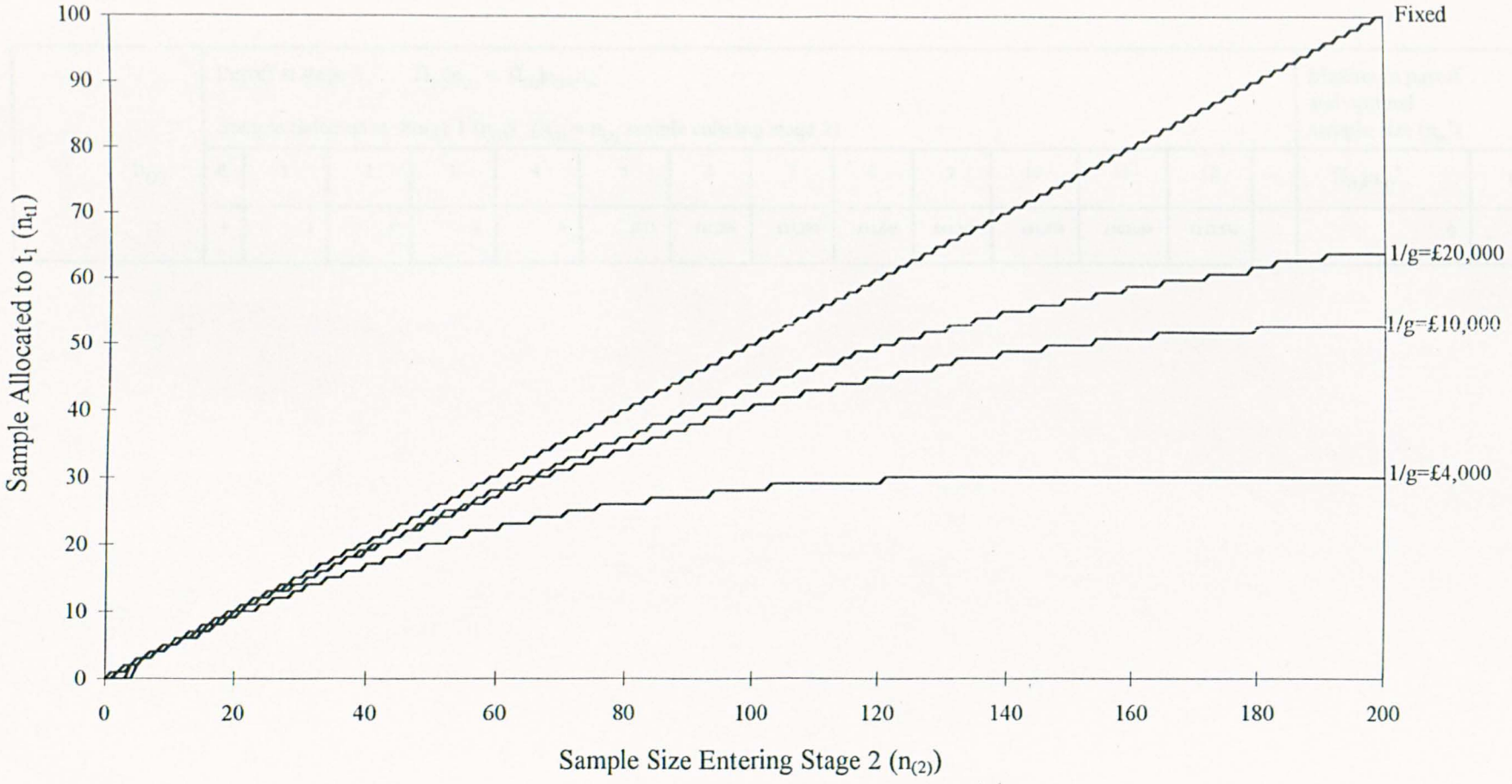


Table 6.2.1b Optimal Sample Size at Stage 1 (1/g=£4,000)

$n_{(1)}$	Payoff at stage 1 $\Pi_{(1)}\|n_{(1)} = \Pi_{(2)}\|n_{(2)},n_{te}^*$ Sample Selected at Stage 1 ($n_{(1)}$) ($n_{(1)} = n_{(2)}$ sample entering stage 2)														Maximum payoff and optimal sample size (n_{te}^*)	
	0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(1)}\|n_{(1)}^*$	$n_{(1)}$
	0	0	0	0	0	£712	£11,289	£27,397	£43,646	£63,540	£81,979	£103,658	£122,536	..	0	0

Figure 6.2.3a Optimal Allocation: ENBS, EVSI, and Cs (1/g=£4,000)

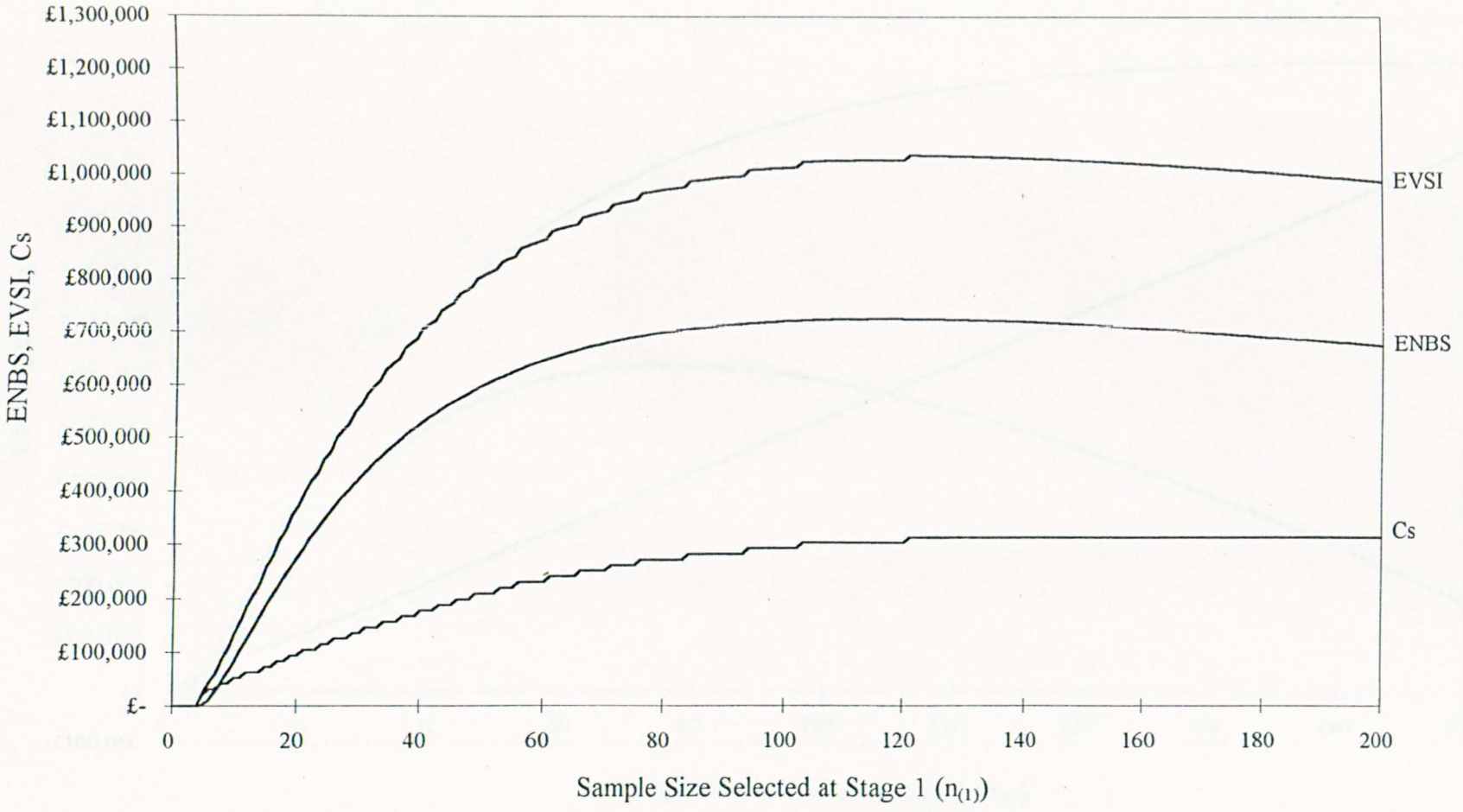


Figure 6.2.3b Fixed Allocation: ENBS, EVSI, and Cs (1/g=£4,000)

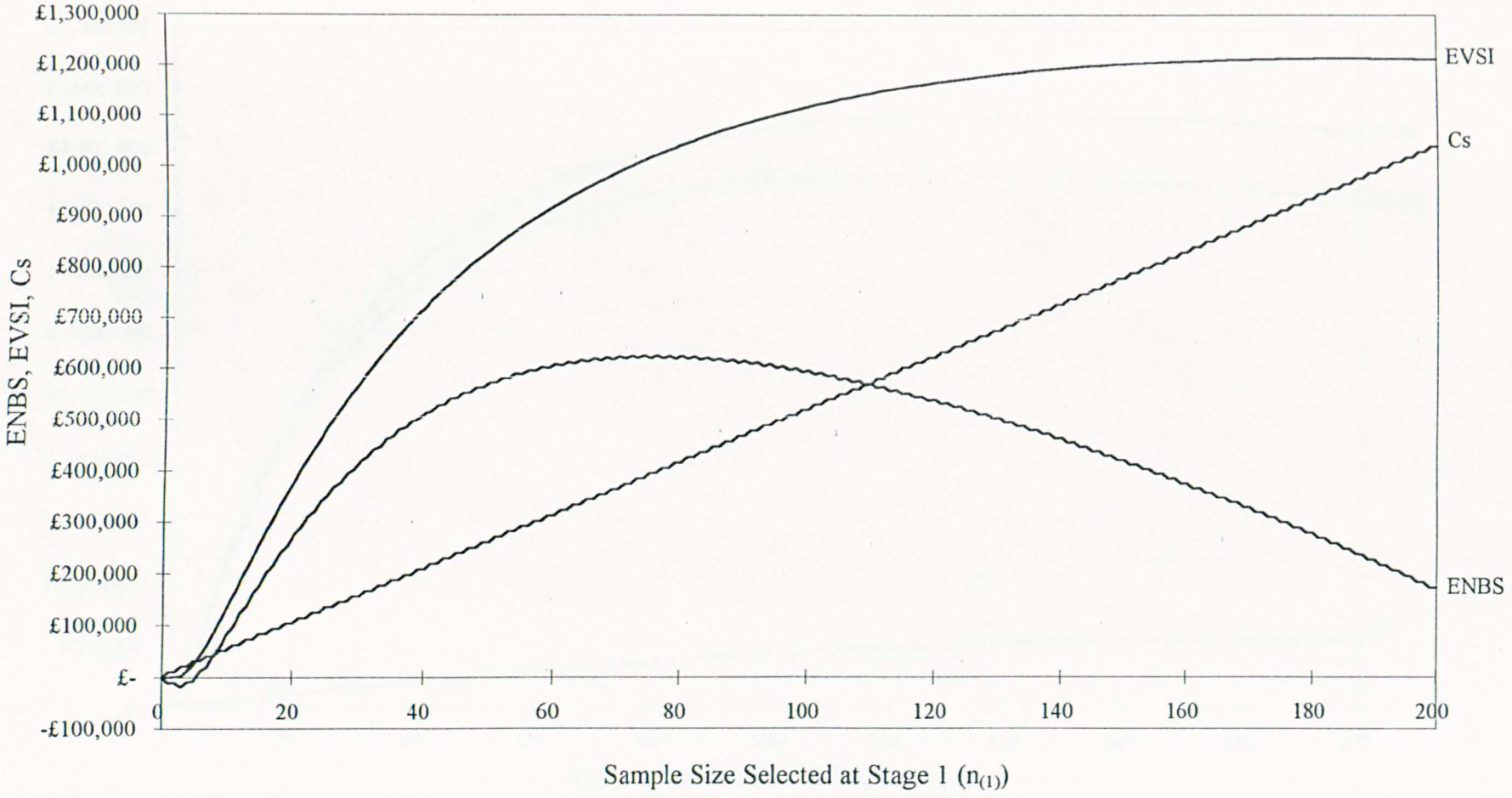


Figure 6.2.4a Optimal Allocation: ENBS, EVSI, and Cs (1/g=£10,000)

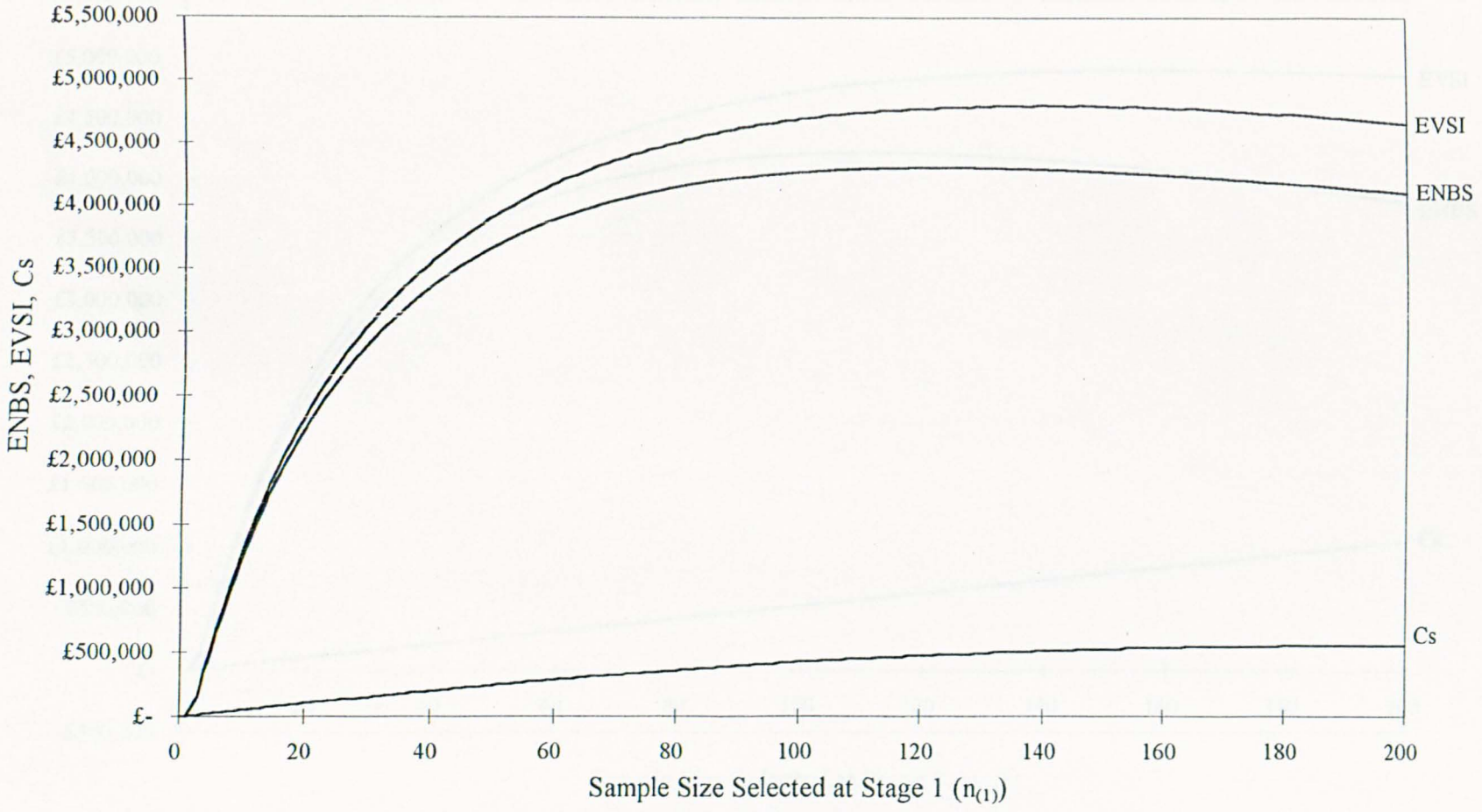


Figure 6.2.4b Fixed Allocation: ENBS, EVSI, and Cs (1/g=£10,000)

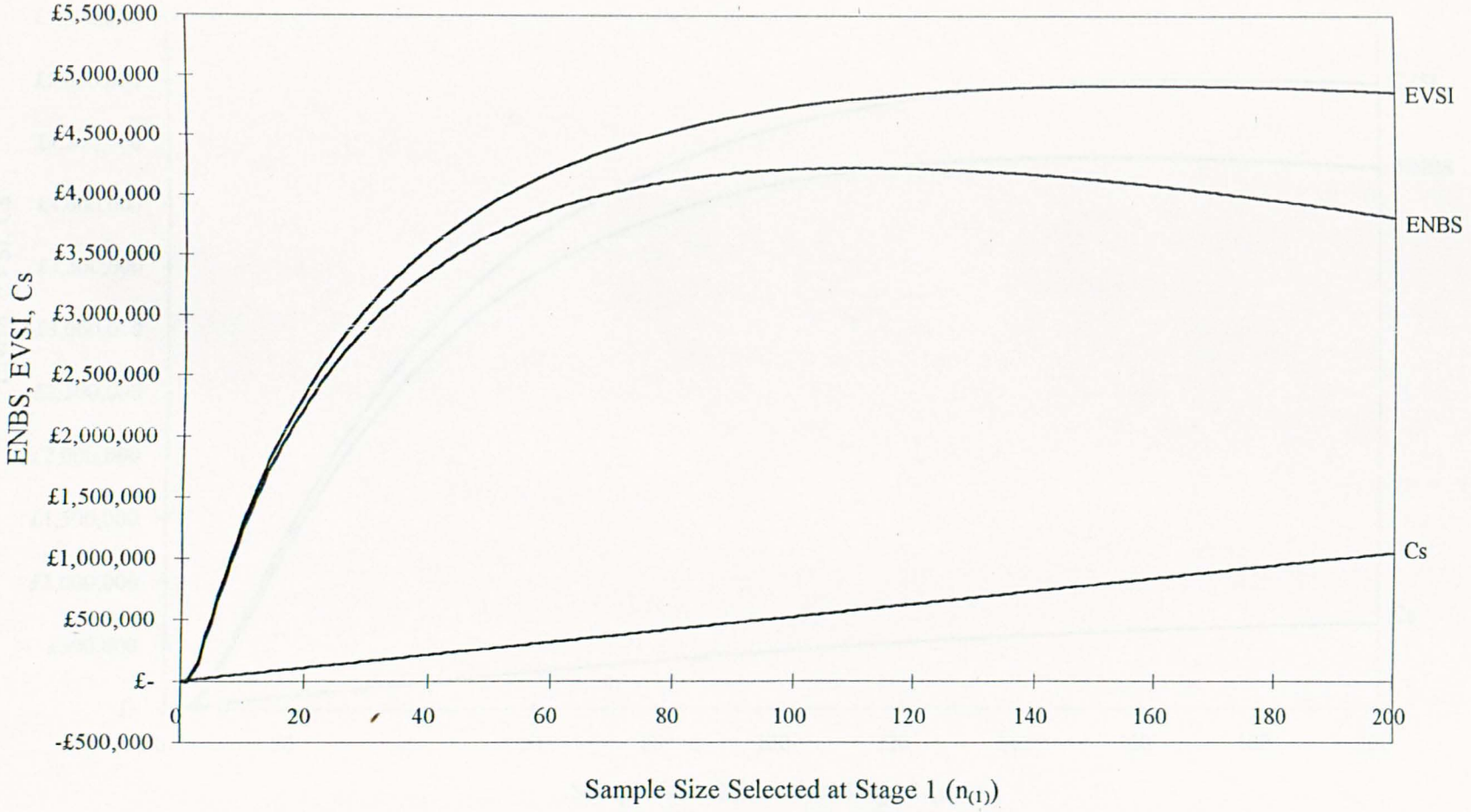


Figure 6.2.5a Optimal Allocation: ENBS, EVSI, and Cs (1/g=£20,000)

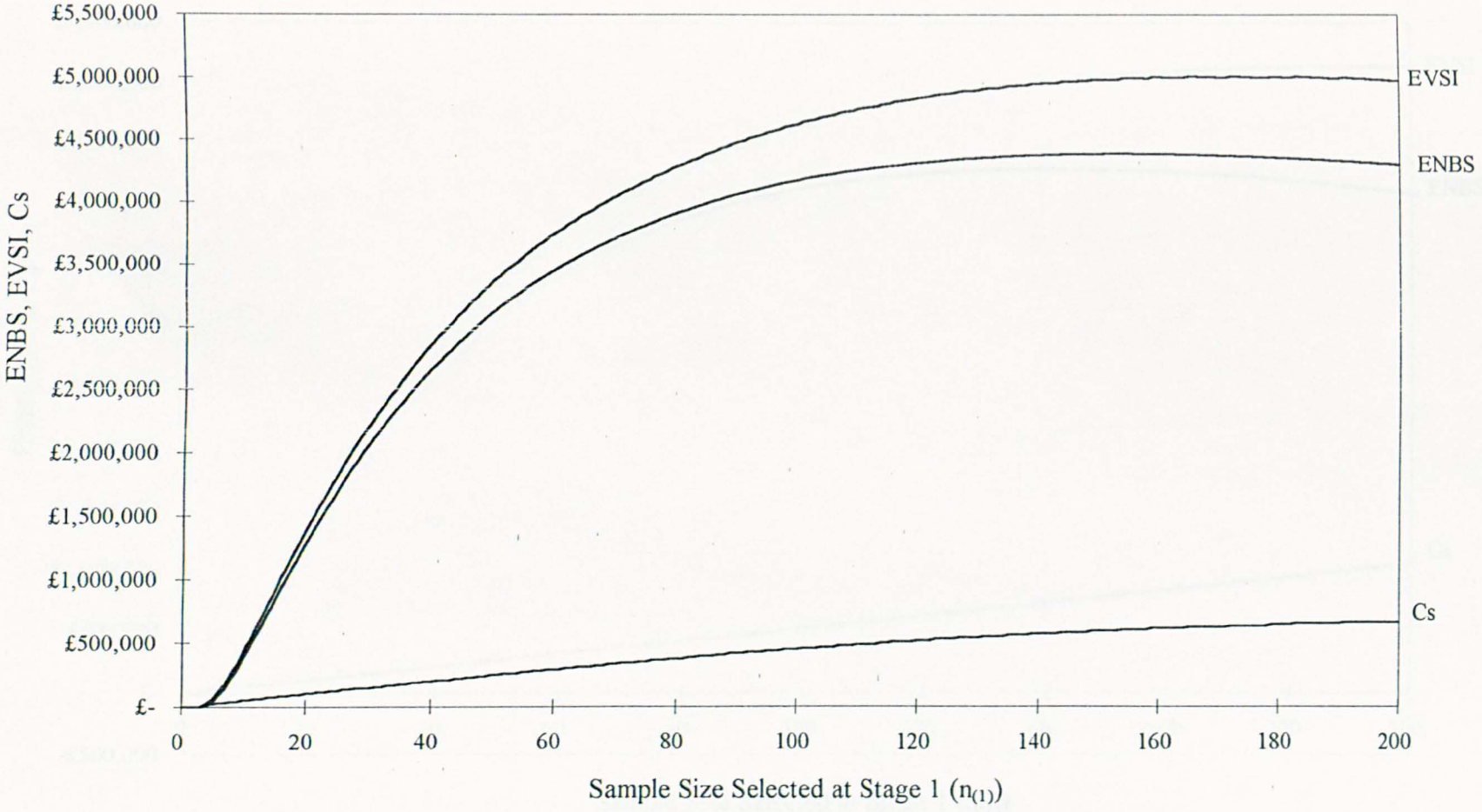


Figure 6.2.5b Fixed Allocation: ENBS, EVSI, and Cs (1/g=£20,000)

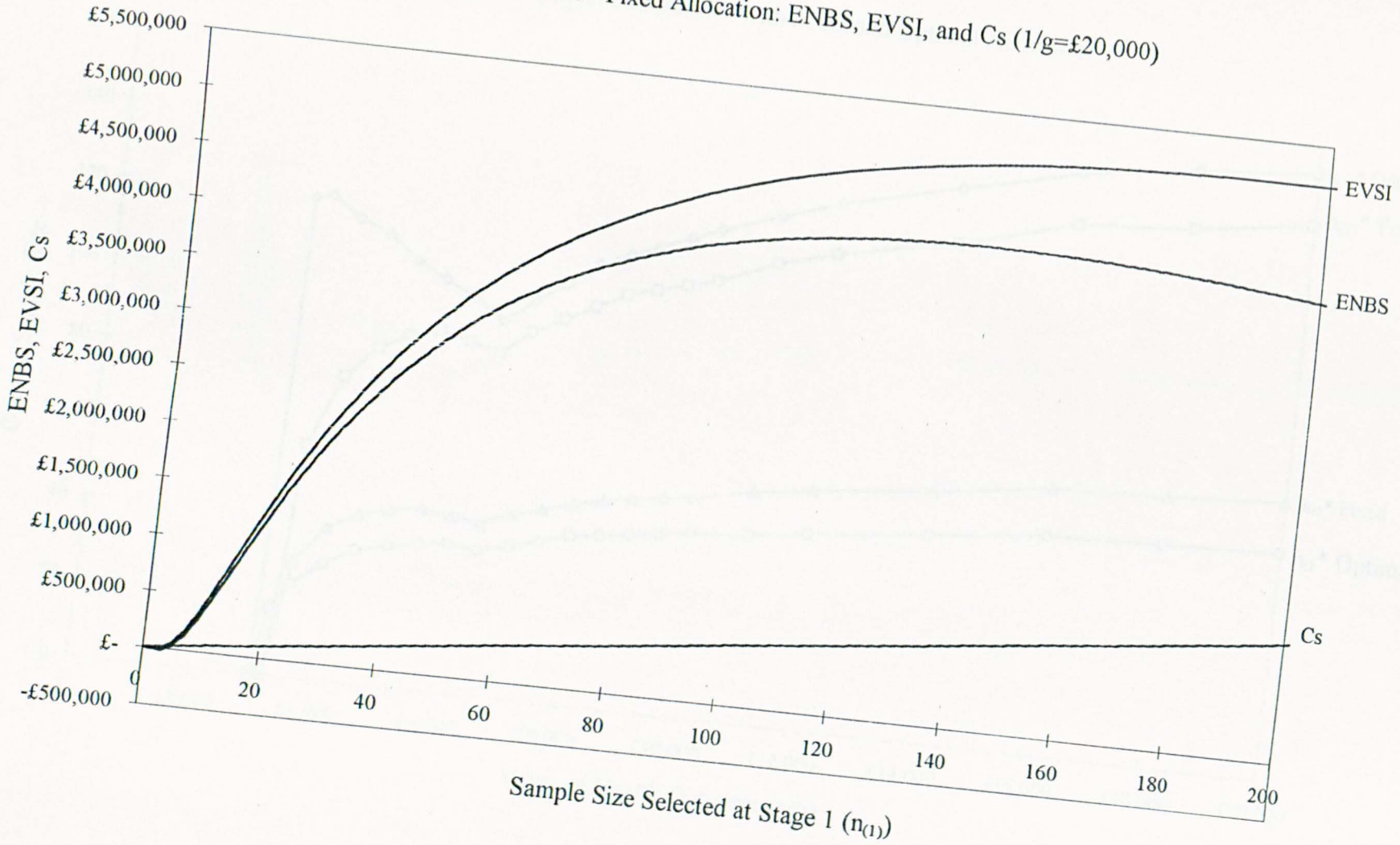


Figure 6.2.6 Optimal Sample Size with Fixed and Optimal Allocation

Figure 6.2.7 Difference Between Optimal and Fixed Allocations

Figure 6.3.1 Decision Tree for the Two Stage Decision Problem

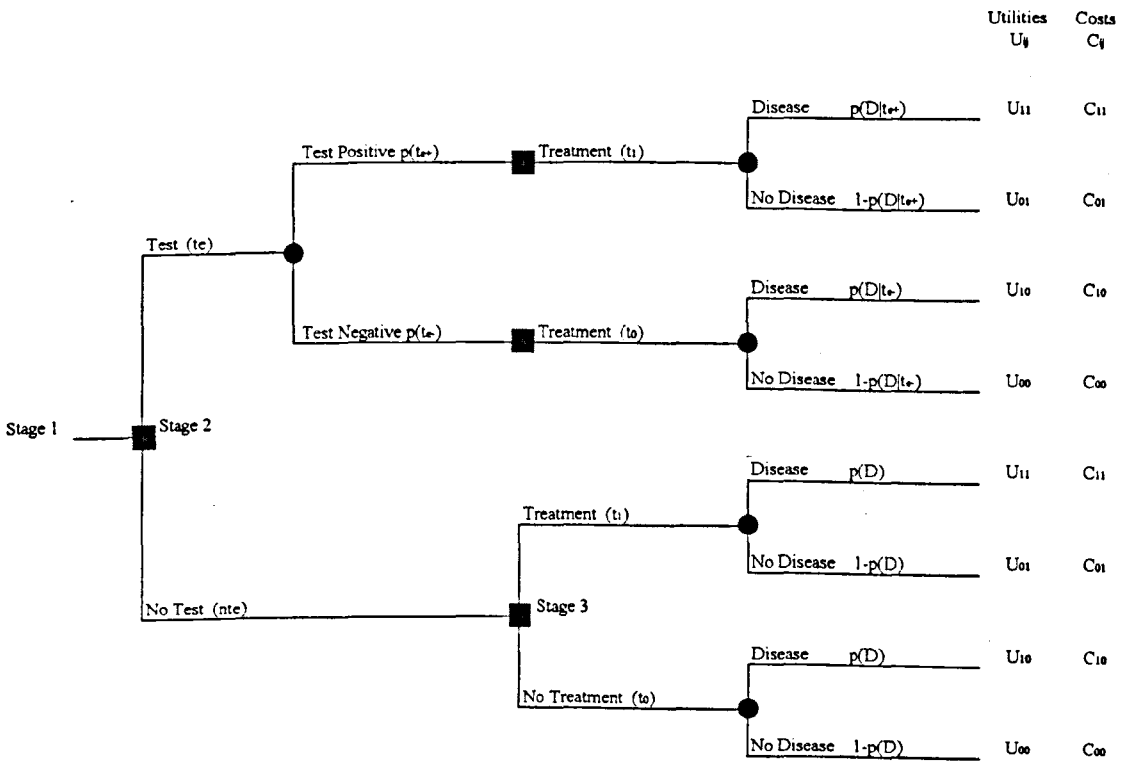


Table 6.3.1a Optimal Allocation of the Sample Entering Stage 3 (1/g=£4,000)

$n_{(3)}$	n_{t1}	Payoff at stage 3 $\Pi_{(3)}\|n_{(3)},n_{t1} = ENBS_{(3)}\|n_{(3)},nt_1$ Allocation to t_1 (n_{t1}) (with n_{t1} allocated to t_1, and $n_{(3)}-n_{t1}$ allocated to t_0)														Maximum payoff and optimal allocation (n_{t1}^*)	
		0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(3)}\|n_{(3)},n_{t1}^*$	n_{t1}^*
Sample Entering Stage 3 $n_{(3)}=n_{(2)}-n_{te}$	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	1	0	-£10,400	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	2	0	-£9,952	-£20,800	-	-	-	-	-	-	-	-	-	-	-	0	0
	3	0	-£8,395	-£18,405	-£31,200	-	-	-	-	-	-	-	-	-	-	0	0
	4	0	-£7,031	-£8,728	-£26,905	-£41,600	-	-	-	-	-	-	-	-	-	0	0
	5	0	-£6.015	£712	-£8,166	-£35,812	-£52,000	-	-	-	-	-	-	-	-	£712	2
	6	0	-£5,257	£8,216	£11,289	-£9,355	-£45,065	-£62,400	-	-	-	-	-	-	-	£11,289	3
	7	0	-£4,677	£14,053	£27,397	£19,266	-£12,344	-£54,572	-£72,800	-	-	-	-	-	-	£27,397	3
	8	0	-£4,224	£18,645	£40,301	£43,646	£24,092	-£16,786	-£64,259	-£83,200	-	-	-	-	-	£43,646	4
	9	0	-£3,860	£22,314	£50,664	£63,540	£55,766	£26,223	-£22,300	-74,084	-£93,600	-	-	-	-	£63,540	4
	10	0	-£3,562	£25,304	£59,072	£79,752	£81,979	£64,182	£26,199	-£28,711	-£84,010	-£104,000	-	-	-	£81,979	5
	11	0	-£3,314..	£27,795..	£66,028..	£93,093..	£103,658..	£96,005..	£69,558..	£24,479..	-35,753£	-£94,013..	-£114,400..	-	-	£103,658	5
	12	0	-£3,108	£29,906	£71,840	£104,237	£121,602	£122,536	£106,255	£72,503	£21,399	-£43,304	-£104,078	-£124,800	-	£122,536	6
	..	0	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

Figure 6.3.2a Optimal Allocation at Stage 3

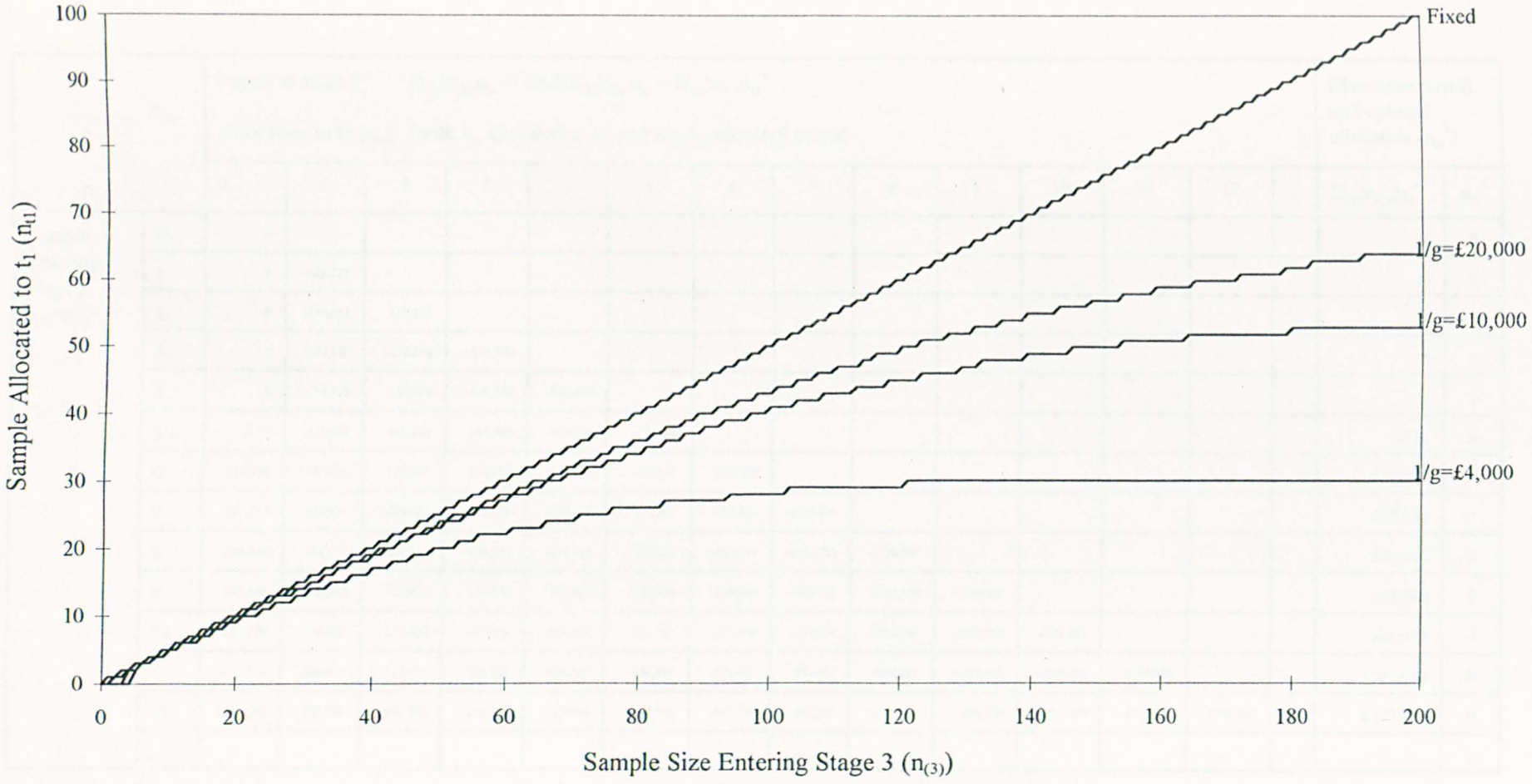


Table 6.3.1b Optimal Allocation of the Sample Entering Stage 2 (1/g=£4,000)

n_{te} / $n_{(2)}$		Payoff at stage 2 $\Pi_{(2)}\|n_{(2)},n_{te} = ENBS_{(2)}\|n_{(2)},n_{t1} + \Pi_{(3)}\|n_{(3)},n_{t1}^*$ Allocation to te (n_{te}) (with n_{te} allocated to te, and $n_{(2)}-n_{te}$ allocated to nte)														Maximum payoff and optimal allocation (n_{te}^*)	
		0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(2)}\|n_{(2)},n_{te}^*$	n_{te}^*
Sample Entering Stage 2 $n_{(2)}=n_{(1)}$	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	1	0	-£15,120	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	2	0	-£15,120	-£30,240	-	-	-	-	-	-	-	-	-	-	-	0	0
	3	0	-£15,120	-£30,236	-£45,360	-	-	-	-	-	-	-	-	-	-	0	0
	4	0	-£15,120	-£30,218	-£45,307	-£60,480	-	-	-	-	-	-	-	-	-	0	0
	5	£712	-£15,120	-£30,212	-£44,990	-£60,274	-£75,600	-	-	-	-	-	-	-	-	£712	0
	6	£11,289	-£14,408	-£30,212	-£44,853	-£58,851	-£75,126	-£90,720	-	-	-	-	-	-	-	£11,289	0
	7	£27,397	-£3,831	-£29,495	-£44,830	-£58,146	-£71,493	-£89,886	-£105,840	-	-	-	-	-	-	£27,397	0
	8	£43,646	£12,277	-£18,917	-£44,054	-£57,969	-£69,524	-£82,953	-£104,585	-£120,960	-	-	-	-	-	£43,646	0
	9	£63,540	£28,526	-£2,811	£33,450	£57,024	£68,932	£78,941	-£93,440	-£119,249	-£136,080	-	-	-	-	£63,540	0
	10	£81,979	£48,420	£13,438	-£17,366	-£46,322	-£67,709	-£77,590	-£86,637	-£103,204	-£133,898	-£151,200	-	-	-	£81,979	0
	11	£103,658	£66,859	£33,329	-£1,103	-£30,261	-£56,799	-£76,017	-£84,169	-£92,958	-£112,465	-£148,545	-£166,320	-	-	£103,658	0
	12	£122,538	£88,538	£51,769	£18,753	-£13,948	-£40,703	-£64,779	-£82,211	-£89,034	-£98,250	-£121,409	-£163,201	-£181,440	-	£122,536	0
	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

Figure 6.3.2b Optimal Allocation of Sample at Stage 2 to te (n_{te})

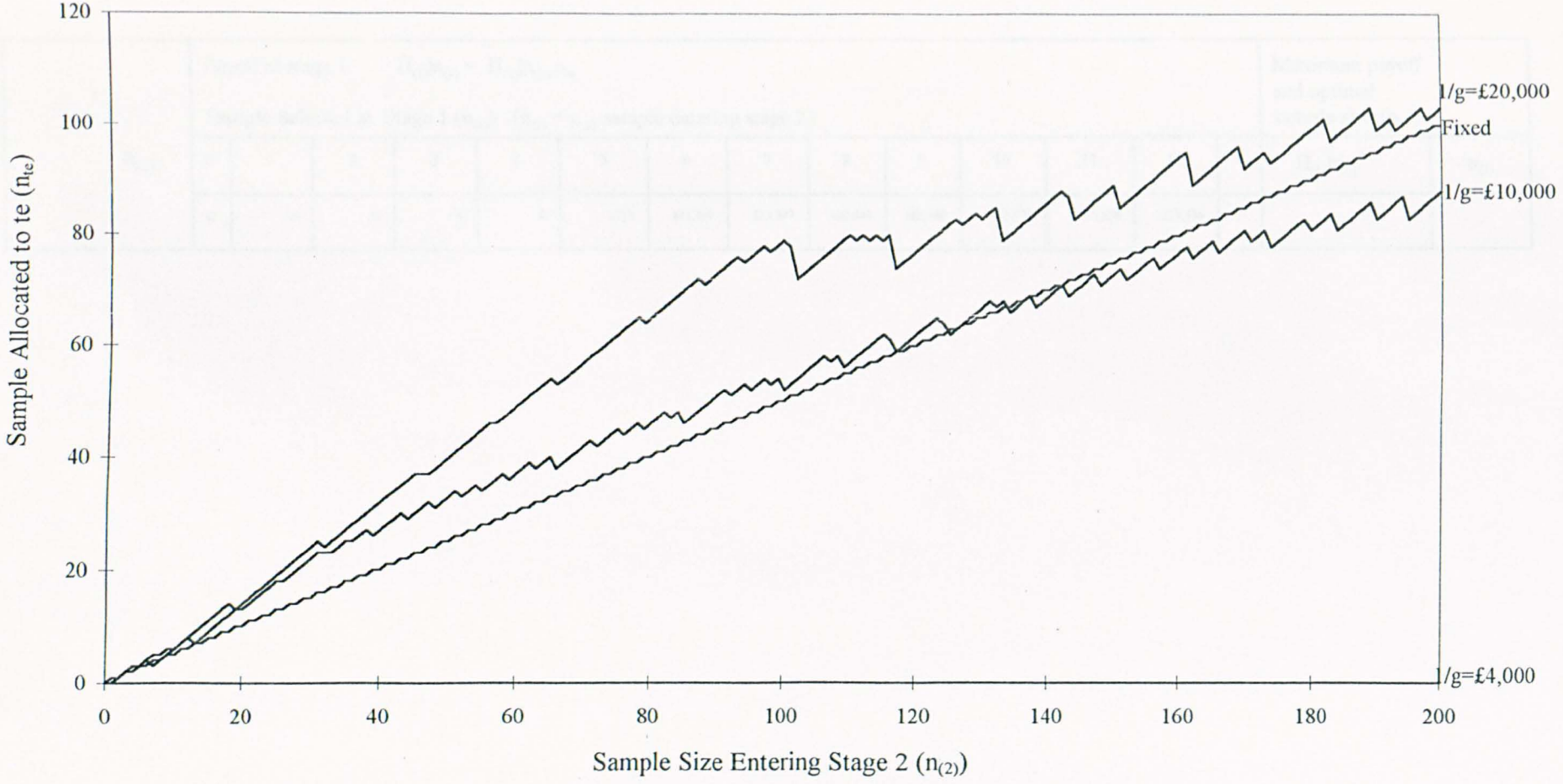


Table 6.3.1c Optimal Sample Size at Stage 1 (1/g=£4,000)

$n_{(1)}$	Payoff at stage 1 $\Pi_{(1)}\|n_{(1)} = \Pi_{(2)}\|n_{(2)},n_{te}$ Sample Selected at Stage 1 ($n_{(1)}$) ($n_{(1)} = n_{(2)}$ sample entering stage 2)														Maximum payoff and optimal sample size (n_{te}^*)	
	0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(1)}\|n_{(1)}^*$	$n_{(1)}$
	0	0	0	0	0	£712	£11,289	£27,397	£43,646	£63,540	£81,979	£103,658	£122,536	-	..	..

Figure 6.3.3a ENBS with Optimal Allocation (1/g=£4,000)

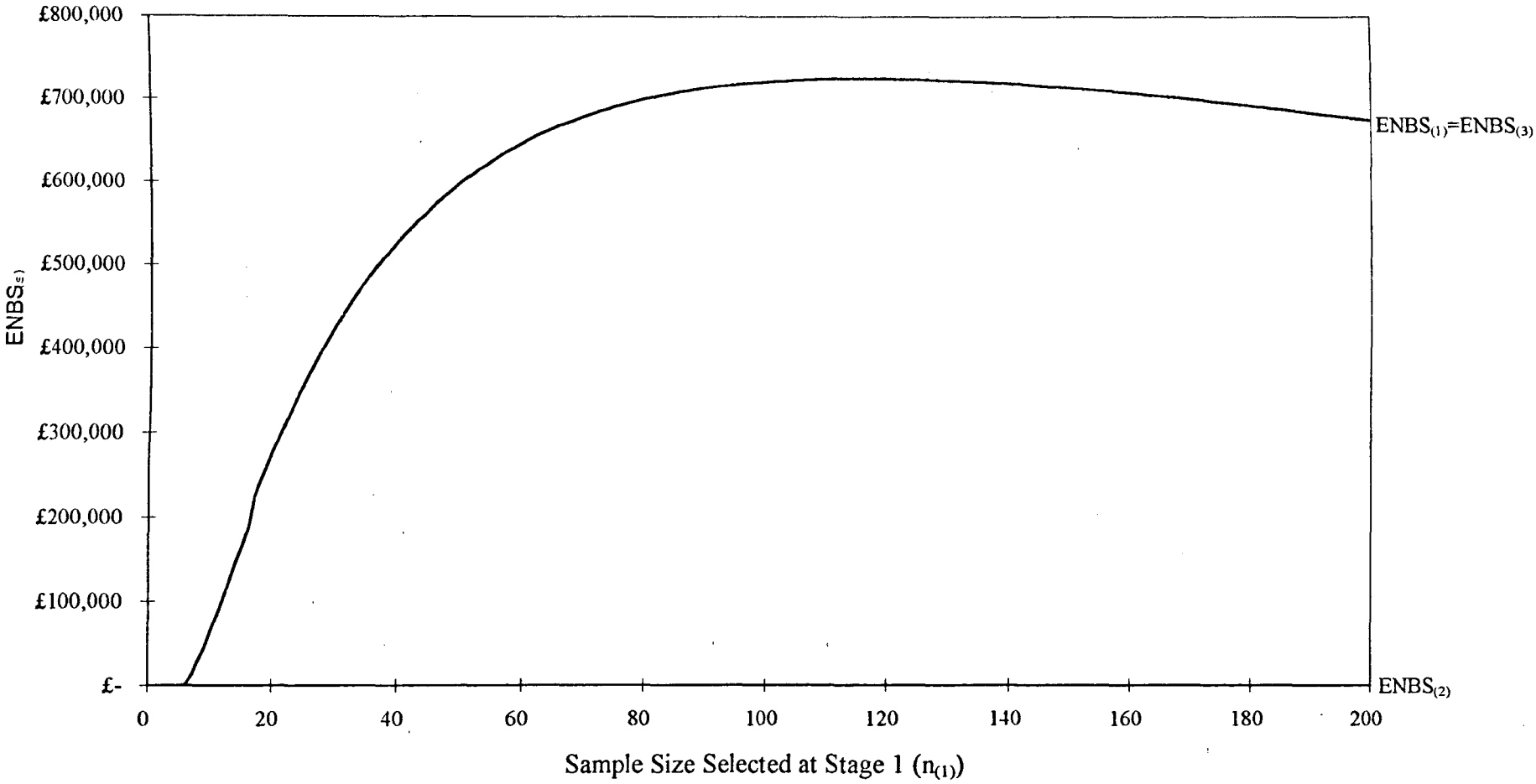
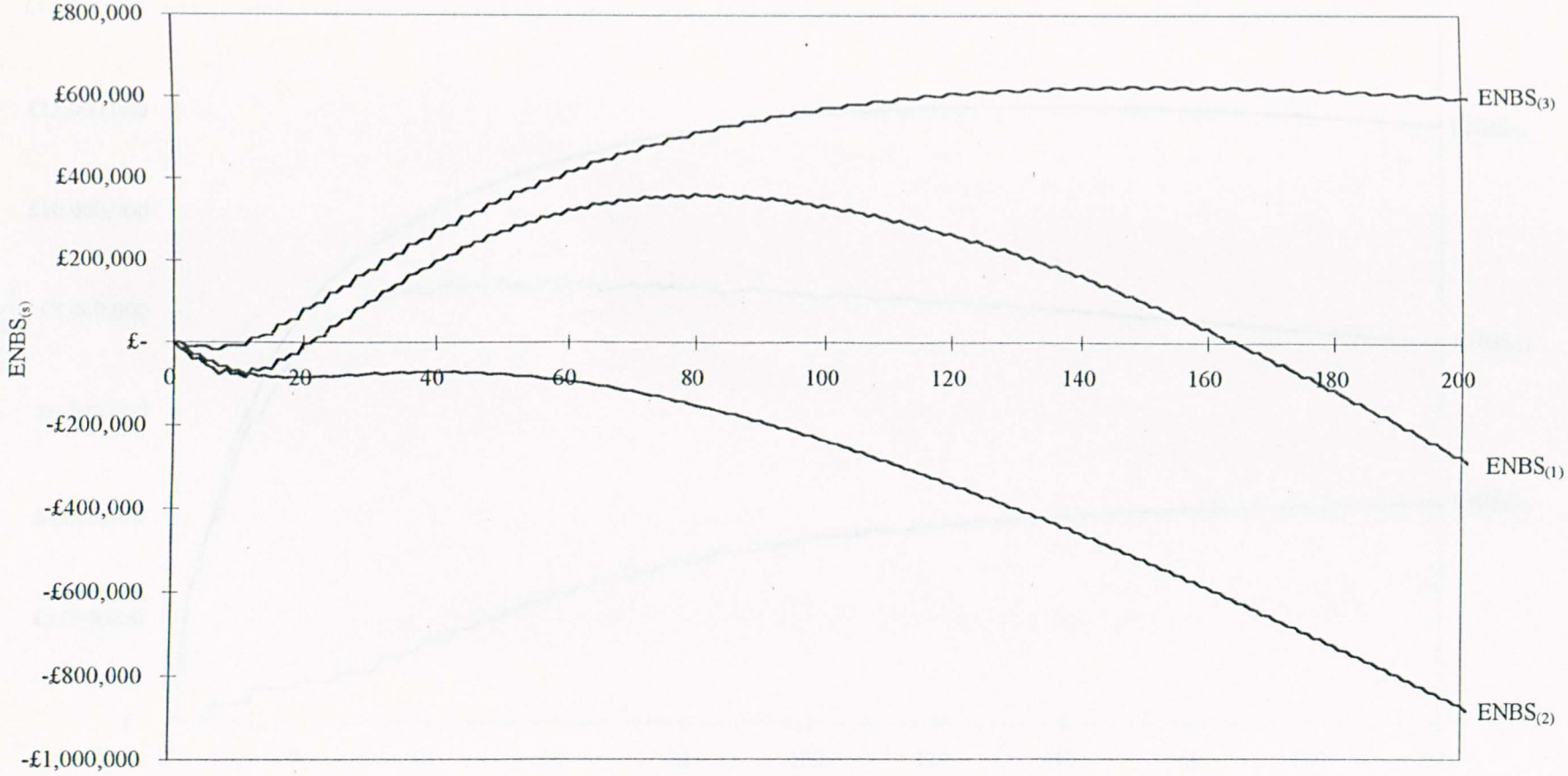


Figure 6.3.3b ENBS with Fixed Allocation (1/g=£4,000)



Sample Selected at stage 1 ($n_{(1)}$)

Figure 6.3.4a ENBS with Optimal Allocation (1/g=£10,000)

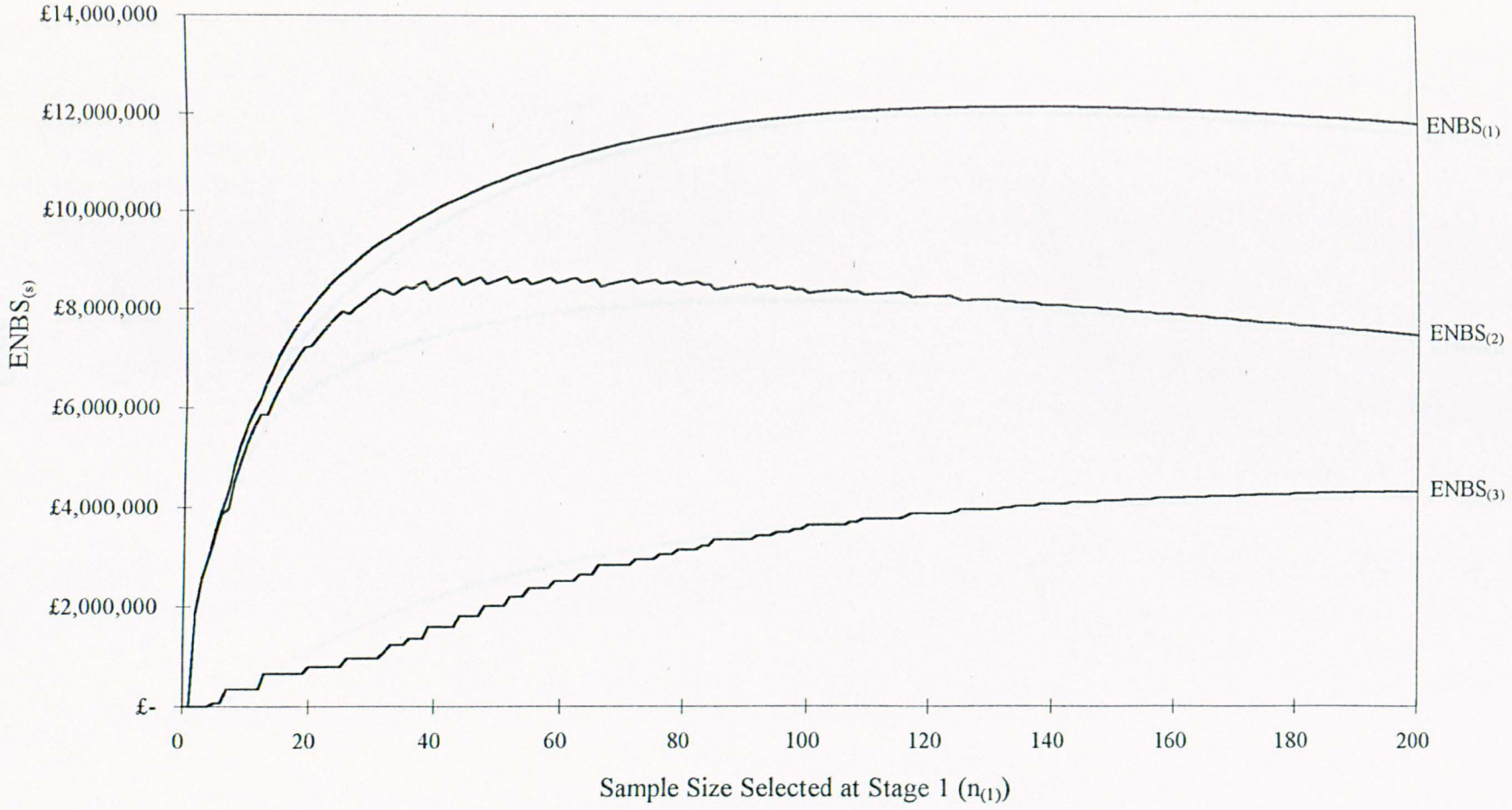


Figure 6.3.4b ENBS with Fixed Allocation (1/g=£10,000)

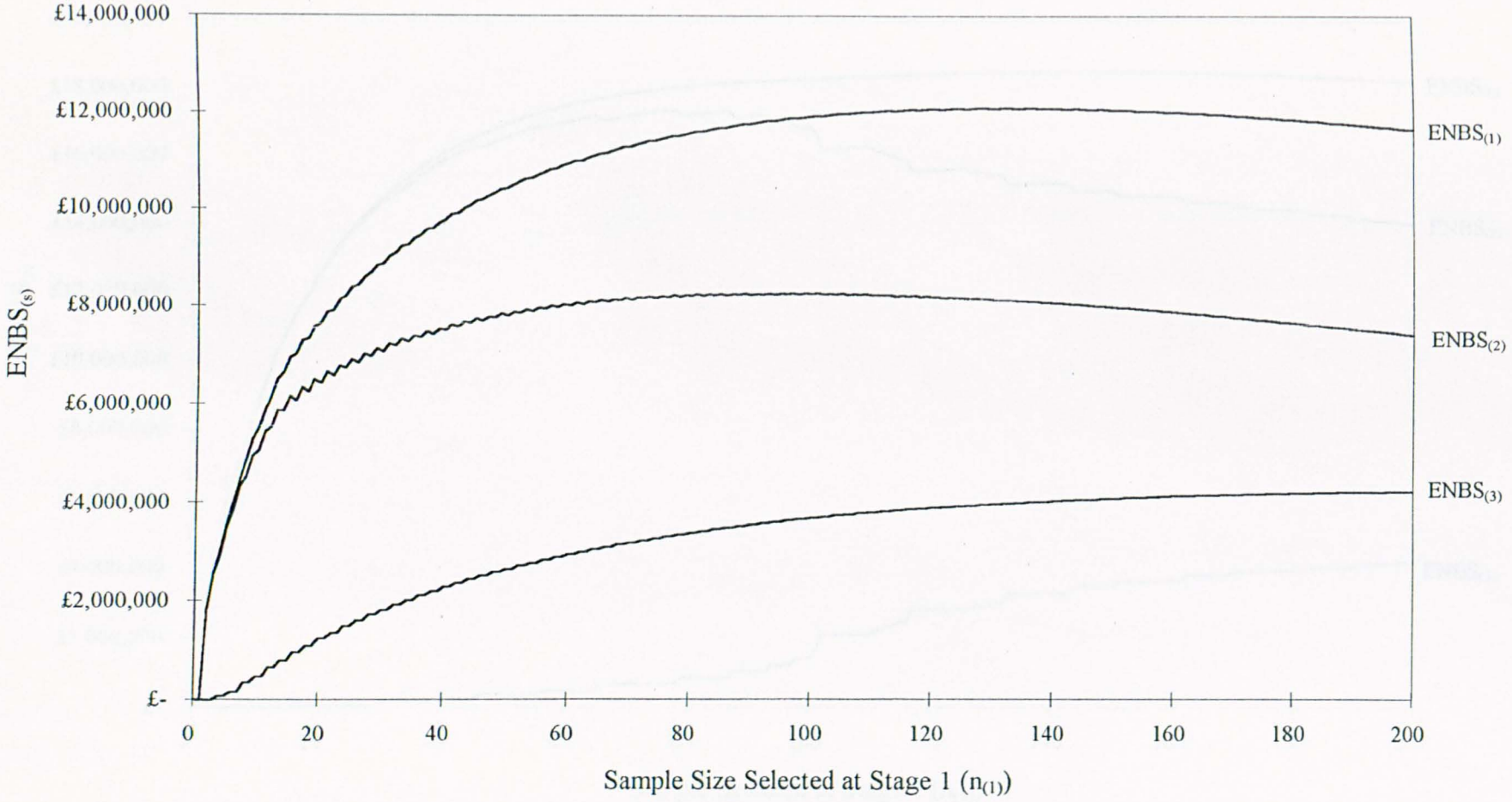


Figure 6.3.5a ENBS with Optimal Allocation (1/g=£20,000)

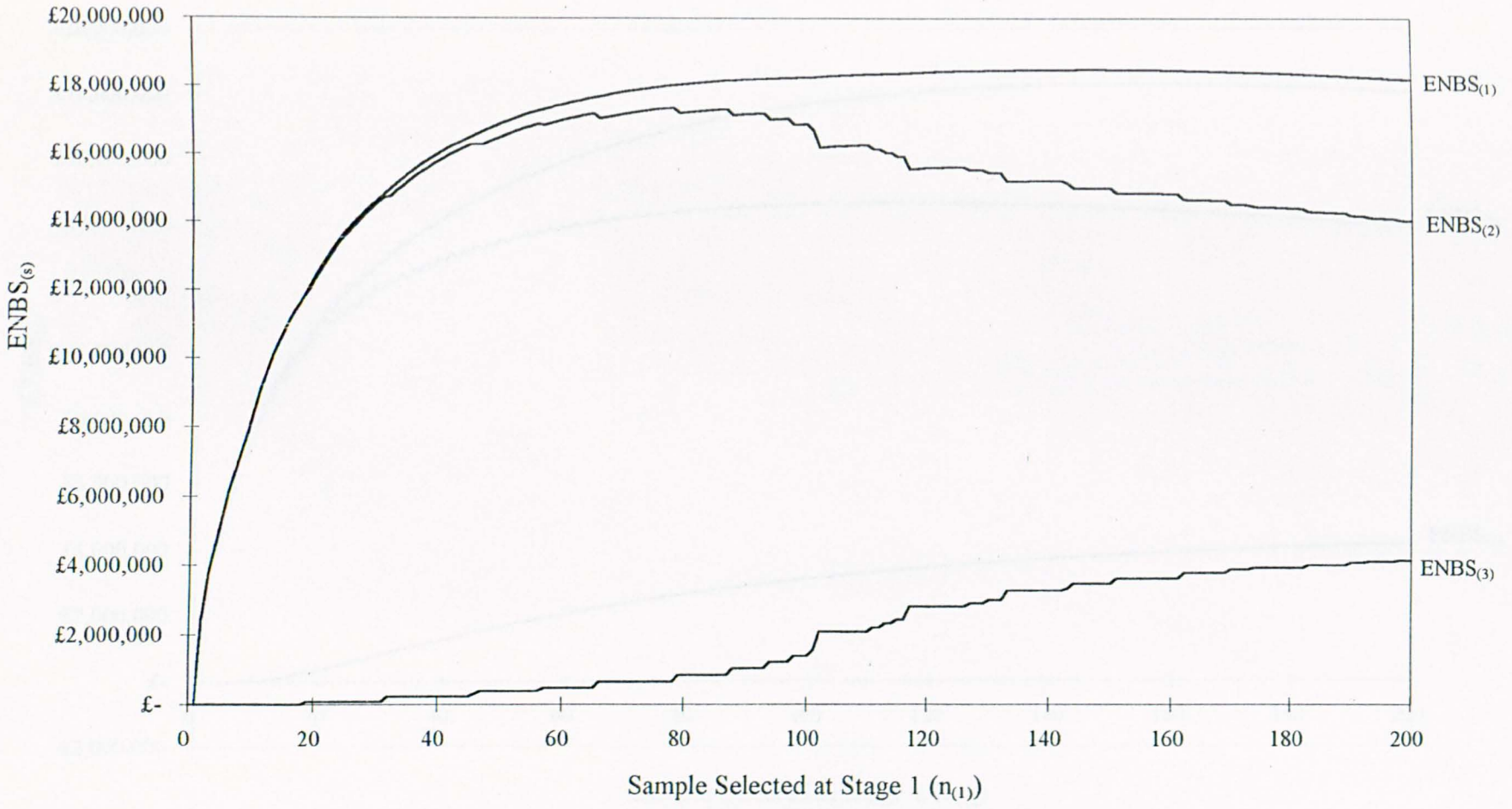


Figure 6.3.5b ENBS with Fixed Allocation (1/g=£20,000)

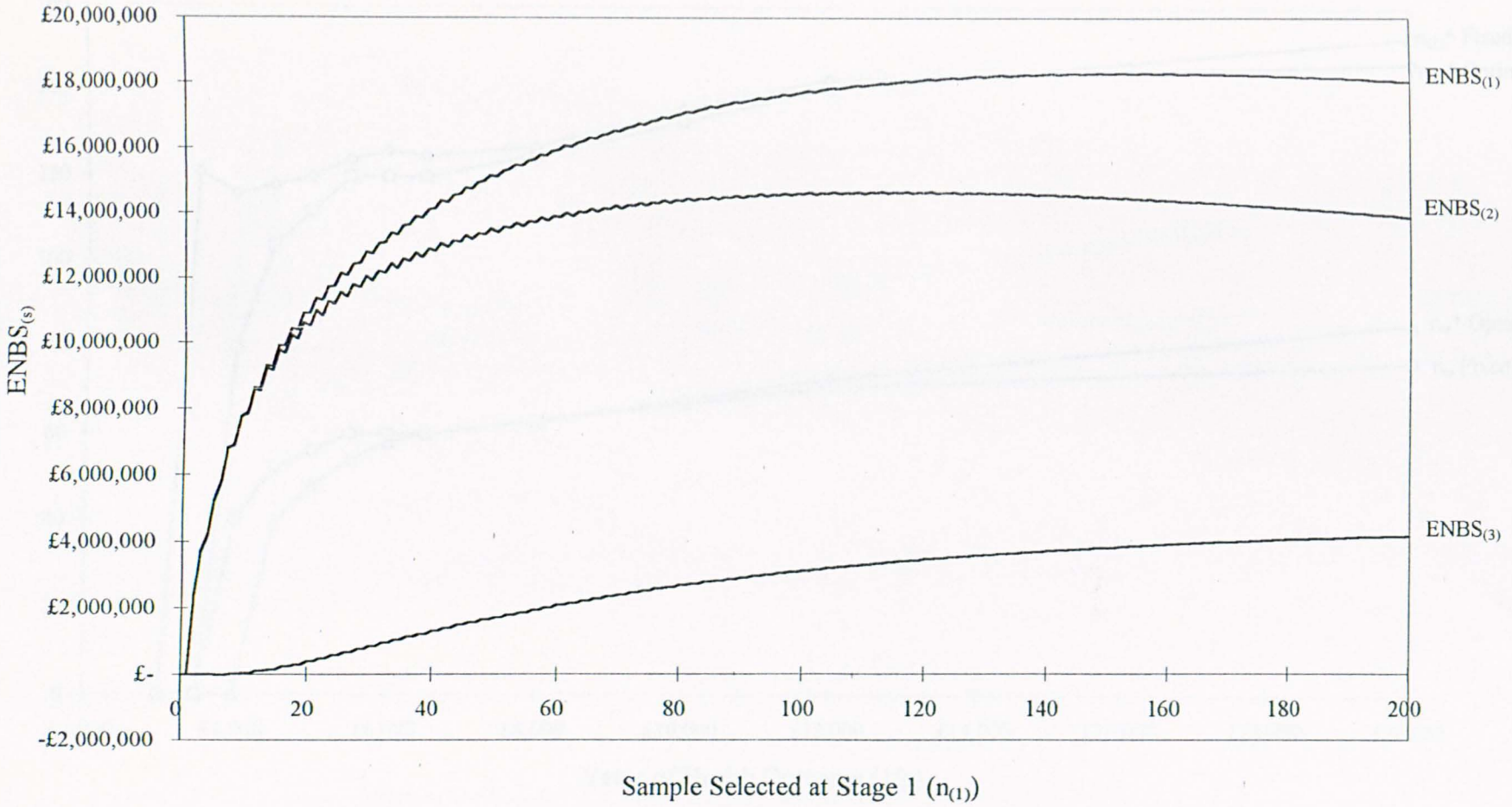


Figure 6.3.6 Optimal Sample Size with Fixed and Optimal Allocation

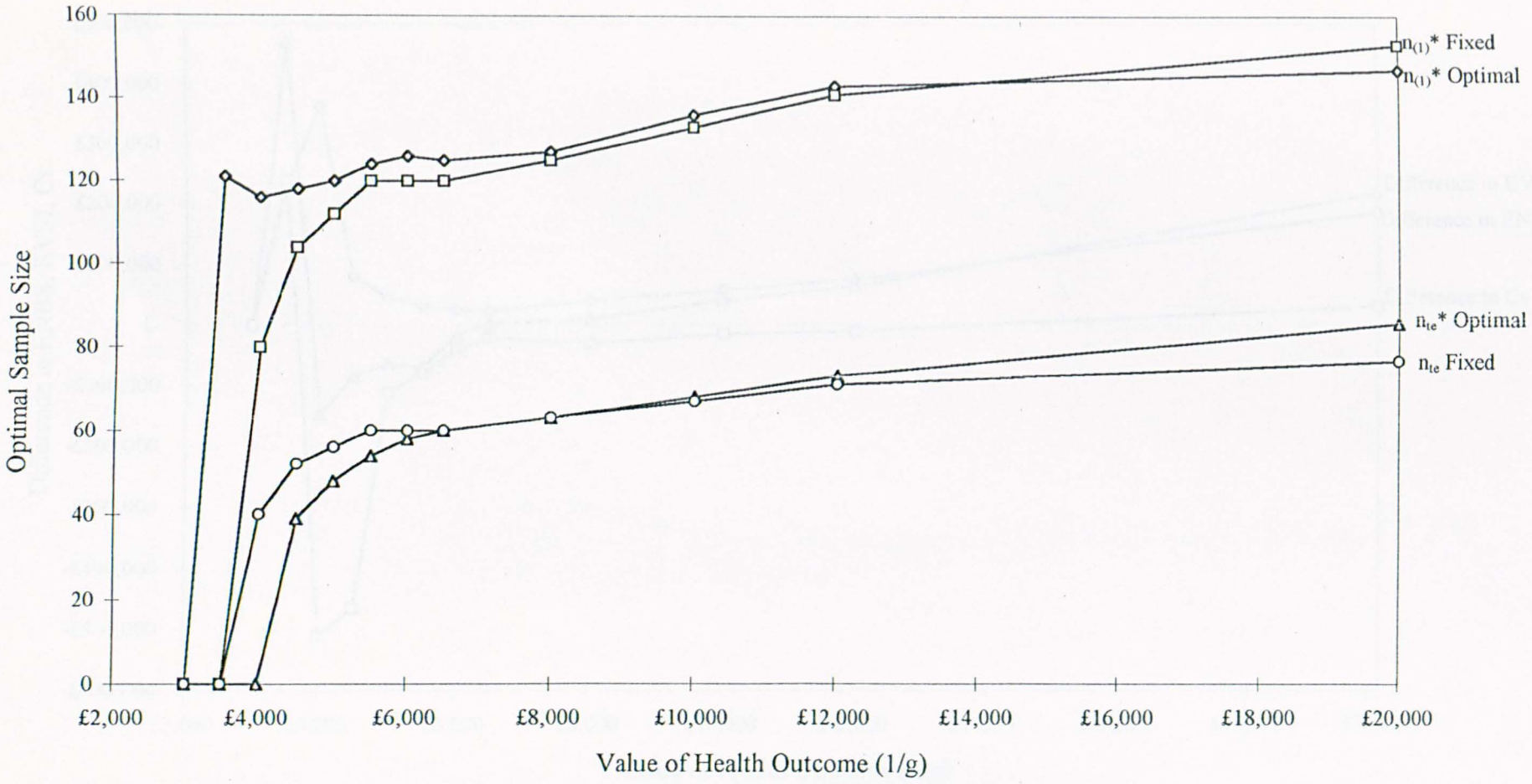


Figure 6.3.7 Difference Between Optimal and Fixed Allocation

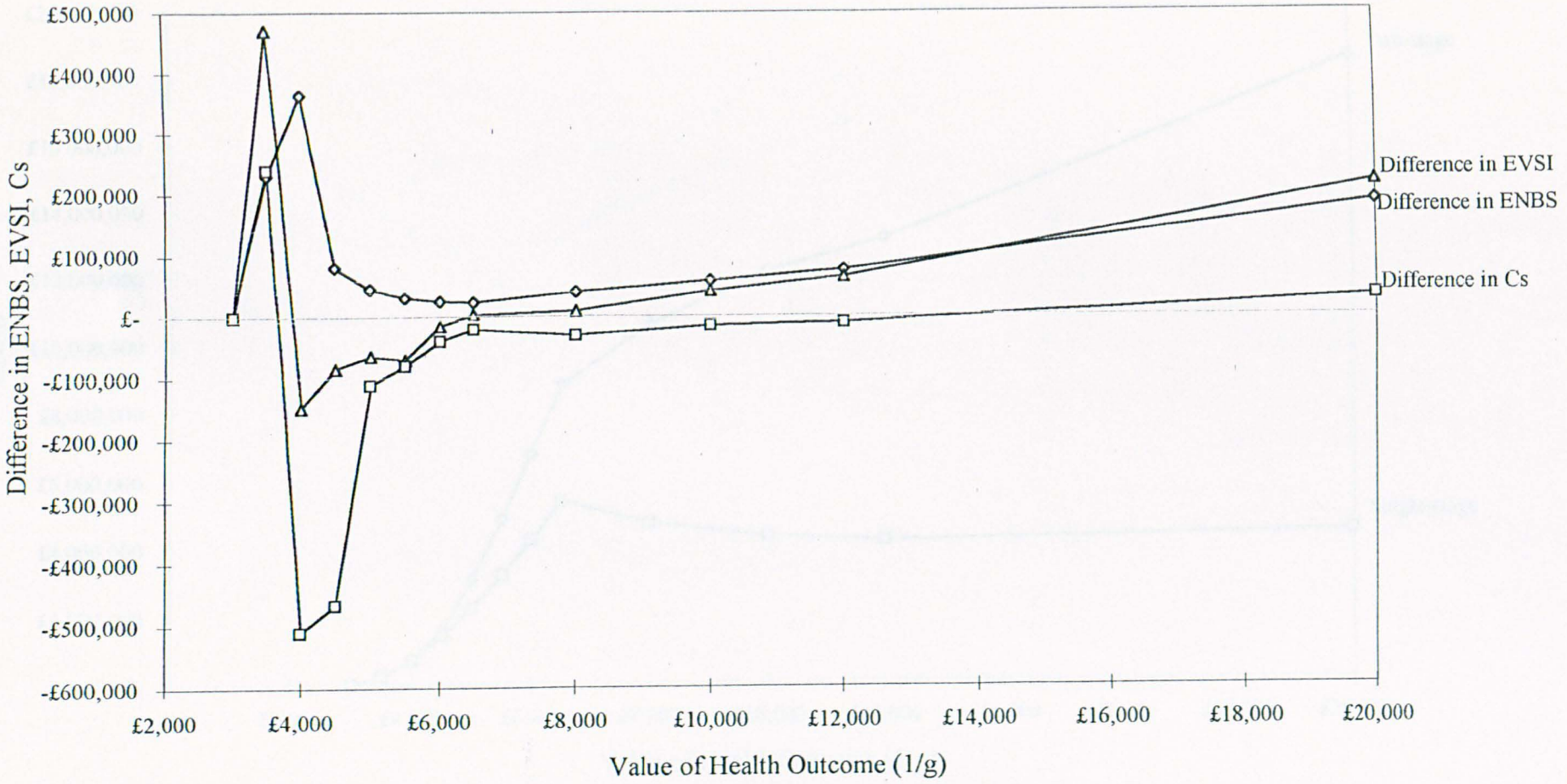


Figure 6.3.8a ENBS for the Single and Two Stage Decision Problems

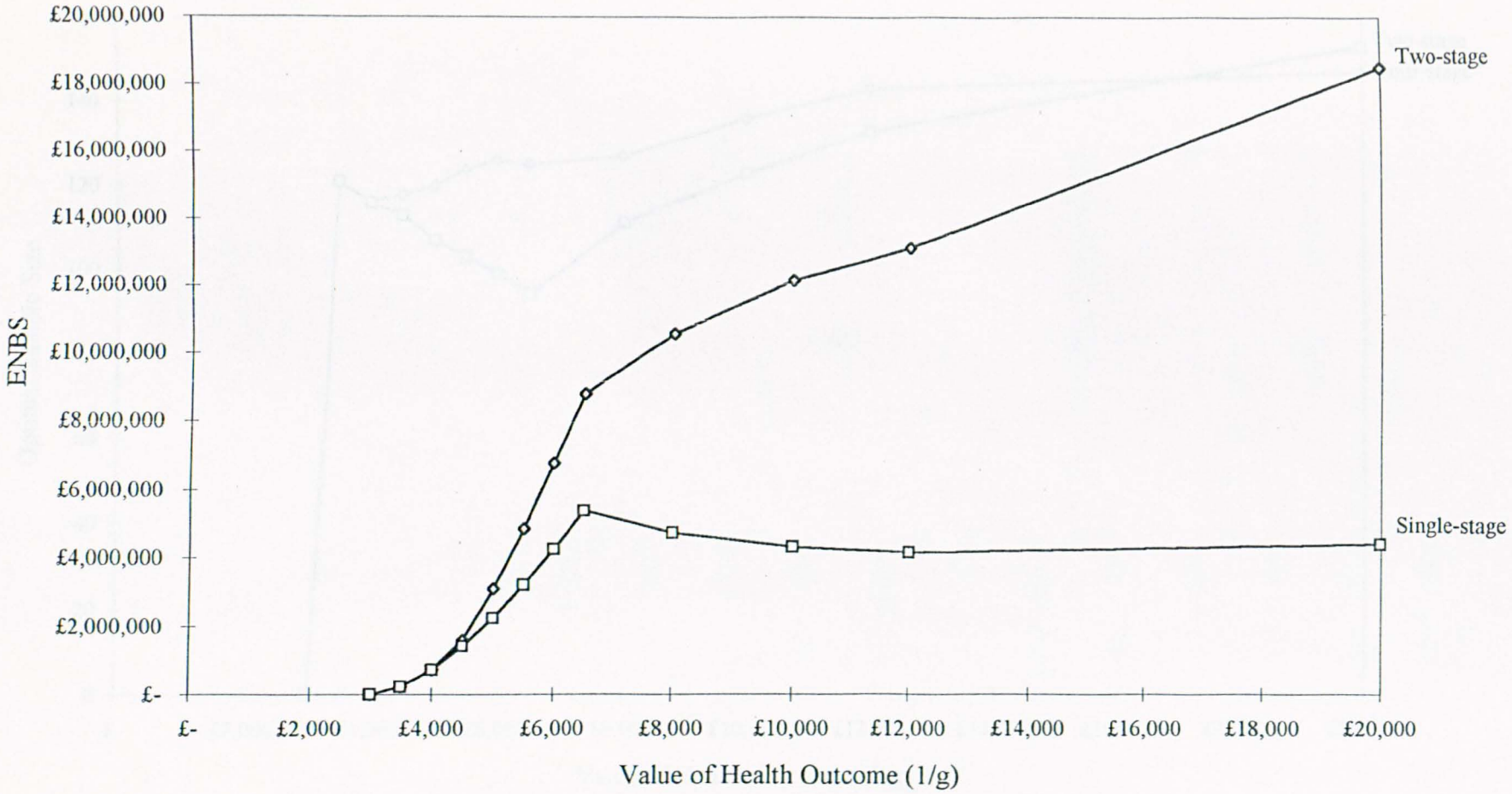


Figure 6.3.8b Optimal Sample Size for the Single and Two-Stage Decision Problems

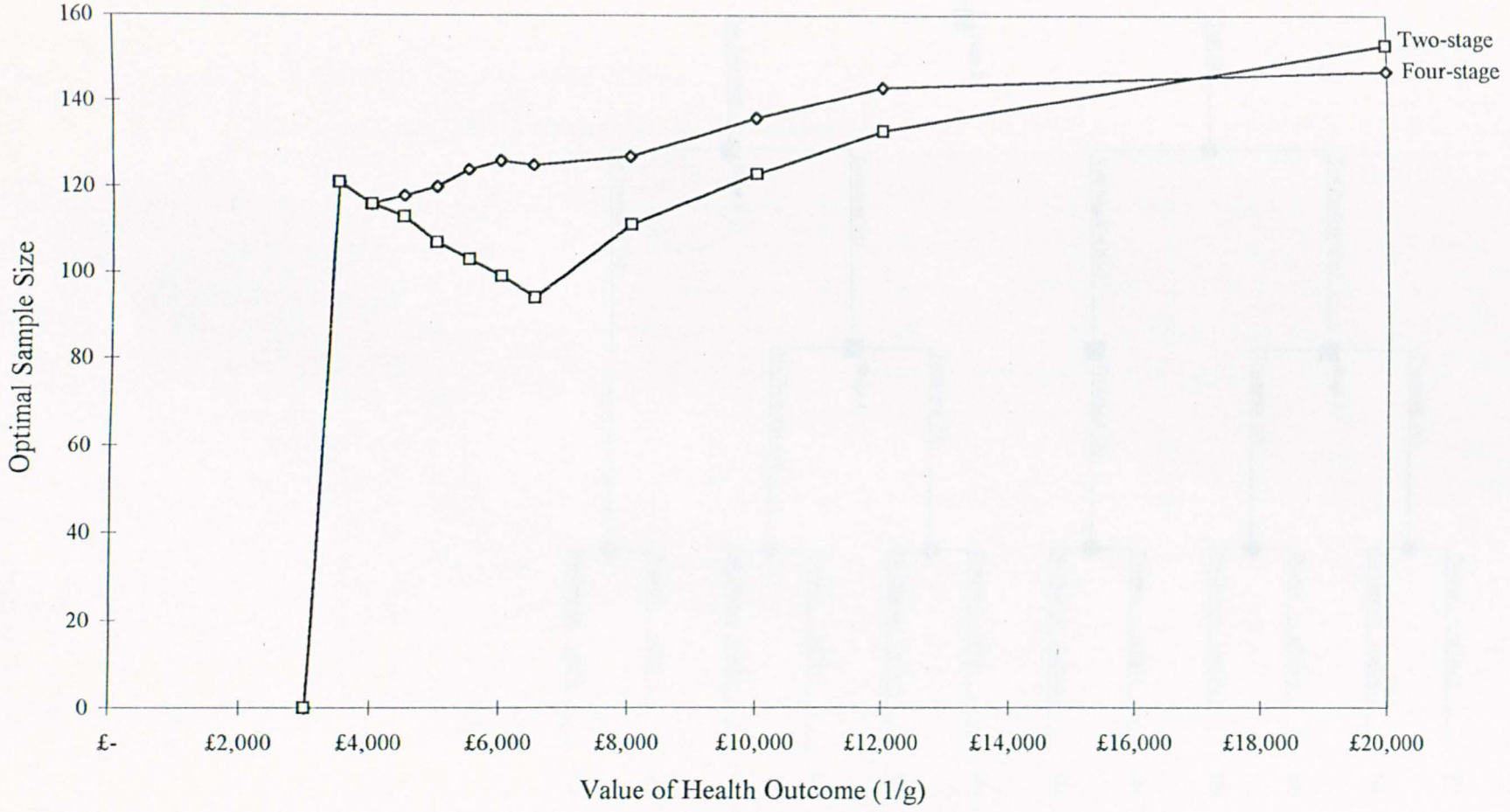


Figure 6.4.1 Decision Tree for the Four Stage Decision Problem

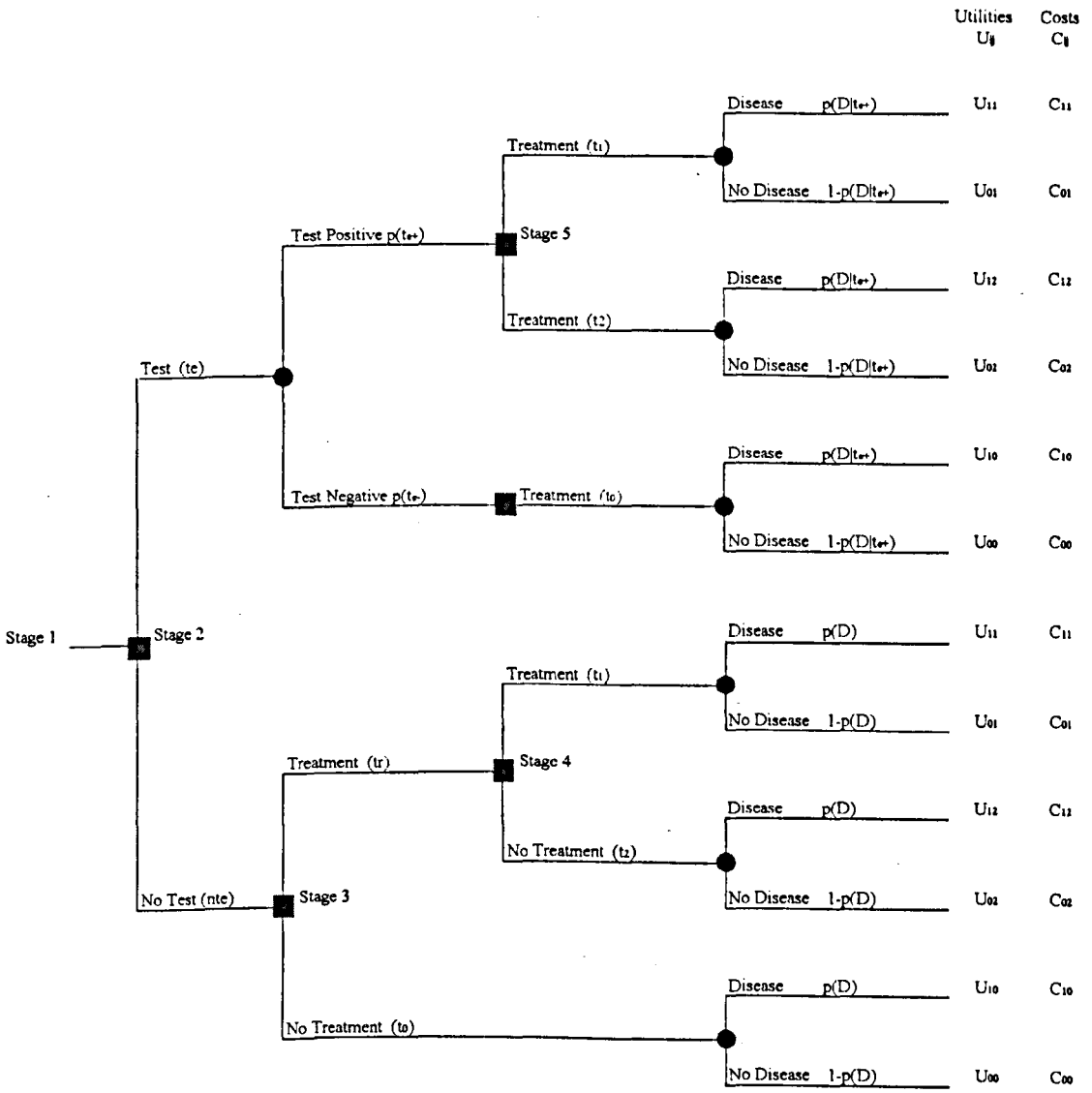


Table 6.4.1a Optimal Allocation of the Sample Entering Stage 5 (1/g=£4,000)

n_{t1} / $n_{(5)}$		Payoff at stage 5 $\Pi_{(5)}\|n_{(5)},n_{t1}$ = $ENBS_{(5)}\|n_{(5)},n_{t1}$ Allocation to t_1 (n_{t1}) (with n_{t1} allocated to t_1, and $n_{(5)}-n_{t1}$ allocated to t_2)														Maximum payoff and optimal allocation (n_{t1}^*)	
		0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(5)}\|n_{(5)},n_{t1}^*$	n_{t1}^*
Sample Entering Stage 5 $n_{(5)}= p(t_e^+).n_{te}$	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	1	-£2,400	-£11,484	-	-	-	-	-	-	-	-	-	-	-	-	-£2,400	0
	2	-£4,800	£1,458	-£22,968	-	-	-	-	-	-	-	-	-	-	-	£1,458	1
	3	-£7,200	£15,631	£10,833	-£34,452	-	-	-	-	-	-	-	-	-	-	£15,631	1
	4	-£9,600	£22,714	£55,929	£12,467	-£45,935	-	-	-	-	-	-	-	-	-	£55,929	2
	5	-£12,000	£26,200	£84,219	£81,378	£9,533	-£57,419	-	-	-	-	-	-	-	-	£84,219	2
	6	-£14,400	£27,780	£102,306	£128,535	£95,897	£3,990	-£68,903	-	-	-	-	-	-	-	£128,535	3
	7	-£16,800	£28,204	£114,376	£160,815	£158,261	£103,413	-£3,172	-£80,387	-	-	-	-	-	-	£160,815	3
	8	-£19,200	£27,935	£122,651	£183,667	£202,633	£177,887	£106,307	-£11,394	-£91,871	-	-	-	-	-	£202,633	4
	9	-£21,600	£27,158	£128,419	£200,367	£235,156	£232,414	£190,583	£106,036	-£20,315	-£103,355	-	-	-	-	£235,156	4
	10	-£24,000	£26,046	£132,436	£212,860	£259,627	£273,256	£253,568	£198,320	£103,488	-£29,766	-£114,839	-	-	-	£273,256	5
	11	-£26,400	£24,683	£135,194	£222,309	£278,462	£304,755	£301,709	£268,332	£202,425	£99,288	-£39,607	-£126,323	-	-	£304,755	5
	12	-£28,800	£23,124	£137,016	£229,593	£293,127	£329,239	£339,165	£322,738	£278,730	£203,808	£93,811	-£49,736	-£137,806	-	£339,165	6
	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

Figure 6.4.2a Optimal Allocation at Stage 5

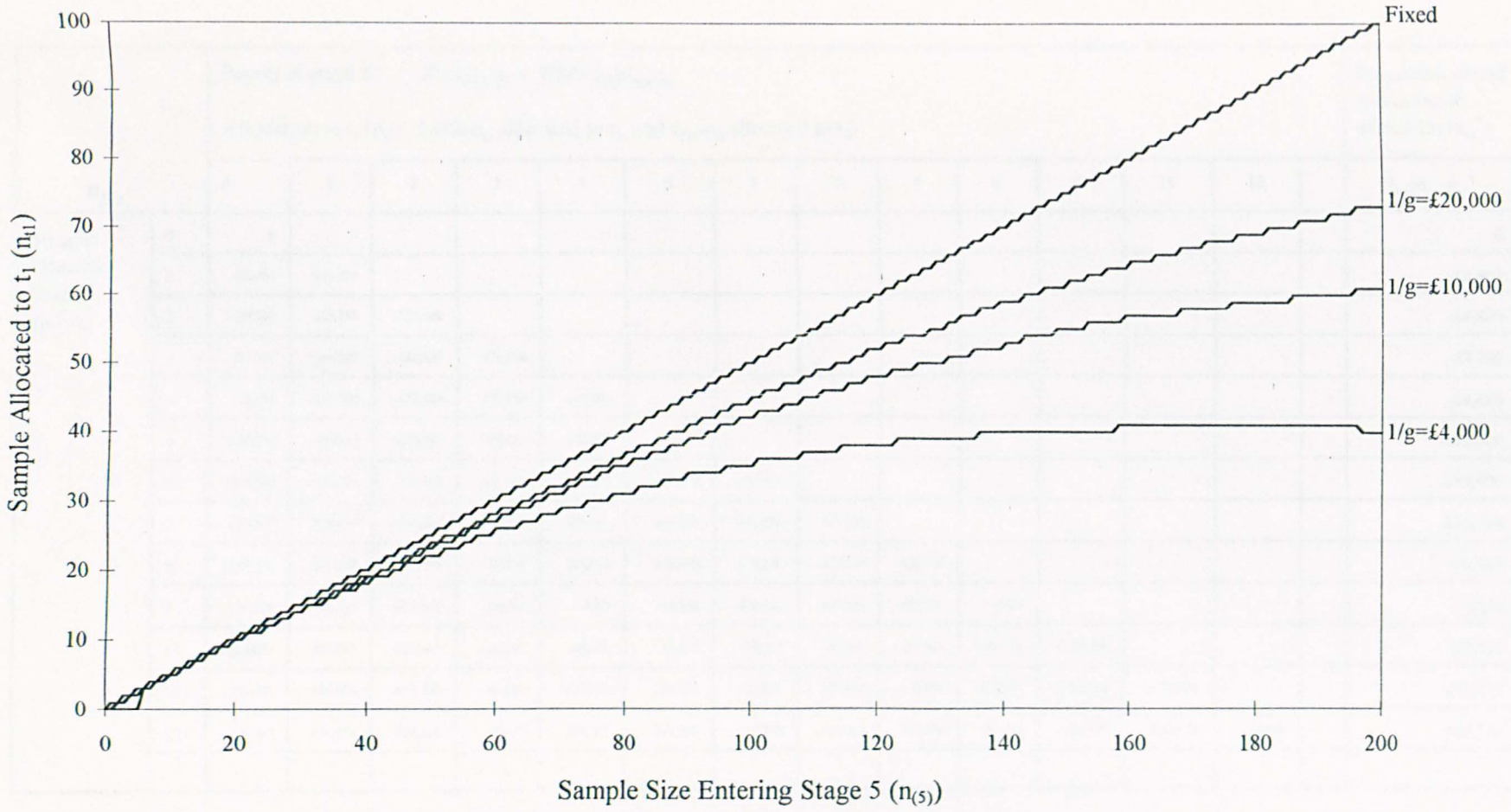


Table 6.4.1b Optimal Allocation of the Sample Entering Stage 4 (1/g=£4,000)

n_{t1} / $n_{(4)}$		Payoff at stage 4 $\Pi_{(4)}\|n_{(4)},n_{t1} = ENBS_{(4)}\|n_{(4)},n_{t1}$ Allocation to t_1 (n_{t1}) (with n_{t1} allocated to t_1, and $n_{(4)}$-n_{t1} allocated to t_2)														Maximum payoff and optimal allocation (n_{t1}^*)	
		0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(4)}\|n_{(4)},n_{t1}^*$	n_{t1}^*
Sample Entering Stage 4 $n_{(4)}=n_{tr}$	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	1	-£2,400	-£10,400	-	-	-	-	-	-	-	-	-	-	-	-	-£2,400	0
	2	-£4,800	-£12,753	-£20,800	-	-	-	-	-	-	-	-	-	-	-	-£4,800	0
	3	-£7,200	-£14,806	-£22,837	-£31,200	-	-	-	-	-	-	-	-	-	-	-£7,200	0
	4	-£9,600	-£16,782	-£22,179	-£32,869	-£41,600	-	-	-	-	-	-	-	-	-	-£9,600	0
	5	-£12,000	-£18,815	-£20,421	-£28,652	-£42,957	-£52,000	-	-	-	-	-	-	-	-	-£12,000	0
	6	-£14,400	-£20,916	-£18,970	-£21,510	-£35,501	-£53,106	-£62,400	-	-	-	-	-	-	-	-£14,400	0
	7	-£16,800	-£23,076	-£18,104	-£14,704	-£23,141	-£42,931	-£63,306	-£72,800	-	-	-	-	-	-	-£14,704	3
	8	-£19,200	-£25,280	-£17,785	£9,069	£10,905	£25,998	-£50,882	-£73,545	-£83,200	-	-	-	-	-	-£9,069	3
	9	-£21,600	-£27,518	-£17,918	-£4,632	-£269	-£8,846	-£30,044	-£59,256	-£83,811	-£93,600	-	-	-	-	-£269	4
	10	-£24,000	-£29,783	-£18,415	-£1,243	£8,572	£6,442	-£8,581	£35,105	-£67,962	-£94,101	-£104,000	-	-	-	£8,572	4
	11	-£26,400	-£32,069	-£19,206	£1,266	£15,872	£19,522	£10,871	-£9,917	-£40,991	-£76,935	-£104,408	-£114,400	-	-	£19,522	5
	12	-£28,800	-£34,372	-£20,226	£3,077	£21,755	£30,518	£27,738	£13,203	-£12,570	-£47,543	-£86,120	-£114,728	-£124,800	-	£30,518	5
	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

Figure 6.4.2b Optimal Allocation at Stage 4

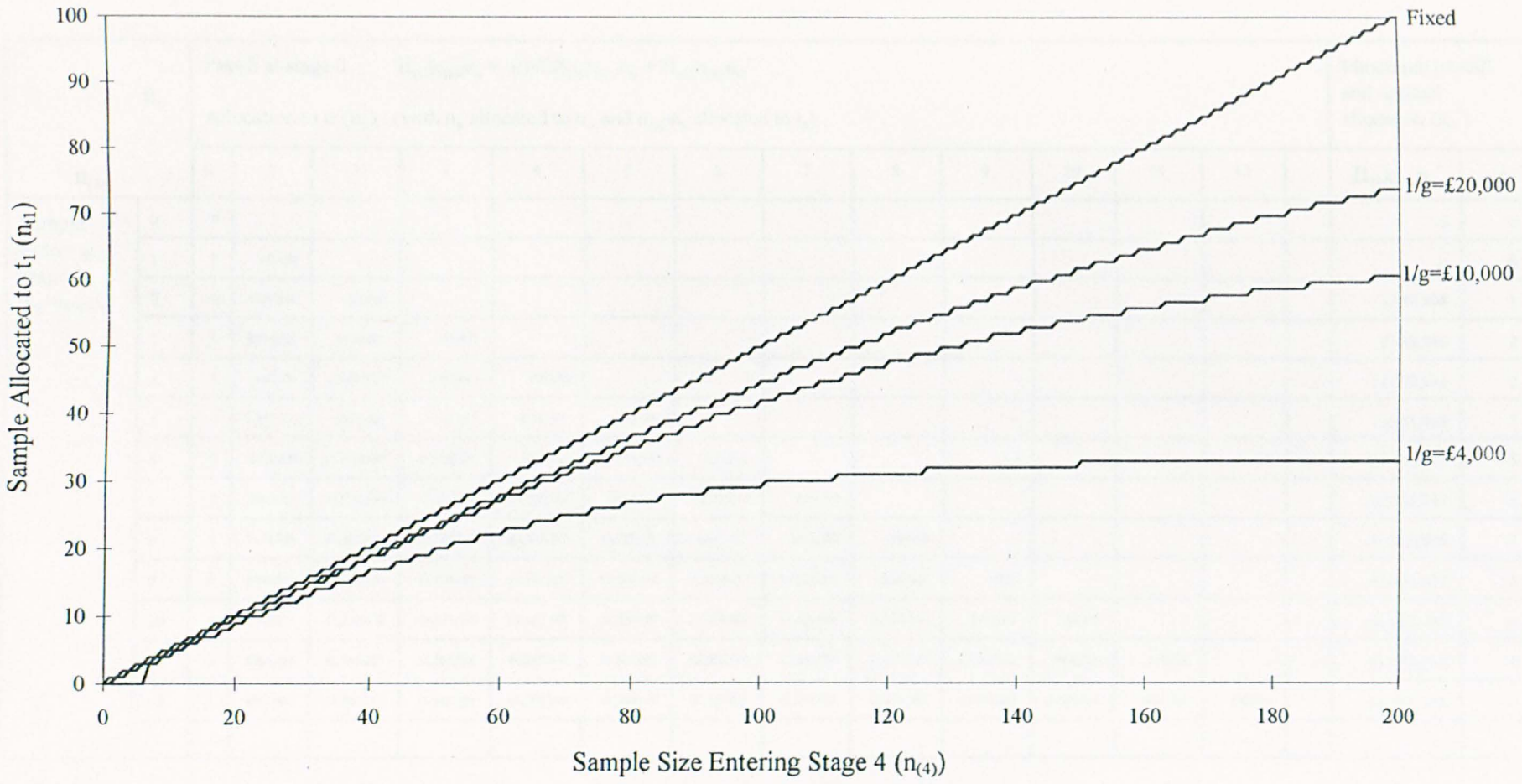


Table 6.4.1c Optimal Allocation of the Sample Entering Stage 3 (1/g=£4,000)

$n_{(3)}$ \ n_{tr}		Payoff at stage 3 $\Pi_{(3)}\|n_{(3)},n_{tr} = ENBS_{(3)}\|n_{(3)},n_{tr} + \Pi_{(4)}\|n_{(4)},n_{t1}^*$ Allocation to tr (n_{tr}) (with n_{tr} allocated to tr, and $n_{(3)}-n_{tr}$ allocated to t_0)														Maximum payoff and optimal allocation (n_{tr}^*)	
		0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(3)}\|n_{(3)},n_{tr}^*$	n_{t1}^*
Sample Entering Stage 3 $n_{(3)}=n_{(2)}-n_{te}$	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	1	0	-£2,400	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	2	0	£359,898	-£4,800	-	-	-	-	-	-	-	-	-	-	-	£359,898	1
	3	0	£464,708	£538,405	-£7,200	-	-	-	-	-	-	-	-	-	-	£538,760	2
	4	0	£511,470	£802,531	£582,041	-£9,600	-	-	-	-	-	-	-	-	-	£802,531	2
	5	0	£537,716	£951,965	£917,413	£586,379	-£12,000	-	-	-	-	-	-	-	-	£951,965	2
	6	0	£554,465	£1,048,880	£1,128,581	£951,155	£578,408	-£14,400	-	-	-	-	-	-	-	£1,128,581	3
	7	0	£566,007	£1,116,999	£1,276,244	£1,192,266	£954,985	£565,915	-£14,704	-	-	-	-	-	-	£1,276,244	3
	8	0	£574,336	£1,167,607	£1,385,955	£1,367,582	£1,210,928	£946,474	£616,005	-£9,069	-	-	-	-	-	£1,385,955	3
	9	0	£580,575	£1,206,393	£1,470,400	£1,501,327	£1,400,610	£1,209,477	£1,003,363	£614,545	-£269	-	-	-	-	£1,501,327	4
	10	0	£585,378	£1,236,970	£1,537,934	£1,607,449	£1,548,189	£1,406,981	£1,266,829	£1,004,560	£634,691	£8,572	-	-	-	£1,607,449	4
	11	0	£589,154	£1,261,625	£1,592,824	£1,693,964	£1,667,061	£1,562,399	£1,462,268	£1,273,647	£1,027,273	£635,286	£19,522	-	-	£1,693,964	4
	12	0	£592,167	£1,281,917	£1,368,369	£1,765,246	£1,764,655	£1,688,839	£1,614,969	£1,475,597	£1,298,402	£1,026,530	£656,028	£30,518	-	£1,765,246	4
	..	0	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

Figure 6.4.2c Optimal Allocation at Stage 3

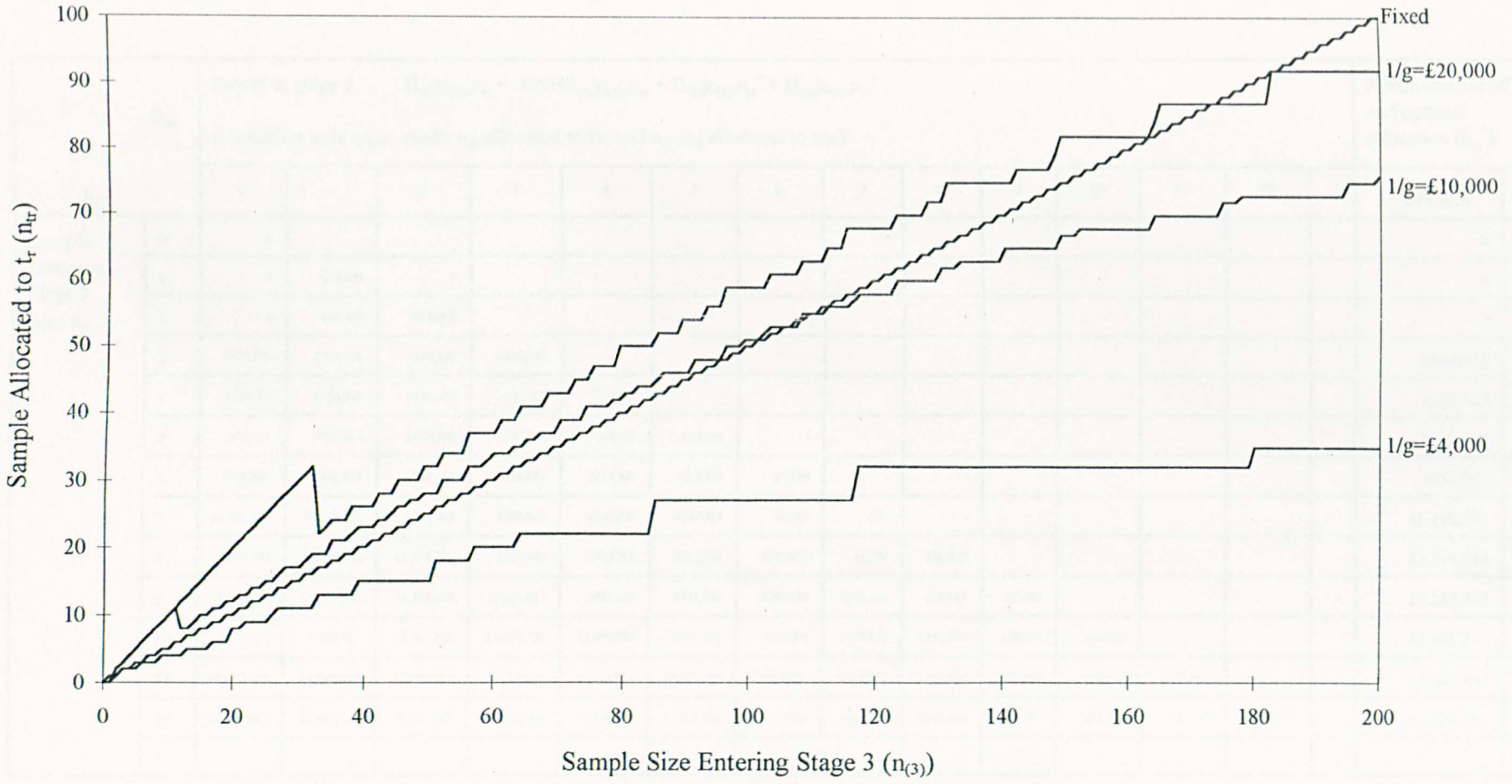


Table 6.4.1d Optimal Allocation of the Sample Entering Stage 2 (1/g=£4,000)

n_{te} / $n_{(2)}$		Payoff at stage 2 $\Pi_{(2)}\|n_{(2)},n_{te} = ENBS_{(2)}\|n_{(2)},n_{te} + \Pi_{(3)}\|n_{(3)},n_{t1}^* + \Pi_{(5)}\|n_{(5)},n_{t1}^*$ Allocation to te (n_{te}) (with n_{te} allocated to te, and $n_{(2)}$-n_{te} allocated to nte)														Maximum payoff and optimal allocation (n_{te}^*)	
		0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(3)}\|n_{(3)},n_{te}^*$	n_{t1}^*
Sample Entering Stage 2 $n_{(2)}= n_{(1)}$	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	1	0	-£10,400	-	-	-	-	-	-	-	-	-	-	-	-	0	0
	2	0	-£10,400	-£18,400	-	-	-	-	-	-	-	-	-	-	-	0	0
	3	£359,898	£349,498	-£18,399	-£22,542	-	-	-	-	-	-	-	-	-	-	£359,898	0
	4	£538,760	£528,360	£341,499	-£22,525	-£30,542	-	-	-	-	-	-	-	-	-	£538,760	0
	5	£802,531	£792,131	£520,360	£337,376	-£30,455	-£24,369	-	-	-	-	-	-	-	-	£802,531	0
	6	£951,965	£941,565	£784,131	£516,241	£329,509	-£23,750	£7,929	-	-	-	-	-	-	-	£951,965	0
	7	£1,128,581	£1,118,181	£933,565	£780,013	£508,430	£337,426	£8,787	-£71	-	-	-	-	-	-	£1,128,581	0
	8	£1,276,244	£1,265,844	£1,110,181	£929,447	£772,213	£517,735	£371,031	£1,198	£20,219	-	-	-	-	-	£1,276,244	0
	9	£1,385,955	£1,375,555	£1,257,844	£1,106,057	£921,650	£781,785	£552,918	£365,161	£22,111	£56,535	-	-	-	-	£1,385,955	0
	10	£1,501,327	£1,490,927	£1,367,555	£1,253,720	£1,098,220	£931,324	£817,291	£549,663	£389,801	£58,591	£48,535	-	-	-	£1,501,327	0
	11	£1,607,449	£1,597,049	£1,482,927	£1,363,431	£1,245,882	£1,107,470	£966,961	£814,581	£581,088	£428,163	£50,977	£72,815	-	-	£1,607,449	0
	12	£1,693,964	£1,683,564	£1,589,049	£1,478,798	£1,355,590	£1,255,138	£1,142824	£964,482	£847,518	£623,788	£423,258	£75,589	£64,815	-	£1,693,964	0
	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

Figure 6.4.2d Optimal Allocation at Stage 3

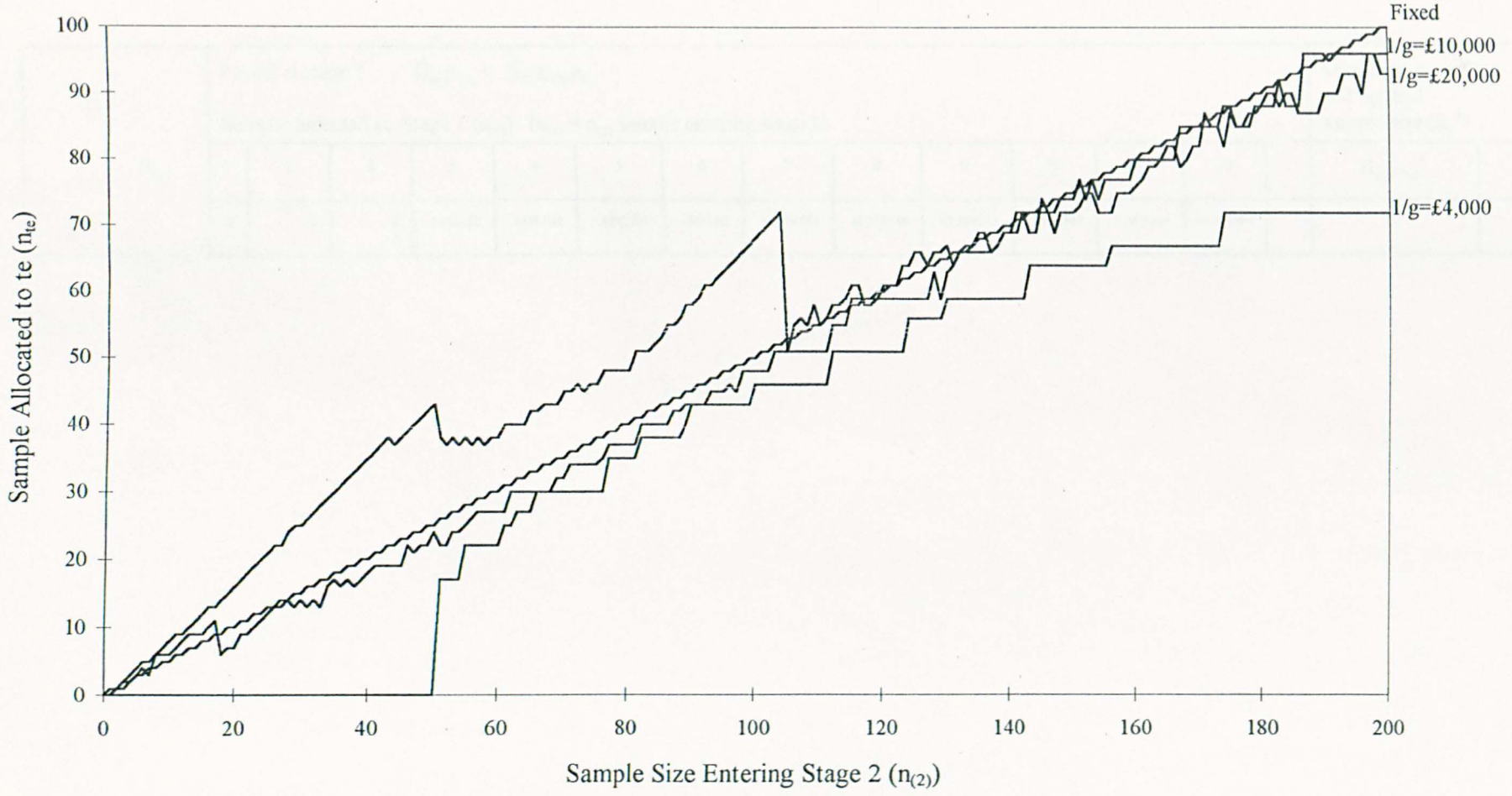


Table 6.4.1e Optimal Sample Size at Stage 1 (1/g=£4,000)

$n_{(1)}$	Payoff at stage 1 $\Pi_{(1)}\|n_{(1)} = \Pi_{(2)}\|n_{(2)},n_{te}^*$ Sample Selected at Stage 1 ($n_{(1)}$) ($n_{(1)} = n_{(2)}$ sample entering stage 2)														Maximum payoff and optimal sample size (n_{te}^*)	
	0	1	2	3	4	5	6	7	8	9	10	11	12	..	$\Pi_{(1)}\|n_{(1)}^*$	$n_{(1)}$
	0	0	0	£359,898	£538,760	£802,531	£951,965	£1,128,581	£1,276,244	£1,385,955	£1,501,327	£1,607,449	£1,693,964	-	..	..

Figure 6.4.3a ENBS with Optimal Allocation (1/g=£4,000)

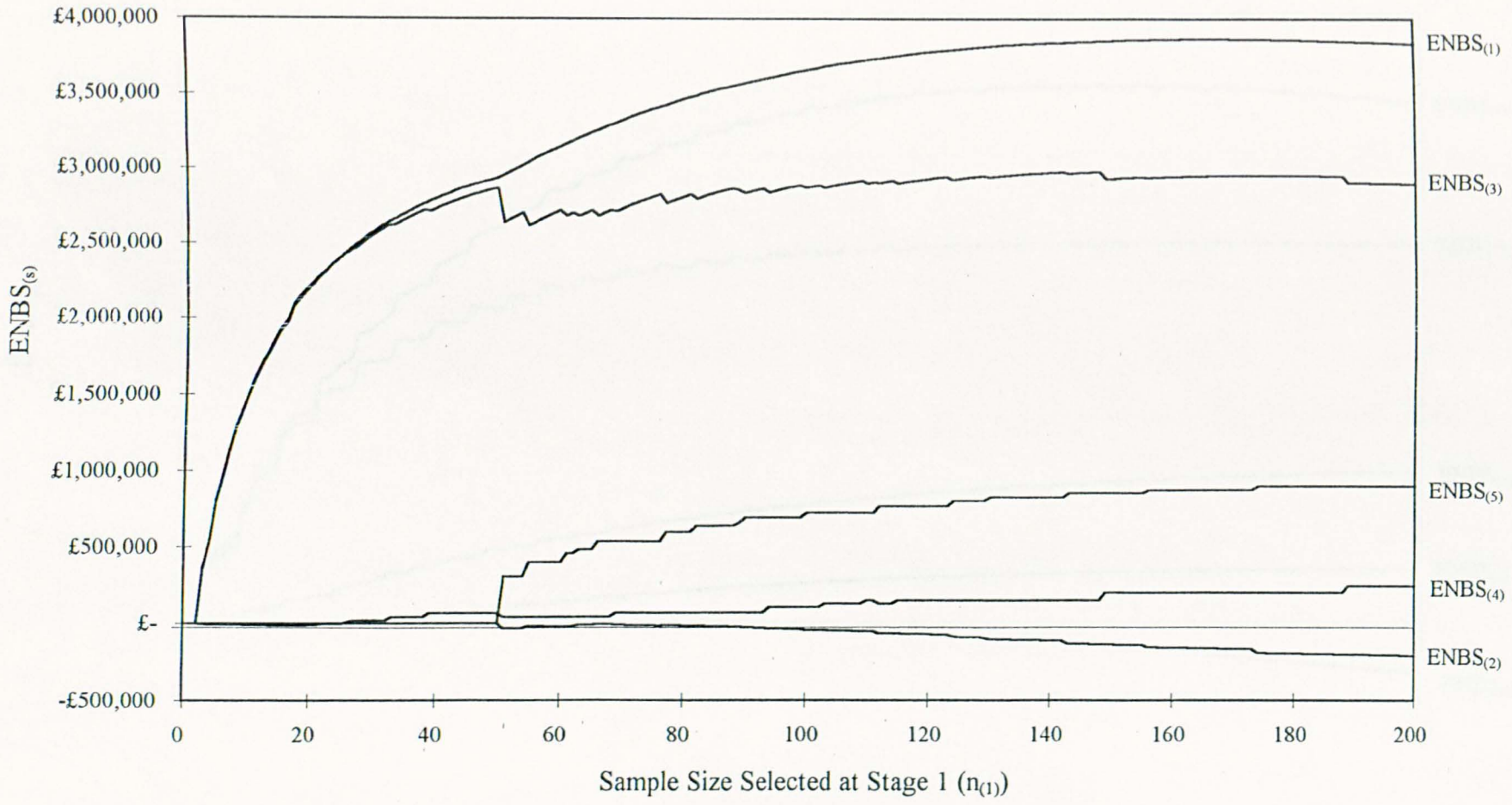


Figure 6.4.3b ENBS with Fixed Allocation (1/g=£4,000)

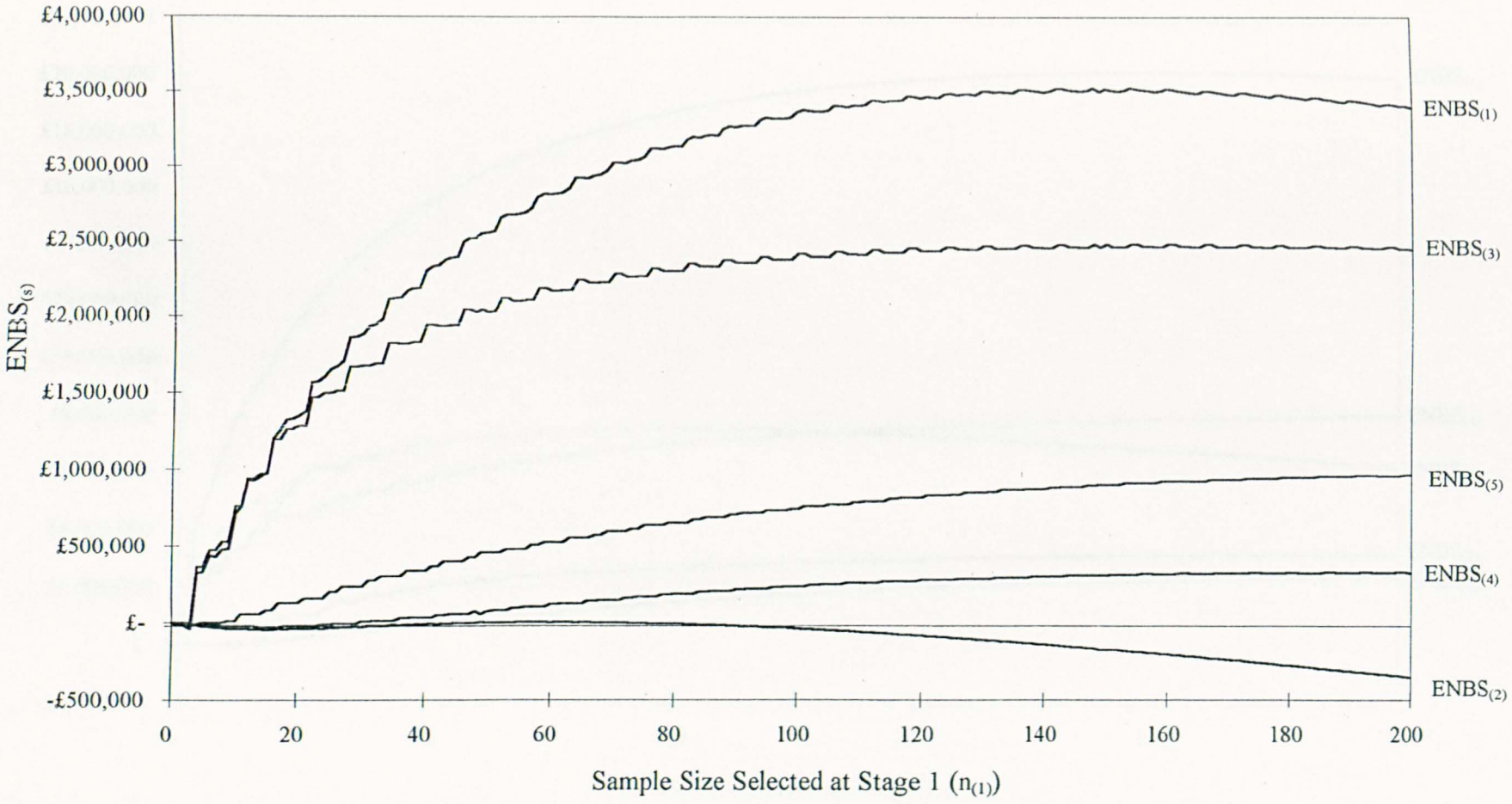


Figure 6.4.4a ENBS with Optimal Allocation (1/g=£10,000)

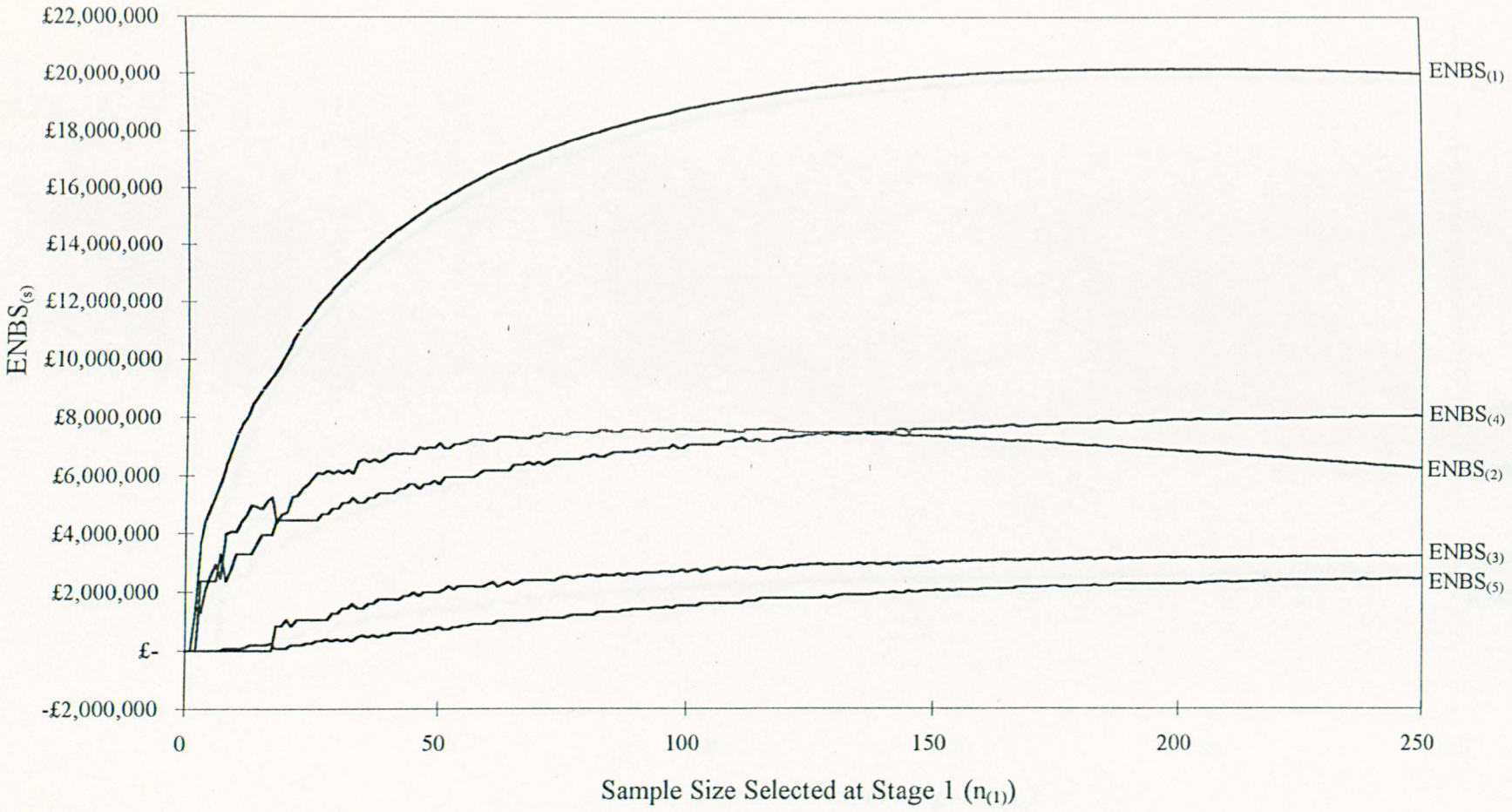


Figure 6.4.4b ENBS with Fixed Allocation (1/g=£10,000)

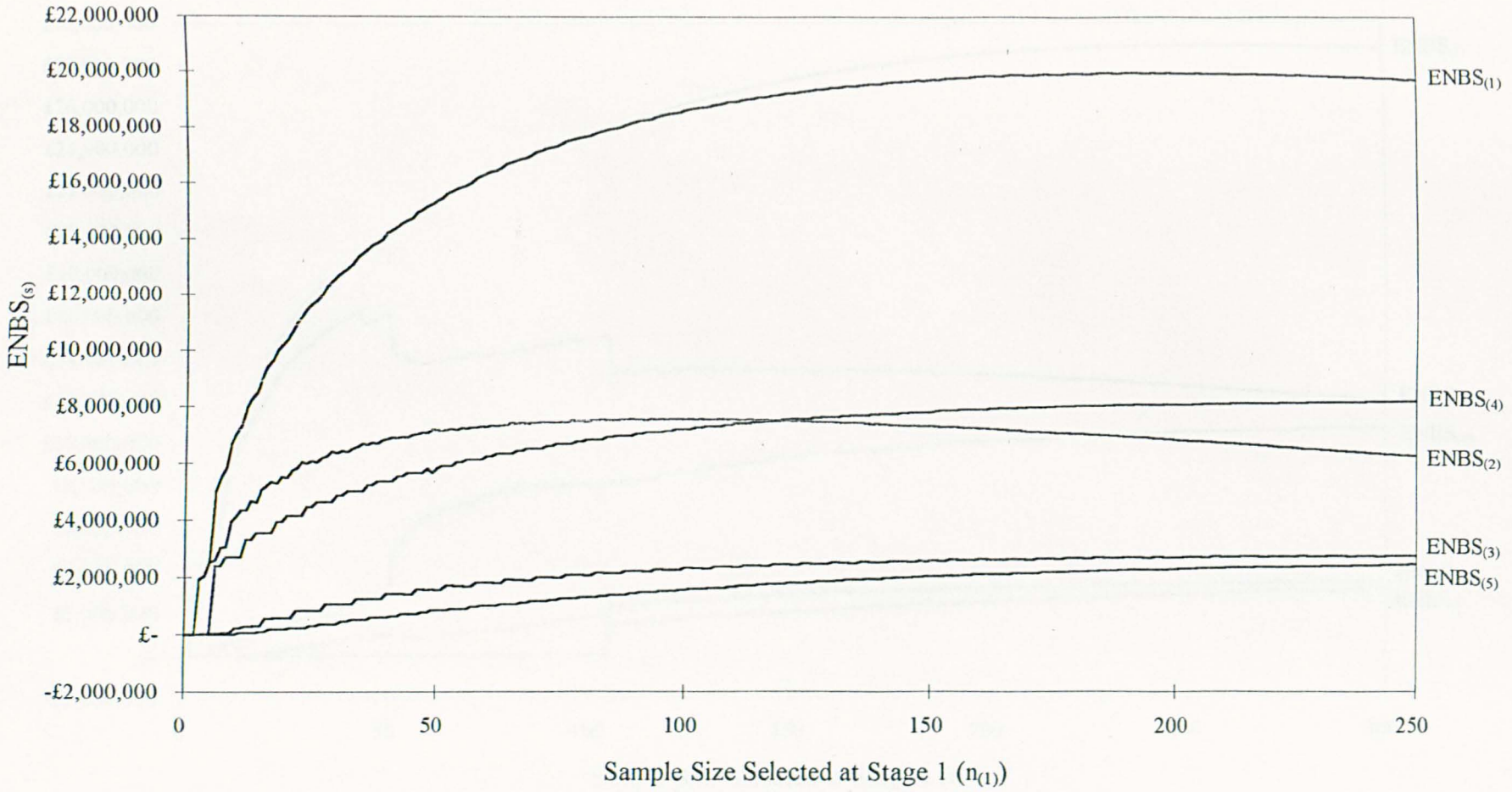


Figure 6.4.5a ENBS with Optimal Allocation (1/g=£20,000)

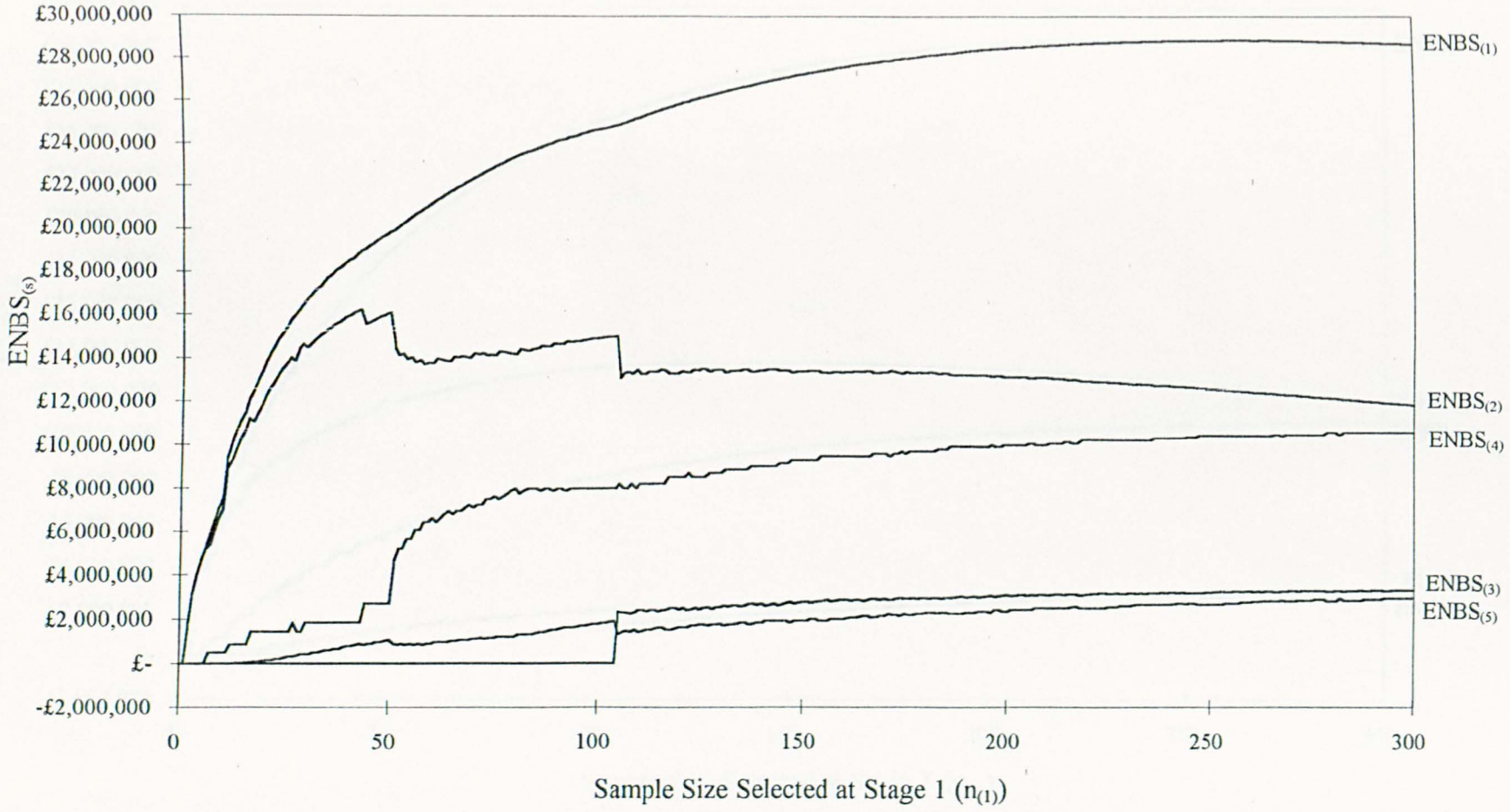


Figure 6.4.5b ENBS with Fixed Allocation (1/g =£20,000)

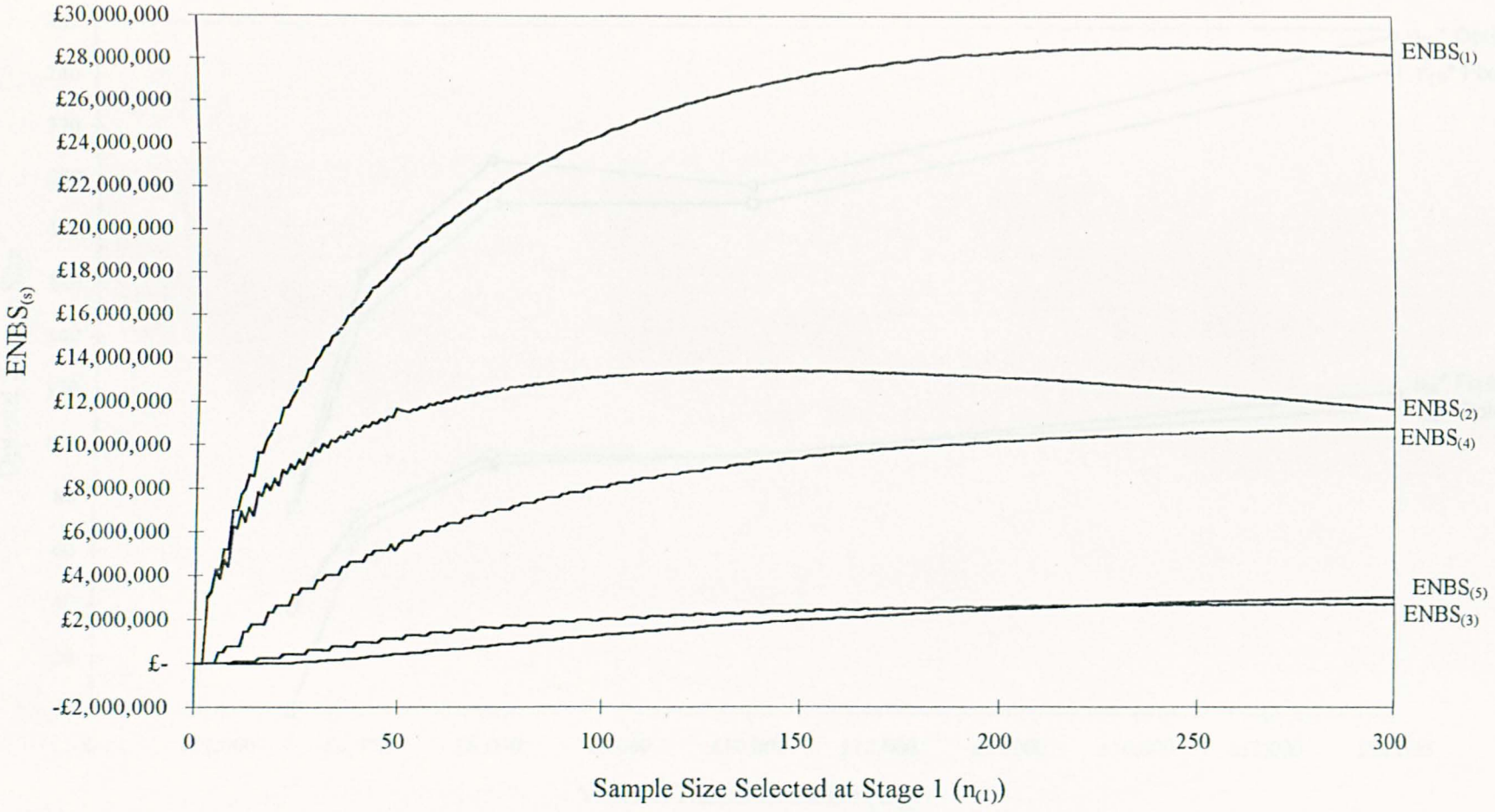


Figure 6.4.6 Optimal Sample Size with Fixed and Optimal Allocation

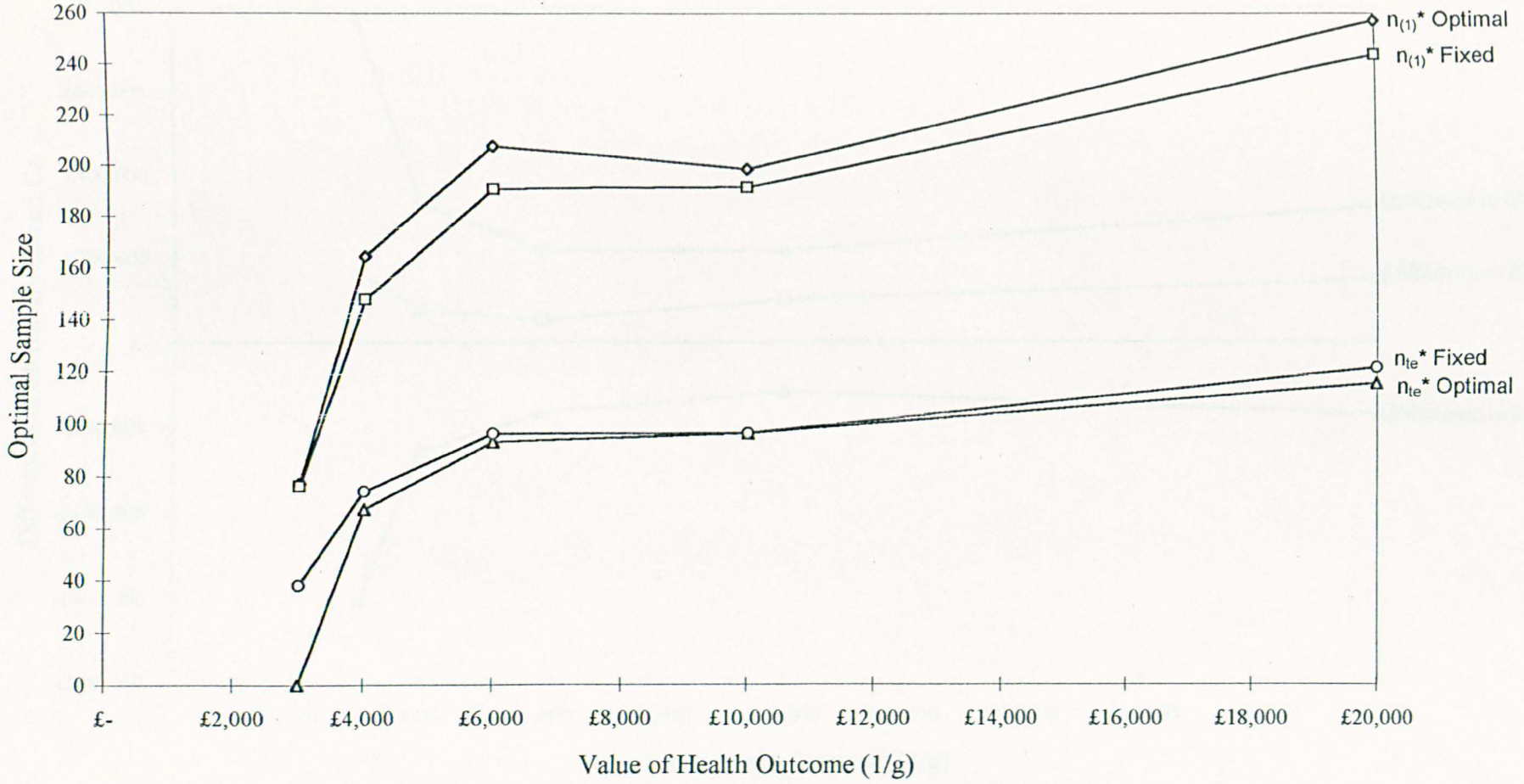


Figure 6.4.7 Difference Between Optimal and Fixed Allocation

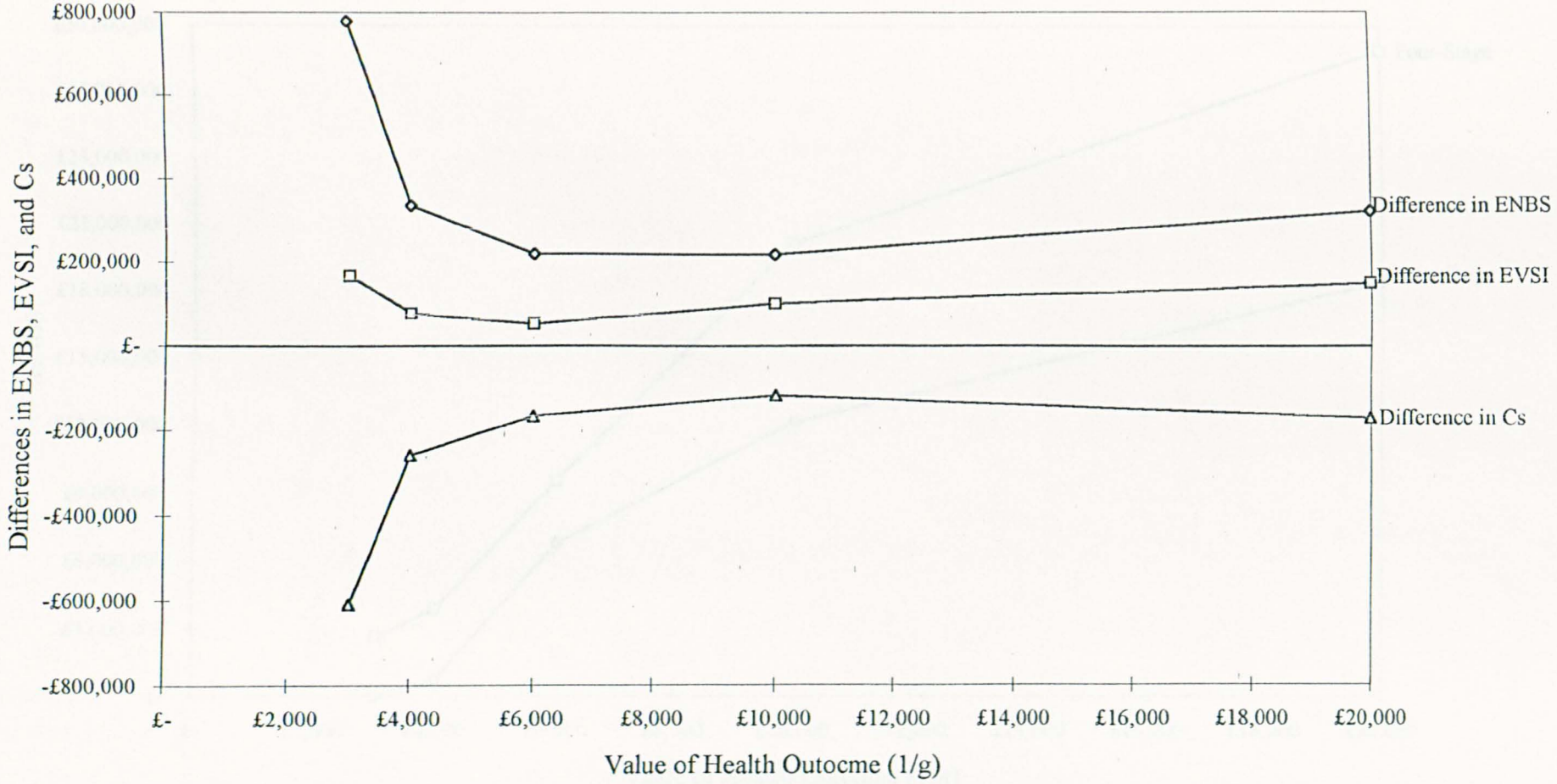


Figure 6.4.8a ENBS for the Four and Two-Stage Decision Problem

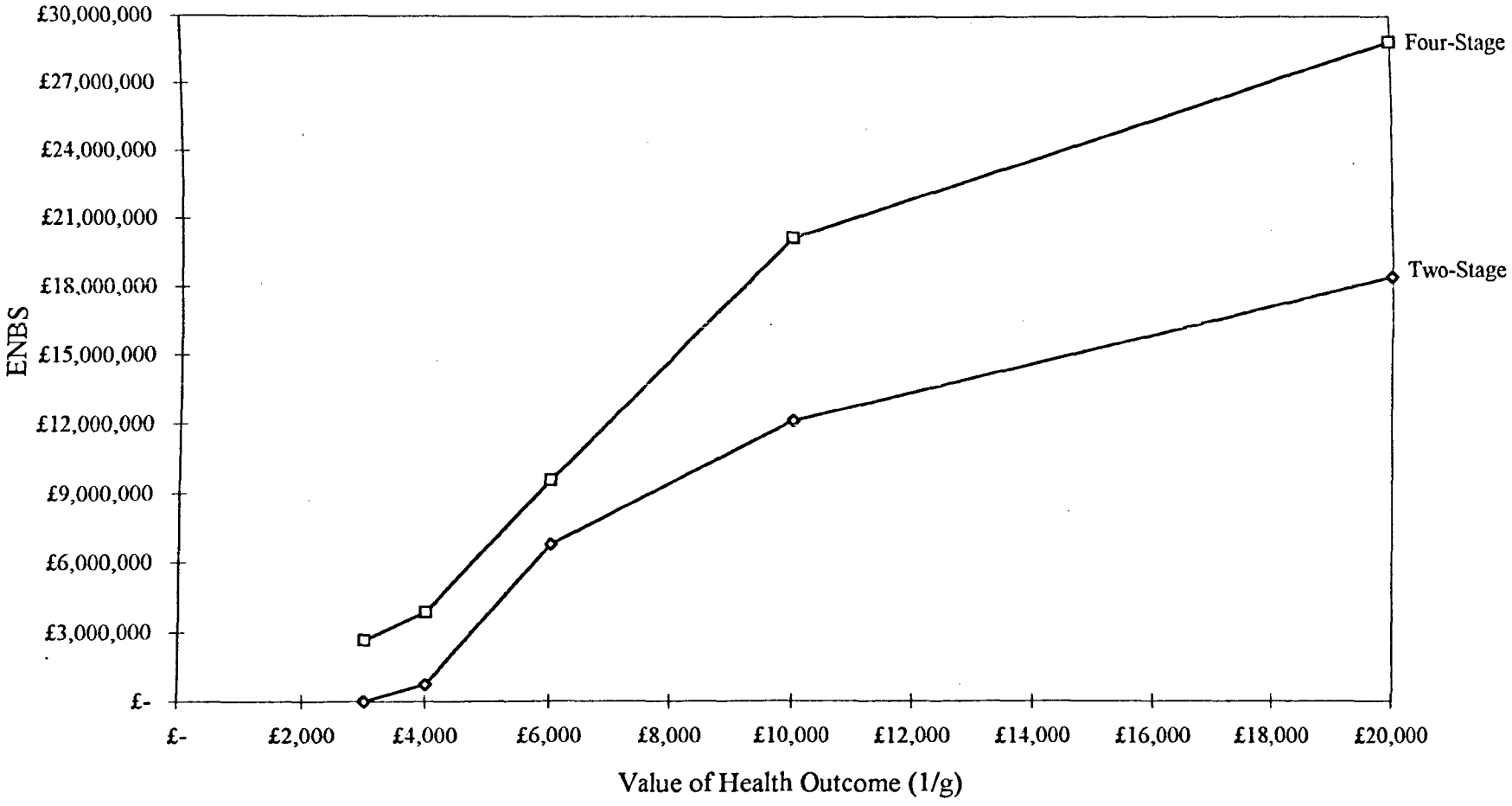
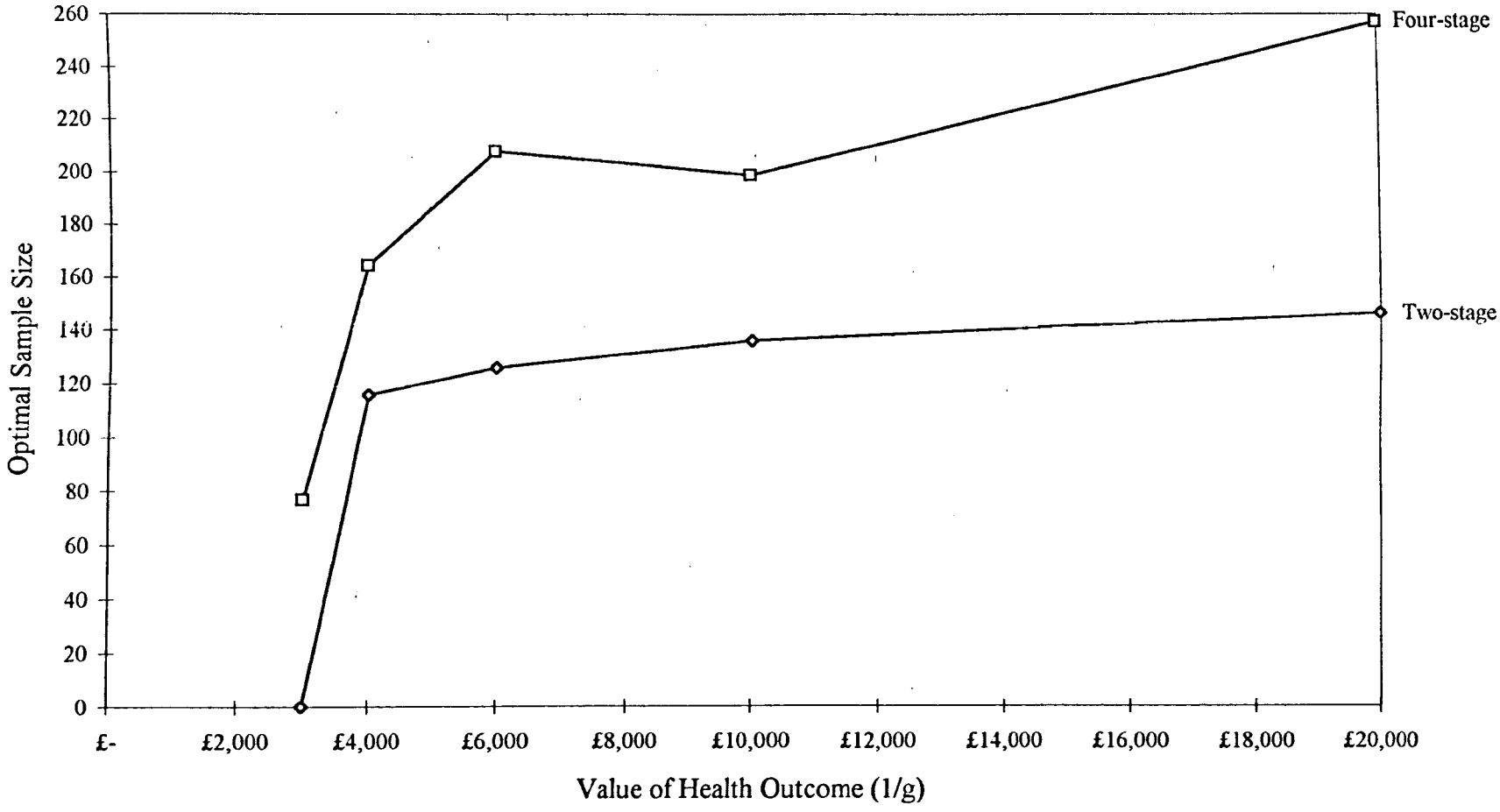


Figure 6.4.8b Optimal Sample Size for the Four and Two-Stage Decision Problems



Chapter 7 Conclusions

Contents: Page

7.1 Introduction

The thesis has considered two aspects of the value of clinical information. The value of information provided by diagnostic technology was considered in chapters two and three. The value of information generated by clinical research was examined in chapters four, five and six.

Background

This thesis was developed in response to the methodological problems which were encountered when faced with the economic evaluation of complex sequential clinical decision problems which include a number of treatment and diagnostic strategies [25]. The issues posed by this type of decision problem are: (a) can diagnostic information be valued without the prospective evaluation of all feasible strategies of patient management; (b) if not then is it worth collecting additional information about this decision problem through prospective research; (c) if it is then what is the optimal scale of this research; and (d) which of the many competing strategies of patient management should be included (regarded as relevant alternatives) in the evaluation. The methodological developments in the thesis are an attempt to address these practical problems which are the issues of allocative and technical efficiency in research and development.

Summary

The thesis aims to provide methods which can address these problems, and these are illustrated throughout the thesis using the same simple numerical examples which are introduced in chapters two and three. In chapter two a strategy for the evaluation of diagnostic information which could in principle avoid prospective evaluation is examined. In chapter three it is generalised to a more complex decision problem. It is argued that this strategy will only provide consistent valuations if critical assumptions which are unlikely to be met in most clinical

settings hold. The consequences of violating these assumptions are demonstrated using the same numerical example. It was found that the value of diagnostic information will be biased: a cost-effective technology many be rejected; and a technology which is not cost-effective may be accepted. It was also found that current clinical practice may not include those treatment strategies which will become optimal once the new technology is adopted. Therefore observing current practice may not be able to provide the information which would be required to evaluate the new technology. This posed a number of problems:

> "If valid inferences can not be based on observing current clinical practice, but the prospective evaluation of all possible alternatives in a sequential clinical decision problem is not possible, efficient, or ethical, then: (a) how should information of different quality from different sources be combined consistently and explicitly; (b) which clinical decision problems will be worth evaluating in a clinical trial; (c) if a clinical decision problem is worth evaluating which of the competing alternatives should be compared in a clinical trial; and (d) what is the optimal scale of this prospective research?"

These are the questions of how to establish both technical efficiency in research design, and how to achieve allocative efficiency in research and development across clinical decision problems and between research and service provision. It is these questions which were addressed in chapters four, five and six.

The analysis in chapter four used a decision analytic approach to the valuation of clinical information which combined a Bayesian view of probability with a framework for decision making. This approach was used to establish the cost of uncertainty surrounding the decision problem (the expected value of perfect information). The marginal benefit and the cost of acquiring sample information was then explicitly considered. This enabled the technically efficient scale of the research to be identified and the expected net benefit of proposed research to be

established. These methods were generalised from the single-stage decision problem considered in chapter four to the two- and four-stage sequential clinical decisions in chapter five. However, although the decision analytic approach taken in these chapters addressed the issues of which clinical decision problems are worth evaluating and what the technically efficient scale of the research should be, it does not allow the relevant alternatives to be identified.

In chapter six the fixed and equal allocation rule which assigned equal numbers of trial entrants to each of the alternative arms of the trial was relaxed. A dynamic programming approach was used to identify the optimal allocation of trial entrants at each stage of the decision problem. By explicitly considering the marginal benefit and marginal cost of assigning trial entrants to the alternative arms of the trial, the expected net benefit of the proposed research is higher than with fixed allocation rules. It enabled relevant alternatives to be identified because it is possible that no sample will be assigned to an arm of the trial and in this case it can be ruled out as an irrelevant alternative. At the end of chapter six the methodological problems which originally motivated the thesis have been addressed and methods proposed which can in principle provide a practical solution.

Methodological and Policy Issues

In the process of addressing these problems some interesting methodological issues have been highlighted. One of the implications from chapter three is that there are problems when using the traditional approach to priority-setting using league tables of cost-effectiveness ratios. Chapter four demonstrates that the traditional approach to clinical trial design is inconsistent with concepts of efficiency even when economic evaluations is conducted alongside a clinical trial. The decision analytic approach to the value of information shows that there are circumstances when it will not be efficient to conduct a clinical trial and clinical practice should be based only on prior information. Establishing the expected net

benefit of research also means ethical judgements about proposed research can be based on a consistent estimate of the opportunity cost of particular ethical concerns.

The thesis has also provided tools which can address some interesting policy questions, in particular methods for research priority-setting. Two hurdles are proposed for clinical research. The first ensures that only potentially cost-effective research is considered. The second ensures that this research will be cost-effective when conducted at the technically efficient scale. The expected net benefits of research can be used to establish allocative efficiency in research and development across clinical decision problems or broader areas of clinical research. Perhaps most importantly, it can be used to establish the optimal allocation of resources between research and development and service provision. Indeed what is clear from the analysis is that the value of information and research priorities cannot be separated from the budgetary constraints on service provision. These methodological issues and the policy implications of this work are discussed in more detail later in this chapter.

7.2 Consistency in the Evaluation of Diagnostic Information

The strategy for the economic evaluation of diagnostic information proposed by Phelps and Mushlin [107] was an attempt to combine information from a number of sources (although not explicitly taking account of the variable quality of this information) and to focus clinical research more sharply by: (a) eliminating new technologies which will not be cost-effective by constructing two hurdles that proposed technology must overcome; (b) by avoiding randomised experimental design where possible; and (c) by focusing on a clinically relevant range of test and patient characteristics.

The Phelps Mushlin strategy was applied to a simple numerical example of a two-stage test/treatment decision problem and two hurdles for the new technology were constructed. The first hurdle compared the expected value of the test assuming that it provided perfect information with an estimate of the cost of the new technology. Once the accuracy of the test is established existing information about current patient management is used to estimate the expected value of this imperfect clinical information. The relationship between the value of diagnostic information and the critical cost-effectiveness ratio (the shadow price of the budget constraint on service provision) was explored. This demonstrated that both hurdles are sensitive to this decision rule and that the explicit monetary valuation of health outcome is unavoidable in the valuation of clinical information.

The approach to the economic evaluation of diagnostic information has a number of advantages and attractions: it is consistent with economic decision rules and it enables research to be focused on new diagnostic technologies which are potentially cost-effective. However the presentation of the approach by Phelps and Mushlin, where treatment following diagnosis is determined only by the test results, implies that no other treatment alternatives are possible. The approach simply added a diagnostic device to existing patient management strategies and clinical practice is changed only to the extent that the test changes the probability

of assigning a patient to a particular diagnosis.

However in most clinical decision problems there is a range of treatment alternatives (and other diagnostic processes) which are at least possible following the results of the test, even if these alternatives are currently not used as part of existing patient management. In these circumstances diagnostic information may change patient management by changing the probability of assigning a patient to a particular diagnosis and by changing the optimal treatment choice.

To establish the circumstances in which this strategy will be appropriate to these less restrictive decision problems the Phelps Mushlin strategy was generalised in chapter three to accommodate more than one treatment for a given diagnosis, and applied to a simple numerical example of a four-stage decision problem. It was argued in chapter three that the Phelps and Mushlin strategy depends critically on two assumptions: Firstly it is assumed that the decision problem facing the clinician prior to the introduction of the test must be identical to the decision problem when the test is introduced and the test results are known. Secondly it accepts current practice as an appropriate baseline (or relevant alternative) to evaluate a new diagnostic device and implicitly assumes that the existing strategies of patient management are correct. In the context of an economic evaluation this means that existing strategies of patient management are efficient (the most cost-effective) at the critical cost-effectiveness ratio. Current practice will only be the relevant alternative if there is consistency between the value of health outcome which is implicit in the selection of current practice (1/g) and the value of the critical cost-effectiveness ratio (CCER), which is the shadow price of the budget constraint

This assumption is unlikely to hold because the appropriate critical ratio is uncertain and depends on which budget is regarded as relevant. It was argued that the value of health outcome implicit in existing patient management may be greater than the CCER because clinicians may only consider clinical effectiveness

or may not have full information about the budget constraint they face and the costs of competing programmes within the budget.

The consequences of violating this assumption were demonstrated in chapter three and two types of error were identified at each of the hurdles. The first type of error occurs when the value of diagnostic information is overestimated and a diagnostic technology which is not cost-effective may be accepted. If both the fallback and the test/treatment strategy are selected using an inconsistent implicit value of health outcome then the value of clinical information is underestimated and a potentially cost-effective diagnostic technology may be rejected. Also a less effective but less costly alternative treatment may exist which would be optimal at the CCER but may not be part of current practice. In these circumstances it can no longer be assumed that treatment alternatives which are optimal following the new test will be part of existing patient management. It will not be possible to estimate the EVPI or the EVCI based on existing information and the investigator may be forced to consider an experimental design which includes both the test and the subsequent treatment strategies.

The analysis of the Phelps Mushlin strategy in chapter three suggests it is likely to fail when applied to more complex decision problems because: (a) the key assumption of consistency between the value of health outcome implicit in current practice and the CCER is unlikely to hold; (b) when this assumption is violated the values of both the first and second hurdles will be biased; and (c) it is not necessarily the case that information about current practice before the test is introduced will be sufficient to construct the first hurdle. The Phelps Mushlin strategy fails, but if the prospective evaluation of all possible alternatives in a sequential clinical decision problem is not possible, efficient or ethical, then this poses the questions of how to establish both technical efficiency in research design, and achieve allocative efficiency in research across clinical decision problems and between research and service provision. The analysis in chapter three also poses some methodological problems for constructing league tables of

cost-effectiveness ratios to set priorities in service provision.

Setting Priorities in Service Provision

The decision problem considered in chapter three includes six possible strategies of patient management and can generate four cost-effectiveness ratios, none of which can be ruled out as extendedly dominated [78, 140]. This poses a problem for the traditional approach to priority-setting and decision making using cost-effectiveness or cost utility analysis. The traditional approach would be to place the cost-effectiveness ratio for this new diagnostic test in a league table along with the cost-effectiveness ratios of other non-mutually exclusive alternatives competing for the same budget. The social decision-maker should implement each in turn until the budget is exhausted [47, 141]. The value of health outcome is set implicity, and will be the cost-effectiveness ratio of the marginal project. In this traditional approach health outcome does not need to be valued explicitly prior to an economic evaluation, and league tables of cost-effectiveness ratios which allow decision-makers to determine the valuation of health outcome implicitly according to their budget assume that the cost-effectiveness ratios are independent of the value of health outcome.

However the example in this chapter shows that when considering sequential decisions problems there is no unique cost-effectiveness ratio and there are a number of ratios which could be placed in a league table [15, 16]. The relevant ratio depends on the valuation of health outcome, so a league table for a particular value of 1/g could be constructed with a unique ratio for each intervention. Howwever, in this case the league table as an aid to decision-making is redundant because 1/g would already have been selected and the projects which should be accepted have already been determined.

One approach to this issue is essentially to ignore the problem, and this is embodied in much of the work on league tables and published cost-effectiveness

ratios [47]. This approach accepts cost-effectiveness ratios which compare alternatives dictated by current clinical practice using some implicit rule, but as we have already seen this will lead to inconsistencies and inefficient allocations. Each of the cost-effectiveness ratios which have been included in published league tables must be based on some implicit decision rule used to identify which alternatives are relevant. There is no reason to believe that these implicit rules are consistent with the shadow price of the budget constraint.

An alternative approach is to evaluate every possible strategy. Sequential clinical decision problems are a comparison of a number of mutually exclusive alternatives. By evaluating each they can be ranked by effectiveness and those alternatives which are dominated and extendedly dominated can be ruled out [78, 140]. Those that remain can be used to generate cost-effectiveness ratios implied by moving to more effective but more costly strategies. These ratios can be placed in a section of a league table and as the budget increases a more effective strategy will be accepted and the less effective strategy will be rejected. Although this approach does provide a theoretical solution if the assumptions of constant returns, non repeatability, and perfect divisibility are accepted it does require that all possible alternatives should be evaluated. In this simple decision problem this would involve the evaluation of six rather than two possible strategies but in more complex decision problems it could involve a very large number of alternatives. Adopting this approach would involve the prospective evaluation of each, a proposal which may well be inefficient, or infeasible (in terms of recruitment into such a trial) and would probably be regarded as unethical.

Health outcome must be valued explicitly prior to an economic evaluation so that relevant alternatives can be selected in a way which is consistent with the decision rules which will be used when the cost-effectiveness analysis is complete. However this also implies that the decision rule must capture all the relevant decision criteria. In this example the value of health outcome is the only criteria, but if social decision-makers [129] wish to include other criteria such as equity and

access these must be included prior to evaluation so they can be used to select the relevant alternatives which are compared. If different criteria are used to decide whether the project should be implemented then there may be other alternatives previously rejected which would meet these new criteria more effectively, and in these circumstances it becomes difficult to separate issues of equity and efficiency [33, 34, 35, 117].

7.3 A Decision-Analytic Approach to Trial Design and Research Priority Setting

The analysis in chapter three demonstrated that a clinical trial may be unavoidable even in the evaluation of a non-invasive diagnostic technology. This posed the problems of allocative efficiency across clinical research and technical efficiency in research design. In chapter four it was argued that the traditional approach to clinical trial design is inconsistent with concepts of efficiency, leads to either infinite or arbitrary sample sizes, and cannot address the issues of allocative or technical efficiency in clinical research. The methods developed in chapter four and five address these problems by using a decision-analytic approach which combines a Bayesian view of probability with a framework for decision-making.

The Traditional Approach to Trial Design

The problems encountered when running an economic evaluation alongside a clinical trial have been well documented. However, the traditional approach to the design of pragmatic clinical trials is inconsistent with concepts of efficiency, because an infinite value is implicitly placed on the benefits of sample information. Furthermore, the traditional approach does not directly address the decision problem faced by clinicians and cannot incorporate prior information explicitly and consistently.

In the traditional approach (assuming a fixed sample design, where all the results are available at the same time at the end of the trial) the key design issue is the number of patients to recruit. Sample size is very sensitive to the reference improvement and if the reference improvement is not well defined or is chosen in an arbitrary way, then sample size will also be arbitrary. The clinical reference improvement has been defined as the smallest worthwhile difference in effectiveness. Very small improvements in effectiveness should be worthwhile, but as reference improvement approaches zero, sample size tends to infinity. The

allocative and technical efficiency in research design.

A Decision-Analytic Approach to Trial Design

The decision-analytic approach developed in chapter four and five combines a Bayesian view of probability with a framework for decision-making. The approach abandons traditional significance testing, confidence intervals and their Bayesian counterparts in favour of minimising the expected costs of making the wrong decision.

The methods developed in these chapters address the problem of allocative and technical efficiency in research design by constructing two hurdles that proposed research must overcome before it can be considered cost-effective. The first hurdle asks if the cost of proposed research exceeds the maximum possible benefits (the expected cost of uncertainty). If the cost does not exceed the maximum benefit then it is potentially cost-effective. Whether the proposed research is cost-effective can be established by constructing the second hurdle which explicitly considers the marginal cost and marginal benefits of sample information. The second hurdle ensures that the research is conducted at the technically efficient scale and provides a measure of the expected net benefit of the proposed research. This approach was illustrated by application to the simple single-stage fallback treatment decision in chapter four before it was generalised to the more complex two and four-stage test/treatment decisions in chapter five.

7.3.1 Hurdle I: The Expected Value of Perfect Information

The expected value of perfect information (EVPI) was established and this forms the first hurdle for proposed research. It is the maximum benefit that could be provided by additional information and the maximum return to research effort. This gives a method for focusing research priorities because it can be used to

identify those clinical decision problems (or areas of clinical research) where the costs of uncertainty are highest and where the information from research will be most valuable. If the fixed costs of research are known, the EVPI can be used to eliminate proposals (where the costs exceed the EVPI) which will not be cost-effective. The EVPI can also be used in the same way to identify priority areas for scientific reviews and Meta-analysis: clinical decision problems where the costs of uncertainty are greatest derive the most benefit from a review of existing research. The EVPI is a powerful tool for identifying research priorities in support of a move towards evidence-based medicine. Indeed this approach can set the limits to evidence-based medicine and provide a framework within which it can be applied consistently.

The expected value of perfect information can be established based only on prior information, including evidence from previous intervention and observational studies, but it can also include expert judgements. The decision-analytic framework focuses attention on those variables where evidence or judgement is required and by making prior information and judgements explicit they are open to empirical falsification. This is not necessarily the case in input/output models for assessing payback in clinical research or in Delphic studies of research foresight which elicit preferences which are not open to criticism or empirical testing. The quality of the prior information is reflected in the prior variance and prior information can be regarded as a quasi-sample where a smaller sample size indicates a more sceptical prior. This framework makes the prior information which is required explicit and allows evidence from a variety of sources to be combined and handled consistently using Bayes Theorem.

The relationship between the EVPI and the value of 1/g was examined for the single, two and four-stage decision problems and this demonstrated that the value of information is crucially dependent on the value of health outcome used to set priorities in service provision. This is because the value of 1/g is determined by the budget constraint faced by clinical practitioners. If the budget constraint is

relaxed then more costly but effective health services can be provided, the cost-effectiveness ratio of the marginal service will increase, and the EVPI will rise. If the budget is tightened the cost-effectiveness ratio of the marginal project will fall, service providers will be unable to take advantage of the information provided by clinical research, and the value placed on this information will diminish. The value of information, research priorities and the optimal level of research and development expenditure are all dependent on the budgetary constraint on the provision of health services.

The relationship between the EVPI and the quality of (or confidence in) the prior information was also explored and this showed that when there is less confidence in the prior information the EVPI is higher because there is more uncertainty surrounding a decision based only on prior information. Similarly when the prior is less sceptical the decision will be less uncertain and the EVPI is lower. The point at which the clinician would be indifferent between the alternative strategies based on prior information is where she is most uncertain and at this point the EVPI reaches a maximum.

Sequential Decision Problems

The analysis of the sequential decision problems in chapter 5 also demonstrated that the cost of uncertainty for a clinical decision problem will be underestimated if some alternatives are ruled out as not relevant and a sequential decision problem is simplified to a single-stage problem. This is because the EVPI for the whole decision problem is the sum of EVPI at each of the contingent decisions and at the initial decision. An analysis which simplified a sequential clinical problem to a single-stage problem would underestimate the EVPI because the process of simplification excludes some alternatives which are feasible and relevant and in certain circumstances could become the preferred strategy.

Chapter five demonstrated that by calculating the EVPI at each stage of a

sequential decision problem, those points in the sequence of decisions where the cost of uncertainty is highest can be identified. This is not necessarily the case with conventional sensitivity analysis because the prior distributions for the key variables and the value placed on opportunity losses at sensitive decisions are not necessarily taken into account. It was shown that simple measures of sensitivity may be misleading particularly if they are used to identify those points where information may be most valuable

7.3.2 Hurdle II: The Expected Net Benefit of Research

Proposed research which passes the first hurdle can be regarded as potentially cost-effective. To demonstrate that it will be cost-effective the optimal scale of the research (in this case sample size) was established. The expected benefit of sample information was measured by the reduction in expected opportunity loss. This can be calculated for a particular sample size based on prior information and an estimate of the sample variance of the incremental net benefits. The marginal cost of sampling includes the additional cost of treatment when patients entering the trial are allocated to the experimental treatment. In chapters four and five patients are allocated equally between the control and experimental arms of the trial. The expected net benefit of sampling (ENBS) was defined as the difference between the total benefit and the total variable cost for a particular sample size. Sample size is optimal when ENBS is positive and at a maximum. The ENBS is the expected net present value of research and can be used to prioritise research proposals. If this is positive then the research passes the second hurdle and is cost-effective when conducted at the technically efficient scale.

The ENBS also represents the opportunity cost of failing to implement cost-effective proposals. If a proposal with positive ENBS was rejected on the grounds of medical ethics then the implicit opportunity cost of this ethical position can be established either in monetary terms or in terms of health benefits forgone.

Consideration of medical ethics is an essential element in trial design, but this approach makes it possible to estimate the opportunity cost of particular concerns for the individuals involved in a trial. In this way the trade-off between individual and collective ethics can be made explicit. If these trade-offs are explicit they can be made consistently and be open to criticism and debate.

The analysis in chapters four and five explored the relationship between the value of health outcome and the ENBS and showed that the value of sample information and the technically efficient scale of clinical research is dependent on the budgetary constraints on service provision. The issues of allocative and technical efficiency cannot be addressed before health outcome has been valued in monetary terms. The relationship between the optimal sample size, the expected net benefits of research and the quality of prior information was also examined. This showed that when the value of health outcome is low and the prior is less sceptical the optimal sample size is zero and the decision should be based only on prior information. In these circumstances the prior decision will be to reject the experimental treatment.

This analysis also showed that when the prior is less sceptical and the value of health outcome is high then the optimal sample size will also be zero. But the prior decision is now to treat using the experimental treatment, suggesting that there may be circumstances in which a new treatment should be adopted without gathering sample information through a clinical trial. This demonstrates that a decision-analytic approach can be used to set rational limits to evidence-based medicine and to provide a framework where new treatments of potentially great benefit can be adopted without incurring the cost (including the opportunity cost of the delay before the results are available) of a clinical trial.

Sequential Decision Problems

The analysis in chapter five demonstrated that the expected net benefits of sample

information and the optimal sample size can be established and the second hurdle can be constructed for sequential decision problems. This also showed that an analysis which did not recognise the sequential nature of a decision problem by simplifying it to single-stage decision and excluding the diagnostic process would bias efficient research design and cause errors at the second hurdle. Similarly an analysis which simplified the four-stage decision problem to the two-stage problem by excluding one of the treatment alternatives will underestimate the ENBS and bias the optimal sample size. Once again this illustrates that arbitrarily excluding feasible alternatives and using implicit rules and inconsistent judgements to identify which alternatives are regarded as relevant will bias research design. In this example the value of information and the technically efficient scale of proposed research will be underestimated. There is a danger that cost-effective proposals will be rejected and those that are accepted will be conducted at less than the technically efficient scale.

7.3.3 Setting Priorities in Research and Development

The decision-analytic approach which was developed in chapters four and five can provide practical policy tools for research priority-setting. The information generated by clinical research is valued in a way which is consistent with concepts of efficiency and with the methods used to set priorities in service provision. The prior information, which is implicit in the traditional approach, is identified and handled consistently so that it is open to criticism, alternative formulation and empirical testing.

The simple numerical example examples of a fixed sample pragmatic trial shows that these techniques can be used to identify areas of clinical practice where the cost of uncertainty is high and where the potential benefits of clinical research will also be high. Estimates of the EVPI and the ENBS can be used to construct two hurdles which proposed research must overcome before it can be considered cost-

effective. The first hurdle is based only on prior information and asks if the EVPI (the cost of uncertainty or the maximum value of sample information) is greater than the costs of the proposed research. This hurdle can eliminate proposed research which will not be cost-effective before issues of research design must be addressed. Those proposals which pass the first hurdle can be regarded as potentially cost-effective and can be considered at the second hurdle. The second hurdle ensures that the design of potentially cost-effective research is technically efficient and that it will be cost-effective when conducted at the optimal scale. The ENBS represents the value of the proposed research to the providers and consumers of health services. It also represents the opportunity cost of rejecting cost-effective research proposals. This approach provides a means to decide which clinical decision problems are worth evaluating in a clinical trial and what is the technically efficient scale of this research.

Estimates of the expected net benefit of research can be used to rank competing research proposals, and by implementing proposals with higher net benefit first, the maximum health benefits can be gained for limited research and development resources. If all proposals with positive net benefits could be implemented then the returns to research and development expenditure would be at a maximum. The net benefit provided by the marginal research proposal will be zero, and the level of expenditure would be optimal. At this point research and development should only be expanded if there is a corresponding expansion in health service provision.

These tools can be used to address the optimal allocation of resource between research and development and service provision. The optimal level of research and development expenditure is determined by the budgetary constraints on service provision, because it is the cost-effectiveness ratio of the marginal service which determines the value placed on the benefits of clinical research. The level of research funding would be less than optimal if the net benefit of the marginal research proposal was positive. Expenditure on research and development

should be increased, and in these circumstances health benefits to patients would improve if resources were transferred from service provision to research and development. This would reduce 1/g as resources were transferred from service provision and reduce the value of proposed research. The optimal level of research and development expenditure would fall and converge on the efficient allocation between service provision and research and development.

The estimates of the expected net benefit of proposed research can be used to allocate research resources between broad areas of clinical research. If there are areas of clinical research where the marginal expected net benefit is low then resources should be transferred to other areas of research where the marginal net benefits are higher. This transfer of resources would increase the health gains for patients within a fixed research and development budget and should continue until the marginal net benefit across all areas is the same and the share of research resources is optimal. This approach could also inform policy where there is joint commissioning of research in a clinical area. If there is evidence that the current level of research funding is less than optimal this will provide a framework for negotiation between commissioning agencies. If there is evidence that the expected net benefit of marginal research commissioned by one agency is higher than other agencies this suggest that the former should increase their share of research effort in this clinical area. These techniques provide a framework within which these broad policy issues can be discussed, although further work would be required implement this approach because it has implications for those who commission, design and use clinical research.

7.4 A Dynamic Programming Approach to Optimal Patient Allocation

The approach taken in chapters four and five solved, at least in principle, all but one of the problems which motivated this thesis: the Bayesian view of probability can explicitly incorporate prior information from different sources and of different quality; the first and second hurdles can identify which clinical decision problems should be considered for prospective clinical research; and the construction of the second hurdle identifies the efficient scale of this research.

However there remains the problem of which of a number of competing alternatives should be compared within a clinical trial. This was not addressed in chapter four or five because that analysis assumed an fixed and equal allocation of trial entrants at each stage. This means that a sample is allocated to each arm of a trial irrespective of the cost and benefit. This arbitrary rule forces part of the sample to be allocated to each alternative and does not provide a method to identify which of the alternatives are irrelevant. In chapter six this arbitrary allocation rule was relaxed and a simple dynamic programming approach was used to establish optimal patient allocation and provide an explicit and consistent method to identify relevant alternatives which should be compared in the trial.

The Traditional Approach

The equal allocation of patients between experimental and control arms of a trial is often used and is implicitly justified by assuming that the variance of the outcome of interest for the control and experimental arm is the same, so that the benefits (reduction in sample variance) of assigning an additional trial entrant to either arm of the trial will be the same. However there is little justification for this rule of precedent when the costs and benefits of allocating a trial entrant to the alternative arms of the trial are explicitly considered. Whether an additional trial entrant should be allocated to a particular arm of a trial should be determined by the

marginal benefit of assigning the patient to that arm (which will be determined by the variance of the net benefits of that arm) and the marginal costs of assigning the patient to that arm (which is determined by the additional treatment costs). The only circumstances in which an equal allocation could be justified would be when the variance and the marginal sampling costs of both arms of the trial are the same. In the examples considered the variance of the net benefits of the alternatives at each stage are not assumed to be the same and the marginal cost of assigning a trial entrant to either of the alternatives at each stage will differ.

There are established allocation methods which are only concerned with clinical outcomes in sequential clinical trials, were the accumulated results from earlier participants in the trial are available and are used to allocate those entering the trial. Chapter six addressed a more fundamental problem of optimal allocation in a fixed sample design where sample information is only available at the end of the trial. The benefit and cost of sample information, optimal sample size and the allocation of patients at each stage must be established before any sample information is available.

A Dynamic Programming Approach

The numerical examples in chapter six demonstrate that the optimal allocation of trial entrants is not simply an issue for sequential clinical trials but is a more fundamental problem which is also relevant to the fixed sample designs considered here. The approach to optimal patient allocation taken in this chapter explicitly considered the marginal cost and the marginal benefit of assigning trial entrants to the alternatives arms of the trial at each stage. The single-stage decision problem was solved in two stages and required the full enumeration of all feasible allocations of each sample considered. However the full enumeration of all feasible allocations in more complex sequential decision problems is not tractable and the two and four-stage decision problems were solved using three and five-stage dynamic programmes. This simple dynamic programming approach uses the

recursive relationship between the payoffs at each stage of the decision problem and provides a practical solution to the problem. This may not be the optimal solution (because although there is a recursive relationship between the payoffs at each stage the prior variance and incremental net benefit at earlier stages are partly determined by the contingent allocation at later stages), but the dynamic programming approach is the only feasible solution and clearly provides a better solution than the arbitrary fixed allocation rule used in chapter five.

The optimal allocation of trial entrants will increase the expected net benefit of sampling and the value of the second hurdle. This was achieved by a reduction in the cost of sampling (possibly with a reduction in the expected benefits of sampling) or alternatively by an increase in the expected benefits of sampling (possibly with an increase in the cost of sampling). The analysis in chapter six showed that the optimal allocation at each stage, optimal sample size, and the expected net benefits of sampling are determined by the value of health outcome because this determines the relative weight placed on the additional benefits and costs of assigning entrants to each arm of the trial. The numerical examples considered in this chapter demonstrate that arbitrary and fixed allocation rules are inefficient and will lead to an underestimate of the value of proposed research and there is a danger that research which should be accepted at the second hurdle will be rejected. Research which is accepted at the second hurdle despite an arbitrary allocation will be designed inefficiently and may include the comparison of irrelevant alternatives. The value of the research will be underestimated.

The simple dynamic programming approach to optimal patient allocation provides an explicit and consistent method to identify relevant alternatives which should be compared in the trial. Optimal allocation provides a method to rule out irrelevant alternatives consistently based on an assessment of the expected benefit and cost of comparing the alternatives, because it allows no sample to be allocated to an alterative at each stage. If it is optimal not to allocate a sample to an alternative then it can be regarded as irrelevant and can be excluded from prospective

research.

These numerical examples demonstrate that what are relevant alternatives depends on the value of 1/g and alternatives cannot be ruled out as irrelevant before the shadow price of the budget constraint (the monetary value of health outcome) has been established. They also illustrate the danger of ruling out alternatives based on implicit and inconsistent decision rules. An analysis which simplified the two-stage decision problem to the single-stage treatment decision would underestimated the value of the research and seriously bias the trial design. Similarly an analysis which simplified the four-stage problem to the two-stage problem, by excluding one of the treatment strategies, would underestimate the ENBS and the optimal scale of the research at the second hurdle. This supports an argument which has been made in previous chapters: that the selection of relevant alternatives and the exclusion of some alternatives based on judgements and implicit decision rules can seriously bias the analysis and can lead to errors at the second hurdle and inefficient research design.

7.5 Further Developments

The methods which have been developed and illustrated in each chapter have solved, at least in principle, the problems which originally motivated this thesis: the Bayesian view of probability can explicitly incorporate prior information from different sources and of different quality; the first and second hurdles can identify which clinical decision problems should be considered for prospective clinical research; the construction of the second hurdle identifies the efficient scale of this research; and the dynamic programming approach to optimal patient allocation can identify which alternatives should be regarded as relevant and included in a clinical trial. However there are further methodological developments which could be pursued and the application of these methods to research priority setting suggests a programme of empirical work and poses the problem of how to implement this approach to research design and priority setting.

One way in which these methods could be extended would be to include the expected health benefit or cost to those enrolled in the trial. The analysis in the thesis has focused exclusively on collective ethical concerns by considering the expected benefit and cost of proposed research to future patients and society as a whole. This provides a measure of the opportunity costs of rejecting cost-effective research on the grounds of individual ethical concerns. The valuations of the potential health benefit and cost to those enrolled in the trial is excluded from this analysis and it is left to ethical decision-makers to make this trade-off, albeit with more information about the expected benefits of the research. However it would be possible to include the expected health benefit and cost to those enrolled in the trial, although this would involve making a value judgement about the relative weight attached to the costs and benefits to entrants as compared to the benefits which will accrue to future patients.

This thesis has been concerned with the value of information generated by pragmatic clinical research at the final stage (phase III or IV) of the development

of new technology. However it seems that in principle the same approach can be used to set priorities in the development of new technologies and pharmaceuticals at a much earlier stage of development. Indeed it may be possible to use the same approach to set priorities in fundamental and biomedical research. This could provide a useful method to identify new chemical entities which are most likely to be cost-effective and which will change clinical practice when fully developed. This could reduce the research and development cost of bringing successful technologies to the market and could also inform a regulatory framework to provide incentives for the development of technologies with these desirable characteristics. These techniques have provided a framework within which broad policy issues can be discussed, and in principle prospective empirical work which applied this approach to a sample of research proposals could provide evidence as to whether the current level of research and development is optimal and whether patients would benefit from a reallocation of resources between research and development and service provision.

Clearly the expected net benefits of research can only be realised if this approach can be successfully implemented. Implementing a decision-analytic approach to evaluative design and research priority setting has implications for those who commission, design and use clinical research. Those who commission research require guidelines for selectors to identify priority areas, to ensure that proposals are technically efficient and to establish the expected net benefits of proposed research. Those who design and seek support for clinical research need to be convinced of the value of these techniques and require bidding guidelines to ensure that research is designed efficiently and can demonstrate expected net benefits. Research interests would need to be focused on priority areas. This may be achieved by the incentives created as those who commission research give higher priority to funding research in these areas.

The impact on the users (clinical practitioners and purchasers) of clinical research ultimately determine whether the expected benefits of research are realised by

changing clinical practice. A key issue is the impact on decision-making of alternative approaches to evaluative design and research priority setting. One approach to the impact of research could be described as positive, and this attempts to establish how decision-makers make decisions and what type of information will have an impact on their practice. The approach accepts the current decision-making process and it implies that methods should be adopted which address decision-makers current concerns. This approach implicitly assumes that either the current decision-making process is optimal, with regard to the objectives of the NHS, or it is not amenable to change. This is the rationale for specifying a reference treatment difference in the traditional approach to evaluative design.

The decision-analytic approach takes what could be described as a normative and extra-welfarist [33, 34, 129] approach to the impact of research findings. This approach does not accept that the current decision-making process is necessarily optimal. It suggests how decisions should be made given specified objectives, prior evidence and the explicit assumptions and judgements that must be made. The aim is to establish methods and decision rules which best meet the specified objectives of social decision-makers (the objective which is embodied in the decision rules used throughout this thesis is the maximisation of health benefits). To change the existing decision-making process and persuade decision-makers of the issues which should be considered when interpreting alternative approaches to evaluative design, a campaign of dissemination, education and incentives is required. If they can be persuaded to abandon the traditional approach to significance testing and confidence intervals then this approach to the value of information could also be used to identify their own informational needs. They would be able to set their own priorities by establishing the cost of uncertainty or the EVPI of the decision problems that they face. This could be used to focus their efforts in searching the published literature, commissioning scientific reviews and acquiring new skills by recruiting personnel and purchasing training and consultancy.

List of References

1 Adams, ME. et al. Economic analysis in randomised control trials. *Medical Care* 1992; **30**: 231-243.

2 Altman, DG. and Bland, JM. Diagnostic tests 3: receiver operating characteristic plots. *British Medical Journal* 1994; **309**: 188.

3 Anderson, J. New approaches to evaluation in UK research funding. *Concept Paper* No 9. London: Science and Policy Support Group, 1989.

4 Anderson, J. Rogers, L. And Driscoll, M. Editors. International Comparisons of research outputs: cardiovascular research 1981-1991. *PRISM Report* No 2, 1994.

5 Anscombe, FJ. Sequential medical trials. *Journal of the American Statistical Association* 1963; **58**: 365-383.

6 Armitage, P. *Sequential Medical Trials*. Oxford: Blackwell, 1975.

7 Armitage, P. The search for optimality in clinical trials. *International Statistical Review* 1985; **53**:15-24.

8 Bather, JA. Randomised allocation of treatments in sequential experiments. *Journal of the Royal Statistical Society* B 1981; **43**: 265-292.

9 Bather, JA. On the allocation of treatment in sequential medical trials. *International Statistical Review* 1985; **53**: 1-13.

10 Bellman, RE. *Dynamic Programming*. Princeton University Press, 1957.

11 Bellman, RE. and Dreyfus, SE. *Applied Dynamic Programming*. Princeton University Press, 1962.

12 Berry, DA. A Bernoulli two-armed bandit. *Annals of Mathematical Statistics* 1972; **43**: 871-897.

13 Berry, DA. Modified two-armed bandit strategies for certain clinical trials. *Journal of the American Statistical Association* 1978; **73**: 339-345.

14 Berry, DA. A case for Bayesianism in clinical trials. *Statistics in Medicine* 1993; **12**: 1377-1393.

15 Birch, S. and Gafni, A. Cost-effectiveness/Utility Analyses: do current decision rules lead us to where we want to be? *Journal of Health Economics* 1992; **11**: 279-296.

16 Birch S, and Gafni A. Changing the Problem to Fit the Solution: Johannesson and Weinstein's (mis) application of economics to real world problems. *Journal of Health Economics* 1993; **12**: 469-476.

17 Boeckenholt, U. and Weber , E. Use of formal methods in medical decision making: a survey of and analysis. *Medical Decision Making* 1992; **12**: 298-306.

18 Briggs, AH. Sculpher, MJ. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health Economics* 1993; **3**: 95-104.

19 Briggs, AH. Sculpher, MJ. Sensitivity analysis in economic evaluation: a review of published studies. *Health Economics* 1995; **4**: 355-372.

20 Brown, BW. et al. Projection from previous studies: a Bayesian and frequentist compromise. *Controlled Clinical Trials* 1987; **8**: 29-44.

21 Bulpitt, CJ. and Fletcher, AE. Economic assessment in randomised controlled trials. *The Medical Journal of Australia* 1990; **153**: 16-19.

22 Buxton, M. and Hanney, S. Assessing payback from department of health research and development: preliminary report, Vol 1 and 2. *Health Economics Research Group Research Report* No 19, 1994.

23 Cain, KC. Greens, RA. and Begg, CB. Patient orientated performance measures of diagnostic tests. 3 U-factor. *Medical Decision Making* 1984; **4**: 33-46.

24 Centor, RM. Signal detect ability: the use of ROC curves and their analyses. *Medical Decision Making* 1991; **11**: 102-106

25 Claxton, K. and Posnett J. The role of decision analysis in structuring the economic evaluation of sequential clinical decision problems. University of York: *Discussion Papers in Economics* No 96/24, 1996.

26 Claxton, K. and Posnett J. An economic approach to clinical trial design and research priority setting. University of York: *Discussion Papers in Economics* No 96/19, 1996.

27 Claxton, K. and Posnett J. An economic approach to clinical trial design and research priority setting. *Health Economics* 1996; **5** (forthcoming)

28 Cochrane A. Effectiveness and efficiency, random reflections on the health serv ices. *Nuffield Provincial Hospitals Trust* 1972.

29 Colton, T. A model for selecting one of two medical treatments. *Journal of the American Statistical Association* 1963; **58**: 388-400.

30 Colton, T. A two stage model for selecting one of two treatments. *Biometrics* 1965; **21**: 169-180.

31 Cornfield, J. A Bayesian test of some classical hypotheses with application to sequential clinical trials. *Journal of the American Statistical Association* 1966; **61**: 577-594.

32 Coyle, D. Statistical analysis in pharmacoeconomic studies. *Pharmacoeconomics* 1996; **9**: 505-516.

33 Culyer, AJ. The normative economics of healtgh care finance and provision. In: McGuire, A. Fenn,P. and Mayhew, K. editors. *Providing health care: the economics of alternative systems of finance and provision*. Oxford: Oxford University Press, 1991.

34 Culyer, AJ. and Evans, RG. Mark Pauly on welfare economics: normative rabbits from positive hats. *Journal of Health Economics* 1996; **15**: 243-251.

35 Deber, R. and Goel V. Using explicit decision rules to manage issues of justice, risk and ethics in decision analysis: when is it not rational to maximize expected utility? *Medical Decision Making* 1990; **10**: 181-194.

36 Detsky, AS. Are clinical trials a cost-effective investment? *Journal of the American Medical Association* 1989; **262**: 1795-1800.

37 Doubilet, P. Begg, CB Weinstein MC. et al. Probabilistic sensitivity analysis using Monte Carlo simulation. *Medical Decision Making* 1985; **5**: 157-177.

38 Dowie, J. Decision analysis: the ethical approach to medical decision making. In: Gillon, R. editor. *Principles of health care ethics*. Chichester: Wiley, 1994.

39 Dowie, J. Clinical Decision Analysis: background and introduction. In: Llewelyn H. And Hopkins (editors). *Analysing how we reach clinical decisions*. London, Royal Collage of Physicians of London 1993.

40 Dowie J. Evidence-based, cost-effective and preference-driven medicine: decision analysis based medical decision making is the prerequisite. Journal of Health Service Research and Policy 1996; **1**: 104-113.

41 Dowie, J. The research-practice gap and the role of decision analysis in closing it. *Health Care Analysis* 1996; **4**: 1-14.

42 Drummond M. Economic analysis alongside clinical trials: problems and potential. *Journal of Rheumatology* 1995; **22**: 1403-1407.

43 Drummond, MF. and Davies, L. Economic analysis alongside clinical trials: revisiting the methodological issues. *International Journal of Technology Assessment in Health Care* 1991; **7**: 561-573.

44 Drummond, MF. and Menzin J. and Oster G. Methodological issues in economic assessments of new therapies: the case of colony stimulating factors. *Pharmacoeconomics* 1994; **6** Suppl.2:18-26.

45 Drummond, MF. and Stoddart, GL. Economic analysis and clinical trials. *Controlled Clinical Trials* 1984; **5**: 115-128.

46 Drummond, MF. Stoddart, GL. and Torrance, GW. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press, 1987.

47 Drummond, MF. Torrance, G. and Mason, J. Cost-effectiveness league tables: more harm than good? *Social Science and Medicine* 1993; **37**: 33-40.

48 Eddy, D. Selecting technologies for assessment. *International Journal of Technology Assessment in Health Care* 1989; **5**: 485-501.

49 Evidence-based Medicine Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine. *Journal of the American Medical Association* 1992; **268**: 2420-2425.

50 Fleming, TR. and Wateley, LF. Approaches to monitoring clinical trials. *Journal of the National Cancer Institute* 1989; **81**: 188-193.

51 Fleming, TR. Harrinton, DP. and O'Brien, PC. Designs for group sequential tests. *Controlled Clinical Trials* 1984; **5**: 348-361.

52 Fletcher, A. et al. Implications for trials in progress of publication of positive results. *Lancet* 1983; **342**: 653-657.

53 Freedman, B. Equipoise and the ethics of clinical research. *New England Journal of Medicine* 1987; **317**: 141-145.

54 Freedman, LS. Lowe, D. and Macaskill, P. Stopping rules for clinical trials incorporating clinical opinion. *Biometrics* 1984; **40**: 575-586.

55 Freedman, LS. and Spiegelhalter, DJ Comparison of Bayesian with group sequential methods for monitoring clinical trials. *Controlled Clinical Trials* 1989; **10**: 357-367.

56 French, S. *Decision theory: an introduction to the mathematics of rationality*. Chicester: Horwood 1986.

57 Gafni, A. and Birch, S. Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem. *Canadian Medical Association Journal* 1993; **148**: 913-917.

58 Gardner, MJ. and Altman, DG. *Statistics with confidence-confidence intervals and statistical guidelines*. London: British Medical Journal, 1989.

59 Geller, NL. and Pocock, SJ. Interim analysis in randomised clinical trials: ramifications and guideline for practitioners. *Biometrics* 1987; **43**: 213-223.

60 Genest, C. and Zidek, V. Combining probability distributions: a critique and an annotated bibliography. *Statistical Science* 1986; **1**: 114-148

61 Gerard, K. and Mooney, G. QALY league tables: handle with care. *Health Economics* 1993; **2**:59-64.

62 Gittins, JC. Bandit processes and dynamic allocation indices (with discussion). *Journal of the Royal Statistical Society B* 1979; **41**:148-177.

63 Greenes, RA. Begg, CB. and Cain, KS. Patient orientated performance measures of diagnostic tests. 2 Assignment potential and assignment strength. *Medical Decision Making* 1984; **4**: 17-31

64 Greenes, RA. Cain, KS. and Begg, CB. Patient orientated performance measures of diagnostic tests. *Medical Decision Making* 1984; **4**: 1-17

65 Gould, AL. Sample size for event rate equivalence trials using prior information. *Statistics in Medicine* 1993; **12**: 2009-2023.

66 Guyatt, G. and Drummond, M Guidelines for the clinical and economic assessment of health technologies. *International Journal of Technology Assessment in Health Care* 1985; **1**: 551-582.

67 Hanley, JA. and McNeil, BJ. The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology* 1982; **143**: 29-36.

68 Hanley, JA. and McNeil, BJ. A method of comparing the areas under receiver operating characteristics curves derived from the same cases. *Radiology* 1983; **148**: 839-843.

69 Hartz, A. and Centor, R. Measures of the value of a diagnostic test derived from stochastic thresholds. *Medical Decision Making*. 1983; **3**: 149-153.

70 Hartz, A. et al. Stochastic thresholds. *Medical Decision Making*. 1986; **6**: 145-148.

71 Henry, D. Economic analysis as an aid to subsidisation: the development of Australian guidelines for pharmaceuticals. *Pharmacoeconomics* 1992; **1**: 54-67.

72 Hilden, J. The area under the ROC curve and its competitors. *Medical Decision Making*. 1991; **11**: 95-101.

73 Hoel, DG. Sodel, M. and Weiss, GH. A survey of adaptive sampling for clinical trials. In: Elashoff, R. editor. *Perspectives in Biometry*. New York: Academic Press, 1975.

74 Van Hoult, BA. et al. Costs, effects and C/E ratios alongside a clinical trial. *Health Economics* 1994; **3**: 309-19.

75 Igiewicz, B. Alternative designs: sequential, multi-stage, decision theory and adaptive designs. In: Buyse, ME. Staquet, J. and Sylvester, RJ. editors. *Cancer clinical trials: methods and practice*. Oxford University Press, 1983.

76 Jennision, C. and Turnbull, BW. Statistical approaches to interim monitoring of medical trials: a review and commentary. *Statistical Science* 1990; **5**: 299-317.

77 Johannesson, M. The Australian guidelines for subsidisation of pharmaceuticals: the road to cost-effective drug prescribing? *Pharmacoeconomics* 1992; **2**: 355-362.

78 Johannesson, M. and Weinstein, MC. On the decision rules of cost-effectiveness analysis. *Journal of Health Economics* 1993; **12**: 459-467

79 Johnson, N. Lilford, RJ. and Brazier, W. At what level of collective equipoise does a clinical trial become ethical. *Journal of Medical Ethics* 1991; **17**: 23-35.

80 Kass, RE. and Greenhouse, JB. Comment: a Bayesian Perspective. *Statistical Science* 1989; **4**: 310-317.

81 Kassirer, JP. The principles of clinical decision making: an introduction to decision analysis. *Yale Journal of Medicine* 1976; **49**: 149-164.

82 Keeler, E. Decision trees and Markov models in cost-effectiveness research. In: Sloan, FA. editor. *Valuing health care: cost, benefits and effectiveness of pharmaceuticals and other medical technologies.* Cambridge: Cambridge University Press, 1995.

83 Kenney, RL. and Raiffa, H. *Decisions with multiple objectives: preferences and value tradeoffs.* New York: John Wiley, 1976.

84 Lachin, JM Introduction to sample size determination and power analysis for clinical trials. *Controlled Clinical Trials* 1981; **1**: 13-28.

85 Lan, KKG. And DeMets, DL. Discrete sequential boundaries for clinical trials. *Biometrika* 1983; **70**: 659-663.

86 Laupacis, A. Feeny, D. and Detsky, AS. et al. How attractive does a new technology have to be to warrant adoption and utilisation? Tentative guidelines for using clinical and economic evaluations. *Canadian Medical Association Journal* 1992; **146**: 473-481.

87 Lee, P.M. *Bayesian Statistics: an Introduction.* Oxford: Oxford University Press, 1987.

88 Lilford, RJ. and Johnson, N. The alpha and beta errors in randomised trials. New England Journal of Medicine 1990; **322**: 780-781.

89 Lilford, RJ. and Thornton, JG. Decision logic in medical practice. *Journal of the Royal College of Physicians* 1992; **26**: 400-412.

90 Lilford, RJ. Decision theory to asses delta. *Lancet* 1990; 483-484.

91 Lilford, R. and Jackson, J. Equipoise and the ethics of randomization. *Journal of the Royal Society of Medicine* 1995; **88**: 552-559.

92 Lusted, LB. Signal Detect ability and medical-decision making. Science 1971; **171**: 1217-1219.

93 Lusted, LB. *Introduction to medical decision making*. Springfield, 1968.

94 Maurice, R. A different loss function for the choice between two populations. *Journal of the Royal Statistical Society B* 1959; **21**: 203-213.

95 Martin, B. and Irvine, J. *Research foresight: priority setting in science*. London: Pinter, 1989.

96 Mendoza, G. and Iglewicz, B. An extension of Colton's model for comparing two medical treatments. *Journal of the American Statistical Association* 1978; **73**: 646-649.

97 Metz, CE. Basic principles of ROC analysis. *Seminars in Nuclear Medicine* 1978; **8**: 283-298.

98 Nease, RF. And Bonduelle, B. Solid recommendations from soft numbers: the test/treatment decision. *Medical Decision Making* 1987; 7: 220-233.

99 Nemhauser, GL. *Introduction to Dynamic Programming*. New York: Wiley, 1966.

100 O'Brien, BJ. Drummond, MF. Labelle, RJ. et al. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. *Medical Care* 1994; **32**: 150-63

101 O'Brien BJ. and Drummond, MF. Statistical versus quantitative significance in the socioeconomic evaluation of medicine. *Pharmacoeconomics* 1994; **5**: 389-398.

102 Office of Science and Technology. *Research Foresight and the exploitation of the Science Base*. London: HMSO, 1993.

103 Pauker, SG. and Kassirer, JP. Therapeutic decision making: a cost benefit analysis. *New England Journal of Medicine* 1975; **293**: 229-234.

104 Pauker, SG. and Kassirer, JP. Clinical application of decision analysis: a detailed illustration. *Seminars in Nuclear Medicine* 1978; **8**: 324-335.

105 Pauker, SG. and Kassirer, JP. The threshold approach to clinical decision making. *New England Journal of Medicine* 1980; **302**: 1109-1117.

106 Phelps, CE. and Hall, WJ. Appropriate characterisation of differences between ROC curves. *Medical Decision Making* 1987; **7**: 286.

107 Phelps, CE. and Mushlin, AI. Focusing technology assessment using medical decision theory. *Medical Decision Making* 1988; **8**: 279-289.

108 Phelps, CE. and Mushlin, AI. On the near equivalence of cost-effectiveness and cost-benefit analysis. *International Journal of Technology Assessment in Health Care* 1991; **7**: 12-21.

109 McPherson, K. and DeMets, DL. Design and analysis of group sequential tests based on the type I error spending rate function. *Biometrika* 1987; **74**: 149-154.

110 McPherson, K. Sequential stopping rules in clinical trials. *Statistics in Medicine* 1990; **9**:595-600.

111 Pocock, SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 1977; **64**: 191-199.

112 Pocock, SJ. *Clinical trials: a practical approach.* Chichester: Wiley, 1987.

113 Pocock, SJ. When to stop a clinical trial. *British Medical Journal* 1992; **305**: 235-240.

114 Pocock, SJ. and Hughes, MD. Estimation issues in clinical trials and overviews. *Statistics in Medicine* 1990; **9**: 657-671.

115 Raiffa, H. *Decision analysis: introductory lectures on choices under uncertainty.* Addison-Wesley, 1968.

116 Raiffa, H. and Schlaifer, R. *Probability and Statistics for Business Decisions.* McGraw-Hill, 1959.

117 Reinhardt, UE. Reflections on the meaning of efficiency: can efficiency be separated from equity? *Yale Law and Policy Review* 1992; **10**: 302-315.

118 Richardson, W. and Detsky, A. Users guide to the medical literature VII. How to use a clinical decision analysis. Are the results of the study valid? *Journal of the American Medical Association* 1995; **273**: 1292-1295.

119 Rosenberg, W. and Donald, D. Evidence based medicine: an approach to clinical problem-solving. *British Medicial Journal* 1995; **310**: 1122-1126.

120 Sackett, D. and Rosenberg, W. On the need for evidence based medicine. *Health Economics* 1995; **4**: 249-254.

121 Schlaifer, R. *Introduction to Statistics for Business Decisions*. McGraw-Hill, 1961.

122 Schwartz, D. and Lellouch, J. Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Chronic Disease* 1967; **20**: 637-648.

123 Schwartz, D. Flamant, R. and Lellouch, J. *Clinical Trials*. London, Harcourt Brace Jovanovich 1980

124 Smith, P. Dynamic Programming in Action. *Journal of Operations Research Society* 1989; **40**: 779-787.

126 Spiegelhalter, DJ. and Freedman, LS. A predictive approach to selecting the size of a clinical trial based on subjective clinical opinion. *Statistics in Medicine* 1986; **5**: 1-13.

127 Spiegelhalter, DJ, Freedman, LS and Parmar, MKB. Bayesian approaches to randomised trials. *Journal of the Royal Statistical Society A* 1994; **157**: 357-416.

128 Spiegelhalter, DJ. and Freedman, LS. Bayesian approaches to clinical trials. *Bayesian Statistics Volume 3*. Oxford: 1988: 453-477.

129 Sugden, R. And Williams, A. *The principles of practical cost-benefit analysis*. Oxford: Oxford University Press, 1978.

130 Swets, JA. et al. Evaluation of diagnostic technologies. *Science* 1979; **205**: 735-759.

131 Sylvester, RJ. and Staquet, MJ. Design of phase II clinical trials in cancer using decision theory. *Cancer Treatment Reports* 1980; **64**:519-524.

132 Thornton, J. Lilford, R. and Johnson, N. Decision analysis in medicine. *British Medical Journal* 1992; **304**: 1099-1103.

133 Upton, GJG. and Lee, RD. The importance of the patient horizon in the sequential analysis of binomial clinical trials. *Biometrika* 1976; **63**: 335-342.

134 Ware, JH. Investigating therapies of potentially great benefit: ECMO. *Statistical Science* 1989; **4**: 298-340.

135 Warner, K. and Luce, B. Cost-benefit and cost-effectiveness analysis in health care. Ann Arbor, Mich.: *Health Administration Press*, 1982.

136 Weinstein, MC. Principles of cost-effective resource allocation in health care organisations. *International Journal of Technology Assessment in Health Care*. 1990; **6**: 93-103.

137 Weinstein, MC. and Fineberg, HV. *Clinical Decision Analysis*. WB. Saunders 1980.

138 Weinstein, MC. Shepard, DS. and Pliskin, JS. The economic value of changing mortality probabilities: a decision -theoretic approach. *Quarterly Journal of Economics* 1980; **95**: 373-396.

139 Weinstein, MC. and Stason, WB. Foundations of cost-effectiveness analysis for health and medical practices. *New England Journal of Medicine* 1977; **296**: 716-721.

140 Weinstein, MC. and Zeckhauser, RJ. Critical ratios and efficient allocation. Journal of Public Economics 1972; **2**: 716-721.

141 Weinstein, MC. From cost-effectiveness ratios to resource allocation: where to draw the line? In: Sloan, FA. editor. *Valuing health care: cost, benefits and effectiveness of pharmaceuticals and other medical technologies*. Cambridge: Cambridge University Press, 1995.

142 Willan, AR. and O'Brien, BJ. Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. *Health Economics* 1996; **5**: 297-305.

143 Williams, A. Health economics: the end of clinical freedom. *British Medical Journal* 1988; **297**: 1183-1187.

144 Williams, A. Cost-effectiveness analysis: Is it ethical? *Journal of Medical Ethics* 1992; **18**: 7-11.

145 Williams, A. Economics QALYs and medical ethics. *CHE Discussion Paper* No 121, 1998.

146 Williams, A. How should information on cost influence clinical practice? In: Delmothe, T. editor. *Outcomes Into Clinical Practice*. British Medical Journal publishing group, 1994.

147 Williams, N. and Anderson, J. The future of cardiovascular research. *PRISM Research Report* No 1, 1993.

148 White, DJ. *Dynamic Programming*. Edinburgh: Oliver and Boyd, 1969.

149 Whitehead, J. *The Design and Analysis of Sequential Trials*. Ellis Horwood, 1992.

150 Wulff, HR. Commentary on Guyatt and Drummond. *International Journal of Technology Assessment in Health Care* 1985; **1**: 567-581.

151 Zelen, M. Play the winner rule and the controlled clinical trial. Journal of the American Statistical Association 1969; **64**: 131-146.