ECONOMIC ANALYSIS AND RANDOMISED CONTROLLED TRIALS: AN INVESTMENT APPRAISAL APPROACH

Martin E Backhouse, BA, MSc.

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

February 2006

ABSTRACT

Randomised controlled trials (RCTs) play a fundamental role in the development and marketing activities of pharmaceutical companies. They are the primary means of evaluating the tolerability, safety and efficacy of a drug, and for providing information relevant for pricing and reimbursement decisions and clinical decision - making. RCTs require a substantial investment by pharmaceutical companies and the financial consequences of poorly or sub - optimally designed trials are potentially substantial. Revenue does not materialise unless a licence to market a product is granted and sales may be restricted if a trial fails to provide evidence of sufficient strength or relevance for those involved in product adoption decisions. From a pharmaceutical company's perspective, the value of RCTs can therefore be judged on the contribution they make to the performance of a drug in the market and hence on their contribution to the performance of the firm. Consequently, the design choices made in the planning of RCTs are effectively investment appraisal decisions. However, the application of investment appraisal techniques to RCT design has not previously been proposed.

The purpose of this thesis is to consider how private sector investment appraisal methods might be applied to RCT design decision-making and to explore aspects of the practicalities of application. A general investment appraisal model is presented and its application to determine profit maximising RCT designs is illustrated. Considering the cost side of the investment appraisal equation, it is shown how decision-makers' requirements for cost-effectiveness evidence derived from trials could have a significant

impact on the major determinants of cost (sample size and study duration) depending on their specific preferences for evidence defined over key components of RCT design. Considering the revenue side of the investment appraisal equation, it is shown how discrete choice analysis could be used to incorporate decision-makers' preferences for RCT designs into the planning of studies. Specifically, it is shown how the predicted probabilities derived from the application of this technique could be used within an investment appraisal framework. Directions for future research into the application of investment appraisal to RCT design are proposed.

PUBLICATIONS

The research presented in Chapter 2 of this thesis was published in the journal *Health Economics* in 1998.[1]

Aspects of the research presented in Chapter 3 of this thesis were presented at the ISPOR Annual European Congress in 2000 where it received an award for 'Best Contributed Poster Presentation'.[2] The research was subsequently published in the journal *Pharmacoeconomics* in 2002.[3]

Aspects of the research presented in Chapter 4 of this thesis were presented at the ISPOR Annual European Congress in 2002 where it was nominated for the award of 'Best Contributed Poster Presentation'.[4]

1. Backhouse ME. An investment appraisal approach to clinical trial design. Health Economics 1998;7:605-19.
2. Backhouse, ME. NICE requirements for cost-effectiveness evidence: potential sample size implications of generating trial-based evidence. 2000. http://www.ispor.org/awards/3euro_.asp.
3. Backhouse ME. Use of randomised controlled trials for producing cost effectiveness evidence: potential impact of design choices on sample size and study duration. PharmacoEconomics 2002; 20:1061-77.
4. Backhouse, M. E. The use of discrete choice analysis in the design of clinical trials. Value in Health 5(6), 529. 2002.

ACKNOWLEDGEMENTS

I would like to thank Paul Fenn for his guidance and supervision and my family, Rebecca, Edward and Elliot, for their support and patience throughout the course of this research. The thesis is dedicated to my father, Ernest Backhouse, who inspires hard work and a positive approach to life.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND

1.1.1 The importance of randomised controlled trials in drug development and marketing

Randomised controlled trials (RCTs) play a fundamental role in the development and marketing activities of pharmaceutical companies.[1-5] The evidence generated by RCTs figures prominently in the product adoption decisions taken by regulatory agencies, pricing and reimbursement authorities, health technology appraisal bodies and individual physicians. Regulatory agencies, such as the US Food and Drug Administration (FDA)[6] and the European Medicines Evaluation Agency (EMEA)[7] are responsible for granting licences to companies enabling them to market their products. The regulatory process requires that companies perform RCTs to demonstrate a product's tolerability, safety and clinical efficacy. If the latter are demonstrated to the satisfaction of the regulatory agencies, the manufacturer is granted a licence to market the product.

Clearly, the granting of a licence to market a product is a necessary condition for a company to obtain a return on its investment, but it is by no means sufficient. In most major markets, a manufacturer must negotiate with national agencies (or large groups of purchasers) responsible for setting the price, reimbursement and formulary status of a drug.[8;9] A company submits information dossiers according to local guidelines.[9] Typically, they all include some requirement for evidence pertaining to unmet therapeutic need, clinical effectiveness, budget impact and in some jurisdictions, cost-effectiveness analyses.[10-13] Of particular relevance here is that the RCT evidence

used in the marketing approval submission is often re-appraised in the course of product adoption negotiations. However, the RCT evidence produced for regulatory purposes may not be sufficient to answer the questions of primary interest to pricing and reimbursement authorities or health technology assessment bodies such as the National Institute for Health and Clinical Excellence (NICE).[14] Most of the major challenges to achieving a positive pricing and reimbursement outcome are directed at the strength and relevance of the clinical evidence base, as illustrated by recommendations made by NICE[15] and commentaries in medical journals.[16;17]

RCTs require a substantial investment by pharmaceutical companies. It has been estimated that the cost of discovering and developing a new drug introduction exceeds US $300 millions at 1995 prices.[18-20] Most of the cost relates to the conduct of RCTs, in particular the major regulatory studies (known as Phase III trials). Therefore the financial consequences of conducting RCTs that do not meet the evidence needs of regulatory bodies, pricing and reimbursement authorities and other health care decision-makers are potentially significant since a return on the development investment may not materialise or be severely restricted until adequate evidence is generated.

1.1.2 Economic analysis and randomised controlled trials

The literature relating to economic analysis and RCTs consists of two broad components. The first component is concerned with using trials as *vehicles for economic evaluation* of health care technologies. Economic evaluation is the term given to a set of techniques for appraising the economic value of health programmes or treatments from a health care funding body or societal perspective. Many

references include good overviews of the techniques[21;22] and their application.[23 24-30] It has been pointed out[31] that an economic evaluation can be wholly deterministic (utilizing decision-analytic techniques[32-35]), wholly stochastic (based on data sampled from studies such as trials[36-38]) or a combination of the two.[39] Since RCTs are the main scientific method for collecting patient level data for evaluating the clinical benefits of an intervention[40], they have attracted research attention from economists exploring how they can be used for collecting data necessary for performing an economic evaluation. There is an extensive literature concerning the use of RCTs as vehicles for collecting and analyzing data for the purpose of economic evaluation (see for example[23;31;36-38;41-68;68-129]).

The second (and much smaller) component of the literature relating to economic analysis and RCTs focuses on the *economic analysis of research project selection and trial design.* A number of researchers have examined how economic considerations might be used to optimise the design of RCTs.[130-142;142-151] However, there are two prominent contributions in this area, each of which will be summarised briefly below.

The first contribution is the work by Detsky, conducted in the 1980s, which considers how economic analysis might be used to assist government agencies with their decisions about awarding research grants for clinical trials.[152-154] His work was motivated by a concern that funds allocated to trials might not be used optimally due to the often arbitrary way in which key study design parameters are chosen. In particular, he noted that trial sample sizes are often inadequate to determine whether a new treatment is effective or not and that they are often calculated based on arbitrarily chosen clinical differences thought to be worth detecting. Using a number of examples based on life-saving treatments, Detsky illustrates how the cost-

effectiveness of trials can be evaluated in the planning stages and how the results can be used to decide whether a particular trial should be funded or not and if so how large the trial should be. To this end, cost-effectiveness ratios are calculated for trials of different sample sizes. In his model, costs are a straightforward function of sample size which in turn is driven by the choice of the clinically important difference a study aims to detect. Clearly, there is a positive relationship between the cost of conducting a trial and its sample size. In the model, the expected effectiveness of a trial is measured in terms of lives saved. It is calculated as the product of the size of the population who would benefit from the new treatment if adopted, the power of the study, the prior distribution of the true reduction in the relative risk of death and a proportionate implementation factor. The latter can be thought of as a crude demand function. In Detsky's model, the proportionate implementation factor takes on values which are such that a new treatment is completely adopted where a trial shows a statistically significant effect and is not adopted if it fails to show statistical significance.

The second and more recent contribution is from Claxton, whose work was developing in parallel with this thesis.[155-162] He proposes the use of value of information analysis in order to set research priorities, determine whether further research is required to inform treatment adoption decisions and, if so, what the optimal design of that research should be. The approach is developed for a jurisdiction in which treatment adoption decisions are based on the results of cost-effectiveness analysis. Consequently, the proposed method for determining the optimal design of trials is also based on cost-effectiveness analysis in order to ensure that the criteria used for decisions about future research are consistent with those used

for treatment selection. The approach proposed by Claxton involves a number of elements, which are summarised briefly in non-technical terms below.

The analysis starts from the position that if an estimate of the cost-effectiveness of a new treatment is acceptable to the decision-maker then it should be adopted regardless of whether the result is statistically significant or not. It is argued that the traditional paradigm of hypothesis testing is irrelevant and that it should be replaced by a method that minimises the societal cost of a wrong decision. Any estimate of cost-effectiveness will be uncertain and consequently a decision based upon it could be incorrect. An incorrect decision will incur costs (known as opportunity loss) and the approach centres on quantifying the societal cost of the uncertainty surrounding a decision. The costs of uncertainty can be interpreted as the expected value of perfect information (EVPI) because perfect information would eliminate the possibility of a wrong decision. So if the EVPI exceeds the cost of any future research aimed at acquiring additional information then further research is worthwhile i.e. it is potentially cost-effective. Once it is deemed that further research is potentially beneficial, the optimal scale (sample size) of the research needs to be determined. Within Claxton's approach, the optimal sample size for a trial is that which gives rise to an expected net benefit of sampling information (ENBS) that is positive and at its maximum. ENBS is the difference between the expected value of sampling information (EVSI) and the cost of conducting a trial at any given sample size. EVSI is measured as the reduction in opportunity loss which results from a trial of given size.

The above works are relevant to this thesis in that they recognise the importance of optimising clinical trial designs and propose possible methods for doing so.

The methods proposed adopt a societal perspective based around the use of cost-effectiveness analysis. However, the approaches are not generally applicable because most major pharmaceutical markets do not currently base their technology adoption decisions on cost-effectiveness analysis. The works of both Detsky and Claxton utilise very simplistic demand functions whereby a new product is effectively assumed to be adopted immediately and completely if a given level of cost-effectiveness is achieved. The relationship between product diffusion and cost-effectiveness is poorly understood and, more generally, there is an absence of research that explicitly explores the relationship between the adoption of a technology and the design of clinical trials. Whilst Claxton rejects traditional hypothesis testing as irrelevant, the fact is that the demonstration of statistically significant benefits of new products remains a key determinant of drug regulatory decisions and cannot therefore be ignored. This fact is recognised by Detsky whose work not only inspired the topic of this thesis but also provided a basis for computing the expected outcomes of clinical trials that will be seen in later chapters.

In summary therefore, a notable feature of the current literature that focuses on the economic analysis of trial design is that none of the approaches to optimizing trial design adopt a private sector (pharmaceutical company) perspective. Equally notable is the fact that the pharmaceutical research and development project appraisal literature ignores the potential importance of RCT *design and results* on the uptake of a product post-approval.[163-165]

1.2 FOCUS AND STRUCTURE OF THE THESIS

Given the importance of RCT evidence for influencing the nature and extent of product adoption and the substantial cost of performing clinical trials, it seems logical that a privately owned pharmaceutical company should seek to optimise RCT designs based on economic (in particular profit) criteria. Yet a review of the literature pertaining to economic analysis and clinical trials reveals that a private sector investment appraisal approach to RCT design has not hitherto been explored. Therefore this thesis attempts to fill a significant gap in the literature by setting out how methods of investment appraisal might be applied to RCT design decision-making and by exploring aspects surrounding the practicalities of application. To this end the remainder of the thesis is made-up of four chapters (summarised below) of which Chapters 2-4 are presented as 'standalone' (but loosely interrelated) manuscript format pieces.

Chapter 2 sets out a general investment appraisal model which shows how pharmaceutical companies could take profit considerations into account when making decisions about the design of randomised controlled trials. A general model is presented based on the net present value (NPV) method of investment appraisal. The application of investment appraisal requires an evaluation of both the costs and expected revenues associated with a given RCT design. Therefore a description of the major determinants of costs and revenues and how they might be estimated is presented. The importance of being able to estimate the demand for a product contingent upon RCT design and expected trial outcomes is emphasized. The approach is illustrated with a hypothetical example showing how optimal (net present value maximising) designs can be determined based on choices about key

RCT design parameters. Directions for further research are suggested and these set the scene for the themes that are explored in the following two chapters.

Chapter 3 focuses on issues associated with the cost side of the investment appraisal equation. It takes as the starting point the special situation where it is assumed that decision-makers a) make product adoption decisions based on cost-effectiveness analyses, and b) require that those analyses be based on sampled data derived from RCTs (so called wholly stochastic cost-effectiveness analysis). The primary purpose of this chapter is to illustrate how the specific nature of decision-makers' preferences relating to wholly stochastic cost-effectiveness evidence could have a significant impact on the major determinants of RCT costs, namely sample size and trial duration. Data collected in a clinical evaluation are used to calculate sample sizes to test cost-effectiveness hypotheses for hypothetical study designs. These are compared with the sample sizes required to test hypotheses based only on clinical endpoints. It is shown that circumstances can be such that a wholly stochastic cost-effectiveness analysis might not be a practical proposition even though its clinical counterpart is. The importance of prior specification of decision-makers' preferences for different components of RCT study design is emphasised.

Chapter 4 focuses primarily on issues associated with estimating the revenue side of the investment appraisal equation. It considers a general situation where decision-makers involved with product adoption decisions will have preferences for the types of RCT evidence they want to see. The extent to which these preferences are satisfied will influence the nature and extent of a treatment's use. The primary purpose of this chapter is to illustrate how the technique of discrete choice analysis (DCA) could be used by companies to consider decision- makers' preferences for RCT designs

when planning their studies. The approach is illustrated using the design of trials of adjuvant bisphosphonates in the management of patients with primary operable breast cancer as a case study. A stated preference survey is conducted to elicit clinicians' preferences for evidence and to model the predicted probabilities of prescribing a product. It is shown how the predicted probabilities of product adoption can be used by a company within an investment appraisal framework to identify a profit maximizing RCT design. Issues for consideration in future research into the application of DCA in this context are discussed.

Chapter 5 presents the main conclusions and contributions of the thesis. It includes a brief summary of directions for future research into the application of investment appraisal to RCT design.

CHAPTER 2: AN INVESTMENT APPRAISAL APPROACH TO CLINICAL TRIAL DESIGN

SUMMARY

In this chapter, a general investment appraisal model is presented which shows how pharmaceutical companies could take profit considerations into account when making decisions about the design of randomised controlled trials. A general model is presented based on the net present value method of investment appraisal. The approach is illustrated with a hypothetical example which shows how optimal (net present value maximising) designs can be determined based on choices about sample size and endpoint measurement. The method could be extended to accommodate considerations about other trial design features, and could be used to determine a portfolio of studies which maximises the expected return on a given development or trial budget. Furthermore, the approach could be used by pharmaceutical companies to evaluate the incremental costs and benefits of incorporating non-clinical objectives into trials, such as quality of life research and economic evaluation studies. A number of practical difficulties would need to be overcome to utilise the approach. Directions for further research are therefore highlighted centred on the key components of the model: a trial cost function, a product demand function, innovation diffusion processes and Bayesian approaches to trial design.

2.1 INTRODUCTION

Randomised controlled trials (RCTs) play a fundamental role in the development and marketing activities of pharmaceutical companies. They are the primary instruments for evaluating the tolerability of a drug, for demonstrating its efficacy, and for providing information relevant for clinical decision - making.[1-5] RCTs require a substantial investment by pharmaceutical companies in terms of the human and financial resources allocated to their design, execution, analysis and reporting. It has been estimated that the cost of discovering and developing a new drug introduction exceeds US $300 millions at 1995 prices.[18;19] A significant proportion of that cost relates to Phase III trials which are conducted to produce evidence about a product's safety and efficacy to a level which at least satisfies the regulatory authorities responsible for granting product licences. Once approval to market has been granted, companies often conduct Phase IV trials designed to address the information needs of decision-makers involved in product utilisation decisions. The financial consequences of poorly or sub - optimally designed Phase III or Phase IV trials are potentially substantial: revenue is lost if a new drug fails to gain access to its intended market or if a trial fails to provide evidence of sufficient strength or relevance to secure or enhance a product's use.

From a pharmaceutical company's perspective, the value of RCTs can therefore be judged on the contribution they make to the performance of a drug in the market and hence on their contribution to the performance of the firm. Consequently, the design choices made in the planning of RCTs, such as which comparators to use, which endpoints to evaluate and which sample size to adopt, are effectively investment appraisal decisions. The nature and scale of a trial will drive the size of the

investment. The return on the investment will depend upon the sensitivity of decision - makers to the different types, strengths, relevance and quality of the evidence provided by it.

Although some applications of decision - analytic techniques to clinical trial design decisions have previously been reported, none has adopted a private sector investment appraisal perspective.[143;149;152;154;155] Therefore the purpose of this chapter is to illustrate the potential role for investment appraisal in assisting with RCT design decisions taken by pharmaceutical companies. The remainder of the chapter is divided into three sections. Firstly, the components of an investment appraisal model are set out in general form based on the net present value method of investment decision-making. Secondly, the application of the approach is illustrated with simulations using a hypothetical Phase IV trial design scenario facing a pharmaceutical company. Finally, the discussion section addresses issues surrounding further research that would be required to develop the approach for practical application.

2.2 INVESTMENT APPRAISAL APPROACH TO TRIAL DESIGN: A GENERAL MODEL

Investment appraisal is the term given to a general framework used by firms to assist with project investment decisions. Whilst a number of possible approaches are discussed in the literature, the superiority of investment decisions based on the net present value (NPV) discounting method is well documented.[166-168] The investment appraisal approach to trial design is therefore illustrated using the NPV method. With this method, choices about trial design would involve consideration of the differences

between the discounted revenue a trial is expected to yield and the expected discounted cost of conducting it i.e. trial design decisions would incorporate expected NPV (profit) considerations. The key components of this approach are set out below in general form.

2.2.1 Cost function

A pharmaceutical company will be faced with a clinical trial cost function of the following general form:

$$C_t(n) = C_t^P\left(Q_t(n)\right) + C_t^T(n) \tag{1}$$

where $C_t(n)$ is the cost in period t associated with conducting a trial of a particular design and sample size, n. Any additional manufacturing and marketing costs incurred as a direct consequence of conducting a trial are denoted by $C_t^P\left(Q_t(n)\right)$, where $Q_t(n)$ is the number of units of the product sold in time t. $C_t^T(n)$ is the cost in period t of designing and conducting a trial, and will partly depend upon the chosen sample size. Note that the term design refers to a trial defined over all relevant characteristics, such as whether it is of a parallel groups or cross - over form, study population, comparators, endpoints and duration of subject follow - up.

The total cost of conducting a trial, $TC(n)$, is given by:

$$TC(n) = \sum_{t=0}^{H} C_t(n) = \sum_{t=0}^{H}\left(C_t^P\left(Q_t(n)\right) + C_t^T(n)\right) \tag{2}$$

where the costs of the investment are evaluated over the period adopted for the appraisal beginning at time $t = 0$ and ending at time $t = H$. Costs will be

incurred in different time periods. For example, trial subjects may be enrolled in different years and the costs associated with data analysis and reporting will take place once data collection and processing are complete. With the NPV method of investment appraisal, the time distribution of costs and the risk associated with the project are taken into account by discounting costs to their present value:

$$PTC(n) = \sum_{t=0}^{H} \frac{C_t^P\left(Q_t(n)\right) + C_t^T(n)}{(1+r)^t} \tag{3}$$

where $PTC(n)$ is the present value of the total cost of conducting the trial and r is the project's risk adjusted opportunity cost of capital (see section 2.4 below).

2.2.2 Demand function

A demand function for a good or service specifies the relationship between the quantity demanded and the factors that influence it. Similarly, a prescribing clinicians' demand function for a drug states the relationship between the quantity of treatments demanded and its determinants. Demand functions therefore figure prominently in the estimation of revenue resulting from the conduct of a trial since they describe, inter - alia, the likely responsiveness of clinicians to the evidence about a drug's attributes provided by it. A prescribers' demand function for a drug under investigation will take the following general form:

$$Q_t^* = Q_t^*(\Delta X, \Delta P) \tag{4}$$

where Q_t^* is the *desired demand* for a treatment under investigation, say drug A. Q_t^* is expressed in terms of the number of units of A given demonstrated statistically significant incremental changes in the vector of product characteristics, ΔX,

compared with alternative treatments, say B. Note that the vector X represents *potential* trial endpoints which might include non - clinical parameters, such as non - drug treatment costs and quality of life measures. Since drug price considerations are becoming an increasingly important consideration in prescribing decisions, differential drug treatment costs are highlighted in the demand function and are denoted by ΔP.

It is unrealistic to assume that the desired level of demand, Q_t^*, will be realised immediately the results of a trial become known. For example, factors such as the perceived quality of the evidence, delays in the dissemination of the information and the learning process associated with the adoption of a new treatment mean that it will take time for prescribers to switch from existing practices to drug A, and the desired level of use may never be realised. Further, the prescribing practices of clinicians might be influenced by other factors, such as the policies proposed by formulary committees and the promotional activities of pharmaceutical companies. Thus, the *actual demand* for drug A in any specific time period will be determined through a diffusion process in which the growth of actual demand, Q_t, towards Q_t^* will be determined by experience gained through past consumption, Q_{t-1}, and the rate of adoption, k:

$$Q_t = Q_t\left(Q_t^*(\Delta X, \Delta P), Q_{t-1}, k\right) \tag{5}$$

Since the results of a trial are uncertain, a decision analytic approach is required to estimate the *expected demand* for the product resulting from the conduct of a trial. This requires the firm to estimate the expected outcomes with respect to potential trial endpoints, ΔX, which need to be established by empirical demand research

to be important. Making use of the approach illustrated by Detsky,[152;154] if only discrete values of ΔX are possible, the expected effect sizes for the product attributes, $E_n(\Delta X)$, likely to be demonstrated by a trial of sample size n, is given by the following formula:

$$E_n(\Delta X) = \sum_{\Delta X=-\infty}^{\infty} \Pr\left(D_n^{\phi} = \Delta X|\Delta X\right).\Pr(\Delta X).\Delta X \qquad (6)$$

where $\Pr\left(D_n^{\phi} = \Delta X|\Delta X\right)$ is the conditional probability that a difference of ΔX will be established in a trial with significance level ϕ if that difference is in fact there, and where $\Pr(\Delta X)$ is the prior probability of a true difference of ΔX. The expression $\Pr\left(D_n^{\phi} = \Delta X|\Delta X\right)$ is otherwise known as the power of a trial.[169;170] In the case where ΔX are continuous variables, the expected trial outcome is given by:

$$E_n(\Delta X) = \int_{\Delta X=-\infty}^{\infty} \Pr\left(D_n^{\phi} = \Delta X|\Delta X\right).\Pr(\Delta X).\Delta X \, d\Delta X \qquad (7)$$

where the definitions given for equation (6) apply. For convenience, the discrete case is used in the equations and hypothetical analyses which follow.

By substituting the expression for $E_n(\Delta X)$ from equation (6) for ΔX in the dynamic demand function of equation (5) we derive the following general expression for *expected demand* in time period t, $Q_t(n)$, resulting from a trial of given design, sample size and significance level:

$$Q_t(n) = Q_t\left(Q_t^*\left(\sum_{\Delta X=-\infty}^{\infty} \Pr\left(D_n^{\phi} = \Delta X|\Delta X\right).\Pr(\Delta X).\Delta X, \Delta P\right), Q_{t-1}, k\right) \qquad (8)$$

2.2.3 Revenue function

The expected revenue accruing to a company from the use of drug A in each time period, $R_t(n)$, is determined by multiplying the expected number of units of the product demanded in that time period, $Q_t(n)$, by the price per unit. Let P_t denote the exogenously determined selling price for drug A. By calculating the product $P_t.Q_t(n)$, the following expression for the expected revenue accruing to the firm in time period t is derived:

$$R_t(n) = P_t.Q_t(n) = P_t.Q_t\left(Q_t^*\left(\sum_{\Delta X=-\infty}^{\infty}\Pr\left(D_n^{\phi} = \Delta X|\Delta X\right).\Pr(\Delta X).\Delta X, \Delta P\right), Q_{t-1}, k\right) \quad (9)$$

The total revenue accruing to the firm as a result of conducting the trial, $TR(n)$, is therefore given by:

$$TR(n) = \sum_{t=0}^{H} P_t.Q_t(n) = \sum_{t=0}^{H} P_t.Q_t\left(Q_t^*\left(\sum_{\Delta X=-\infty}^{\infty}\Pr\left(D_n^{\phi} = \Delta X|\Delta X\right).\Pr(\Delta X).\Delta X, \Delta P\right), Q_{t-1}, k\right) \quad (10)$$

Since the expected revenues occur in different time periods and are subject to uncertainty they must, like costs, be discounted to their present values using the project's risk adjusted opportunity cost of capital (see section 2.4 below):

$$PTR(n) = \sum_{t=0}^{H} \frac{P_t.Q_t\left(Q_t^*\left(\sum_{\Delta X=-\infty}^{\infty}\Pr\left(D_n^{\phi} = \Delta X|\Delta X\right).\Pr(\Delta X).\Delta X, \Delta P\right), Q_{t-1}, k\right)}{(1+r)^t} \quad (11)$$

where $PTR(n)$ is the present value of the total revenue resulting from the conduct of a trial.

2.2.4 Investment appraisal decision rules

The final component of the investment appraisal approach is the calculation of the expected profit accruing to the firm as a consequence of conducting a trial of given design and sample size. With the NPV method, this is simply the difference between the discounted value of a trial's expected total revenue, $PTR(n)$, and the discounted value of expected total cost, $PTC(n)$:

$$NPV(n) = PTR(n) - PTC(n) \tag{12}$$

where $NPV(n)$ denotes the expected net present value resulting from a trial of given design, sample size and significance level. Substituting for $PTR(n)$ and $PTC(n)$ from equations (11) and (3) respectively, and substituting the expression for $Q_t(n)$ given in equation (8) into equation (3), we derive the following general expression for $NPV(n)$:

$$NPV(n) = \sum_{t=0}^{H} \frac{P_t . Q_t\left(Q_t^*\left(\sum_{\Delta X=-\infty}^{\infty} \Pr\left(D_n^{\phi} = \Delta X | \Delta X\right).\Pr(\Delta X).\Delta X, \Delta P\right), Q_{t-1}, k\right)}{(1+r)^t}$$

$$- \sum_{t=0}^{H} \frac{C_t^P\left(Q_t\left(Q_t^*\left(\sum_{\Delta X=-\infty}^{\infty} \Pr\left(D_n^{\phi} = \Delta X | \Delta X\right).\Pr(\Delta X).\Delta X, \Delta P\right), Q_{t-1}, k\right)\right) + C_t^T(n)}{(1+r)^t} \tag{13}$$

It should be noted that the superiority of the NPV method over alternative methods of making investment decisions stems primarily from the choice of H and from the use of discounting.[166-168] With the NPV method, H is the useful life of the

project (in this case the useful life of the information provided by a trial) rather than some arbitrarily chosen time horizon. Discounting is undertaken to allow for both the timing of costs and revenues *and* the risk associated with the project. A project's risk adjusted discount rate can be determined using the capital asset pricing model (CAPM), which states that:[167]

$$r = r_f + \psi(r_m - r_f) \qquad (14)$$

where r is the risk-adjusted discount rate, r_f is the rate of return on risk free investments (e.g. Treasury Bills), r_m is the rate of return on investments of similar risk to the project being appraised (e.g. other pharmaceutical company stock), and ψ is a measure of the risk of the project relative to other similar investments. Examples of the application of this approach can be found in an article in which costs of capital are estimated for a number of pharmaceutical companies and for projects at different stages of development.[171]

NPV criteria can be used to assist with a number of trial planning decisions based on the following decision rules:

a) A trial of given design is worth conducting if it yields a positive expected net present value: $NPV(n) = PTR(n) - PTC(n) > 0$;

b) The optimal choice of trial design, in terms of factors such as sample size and primary endpoints, is that which maximises the (positive) expected net present

value: $Max_n\ NPV(n) = PTR(n) - PTC(n) > 0$;

c) When allocating funds to different studies competing for a limited trial budget, the funds should be allocated across potential trial designs so as to maximise the expected net present value of the overall investment.

It should be noted that the trial design problem facing the firm might be one of constrained optimisation in which the unconstrained profit maximising solutions become unattainable. For example, decision - makers within the firm may impose budget, sample size, study duration or other constraints on themselves. Also, the company might face externally imposed constraints. For example, factors such as regulatory requirements, ethical considerations and insufficient patient numbers could restrict the choice of comparators or the duration of follow - up. Such constraints do not invalidate the approach, although they will, if binding, inhibit the achievement of the unconstrained NPV maximising solution.

The use of the NPV method of investment appraisal to determine whether an RCT is worth conducting and to determine the optimal strategy based on choices about sample size and endpoint measurement is illustrated below with a simple hypothetical example.

2.3 INVESTMENT APPRAISAL APPROACH TO TRIAL DESIGN: A HYPOTHETICAL EXAMPLE

The investment appraisal approach is illustrated with respect to the design of a hypothetical Phase IV trial. It is assumed that a pharmaceutical company has

developed a new product, drug A, for an acute condition for which there is currently only one other treatment available, drug B. Drug A has been approved for marketing on the basis of placebo controlled trials and the national pricing authority has granted the company a selling price P_{At} = \$200. Whilst prescribers of the product would be fully reimbursed if they were to use it, local hospital formulary committees have excluded drug A from their formularies on the grounds that no comparative trials have been conducted against drug B. It is therefore assumed that all patients needing treatment are currently prescribed drug B, which has been shown to produce a success rate of 25% on endpoint X_1 and 75% on endpoint X_2 after 6 months of follow - up. It is known that the incidence of the condition will remain stable, with approximately 200,000 patients presenting for treatment each year. No other treatment alternatives are expected during the next 5 years.

The clinical development department within the company has proposed a trial of parallel groups design to compare drug A with drug B. It has been designed to yield 90% power of detecting, at the 5% significance level, an arbitrarily chosen absolute difference of 30% based on endpoint X_1. Using the sample size formula for comparing two binomial proportions without correction for continuity and assuming equal allocation, 125 patients per treatment arm are required.[169] Withdrawals from the trial are not anticipated. The development department has proposed not to evaluate endpoint X_2 on the grounds that it would significantly increase the costs of the study. In this example, X_2 might be thought of as an outcome which requires for its measurement a complex and costly diagnostic procedure. However, the marketing department is uneasy with this proposal because market research has highlighted the importance of endpoint X_2 as a factor influencing product use. It has therefore

been decided to investigate the problem in more detail in order to determine an optimal trial strategy.

Within an investment appraisal framework, the determination of an optimal strategy would involve choosing between possible trial designs so as to maximise expected $NPV(n)$. In order to illustrate how this is achieved, it is necessary to impose specific functional forms for $C_t(n)$, Q_t^* and Q_t. It is therefore assumed that the company has determined the following cost function for the trial under consideration:

$$C_t(n) = v_t^P . Q_t(n) + f_t^T + v_t^T . n_t \qquad (15)$$

where $C_t(n)$ is a linear cost function with fixed and variable components. The variable marketing costs per unit of the product sold, v_t^P, will include any incremental costs of manufacturing, distributing and promoting the product as a direct consequence of conducting the trial. The trial fixed cost variables, f_t^T, will include the costs of researching, designing and planning the experiment, plus the costs of data analysis and reporting. Costs which vary with sample size, v_t^T, will include components such as expenditures on trial monitors, payments made to investigators for data collection and payments made to centres for the treatments given to patients during the trial.

Clearly, there may be uncertainties surrounding the estimation of some of the cost variables. For example, if company practice is to pay for the treatments given to patients during the trial as they occur, rather than as a fixed per patient payment set by contract in advance, it will be necessary to estimate the expected value of the payment per subject enrolled. This would be achieved by estimating a weighted

average cost per patient enrolled based on probabilities of different treatments occurring and their associated costs. Systems are available to assist companies in the production of such estimates.[172] However, the methods of estimation of the cost function variables are beyond the scope of this chapter. It is sufficient for the purpose of illustrating the approach to assume that the cost variables are either known with certainty, or alternatively represent expected values.

It is assumed that company market research has been conducted which has determined the following desired demand function:

$$Q_t^* = \alpha + \beta' \Delta X \ + \gamma \Delta P \qquad (16)$$

where β' denotes the row vector of demand function coefficients (β_1, β_2) for the potential endpoint variables $(\Delta X_1, \Delta X_2)$ in the vector ΔX. The sign and the size of estimated coefficient vectors, α, β' and γ, will determine the direction and magnitude of the change in Q_t^* resulting from changes in the demand function variables. It is further assumed that the company has determined that the expected diffusion process is best approximated by the following equation, which is an adaptation of the stock adjustment principle: [173]

$$Q_t = k.Q_t^*(\Delta X, \Delta P) + (1-k)Q_{t-1} \qquad (17)$$

where Q_t is the *actual* demand for drug A in time period t, and k is the coefficient of market adoption which takes on a value between 0 and 1. The achievement of the desired level of demand in the market is therefore expected to be gradual. The time taken to reach Q_t^* depends upon the size of k, which in practice will be determined

by complex relationships amongst the various factors which influence prescribing decisions. Some of these factors will be within the control of a company and others will not. Substituting for Q_t^* from equation (16) into equation (17) yields:

$$Q_t = k.\left(\alpha + \beta' \Delta X \ + \gamma \Delta P\right) + \left(1 - k\right) Q_{t-1} \qquad (18)$$

which is the assumed hypothetical dynamic demand function for drug A.

Substituting the above functional forms for $C_t(n)$ and Q_t into equation (13) gives the following specific formulation of the general model that is used to simulate the approach:

$$NPV(n) = \sum_{t=0}^{H} \frac{P_{At}\left\{k.\left(\alpha + \beta'\left\{\sum_{\Delta X=-100\%}^{100\%} \Pr\left(D_n^{\phi} = \Delta X | \Delta X\right) \Pr(\Delta X) \Delta X\right\} + \gamma \Delta P\right) + \left(1-k\right) Q_{t-1}\right\}}{\left(1+r\right)^t}$$

$$- \sum_{t=0}^{H} \frac{v_t^P\left(k.\left(\alpha + \beta'\left\{\sum_{\Delta X=-100\%}^{100\%} \Pr\left(D_n^{\phi} = \Delta X | \Delta X\right) \Pr(\Delta X) \Delta X\right\} + \gamma \Delta P\right) + \left(1-k\right) Q_{t-1}\right) + f_t^T + v_t^T . n_t}{\left(1+r\right)^t}$$

(19)

The hypothetical parameter values and assumptions for the model are summarised in Table 2.1 and the results are given in Table 2.2. The results are also presented graphically in Figures 2.1 and 2.2. All sample size results are taken to the nearest 25 subjects per treatment arm.

Table 2.1
Parameter Values and Assumptions

Parameter values	Assumptions
Cost function: $C_t(n) = v_t^P.Q_t(n) + f_t^T + v_t^T.n_t$	
v_t^P = \$ 40.00	Incremental costs of manufacturing, distributing and promoting the product as a direct consequence of conducting the trial. Assumed to be 20% of the selling price for product A.
$Q_t(n)$ = variable	Expected demand for product A in period t. It varies with trial sample size and endpoints measured, and is calculated by the model.
f_t^T = \$ 1,000,000	Fixed costs of planning the trial (\$ 750,000 incurred in year t = 0) and data analysis and reporting (\$ 250,000 incurred in year t = 1).
v_t^T = \$ 1,500 or v_t^T = \$ 1,875	Cost per patient = \$ 1,500 if only endpoint X_1 is evaluated and \$ 1,875 if only endpoint X_2 or both endpoints are evaluated. This cost includes trial monitors, payments to investigators, and the costs of treatment given to patients during the trial. The year in which costs are incurred depends upon the timing of subject enrolment, n_t .
Demand function: $Q_t = k.(\alpha + \beta_1 \Delta X_1 + \beta_2 \Delta X_2 + \gamma \Delta P) + (1-k)Q_{t-1}$	
Q_t = variable	Expected demand for product A in period t. The form of the demand function needs to be determined by empirical research.
k = 0.30	The coefficient of market adoption which is assumed to have been determined by empirical product diffusion research.
α = 0	The assumed value implies that drug A will not be used in the absence of comparative evidence of the product's benefits in relation to product B.
β_1 = 1000 and β_2 = 4995	The assumed values imply that X_1 is a less important determinant of demand than X_2. Coefficient values were chosen to yield 100% market share for product A if complete success was demonstrated for both endpoints.
$\Delta X_1 = X_{1A} - X_{1B}$ and $\Delta X_2 = X_{2A} - X_{2B}$	The expected demonstrated differences, $E_n(\Delta X)$, are calculated variables which depend upon the trial sample sizes and prior expectations. X_{1B} is assumed to have a baseline value of 25% and X_{2B} is assumed to have a baseline value of 75%.

Table 2.1 (continued)

Parameter values	Assumptions
Demand function (continued)	
γ = -550.00	The assumed value implies that treatment drug prices have a negative impact on the quantity of product A demanded.
$\Delta P = P_{At} - P_{Bt}$	P_{At} = \$200 is assumed to have been exogenously determined by a central pricing authority setting prices at parity with product B. P_{At} is assumed to be constant over the evaluation period.
Discounting	
r = 0.15	The projects risk adjusted discount rate assumed to have been determined using the capital asset pricing model.
$H = t_0 - t_4$	The time horizon for the investment appraisal is assumed to be 5 years.
Prior distributions, power and sample size	
$\Pr(\Delta X_1)$ =0.14(11%); (0.09)12%; 0.10(13%); 0.08(14%); 0.11(15%); 0.05(16%); 0.04(17%); 0.13(18%); 0.12(19%); 0.14(20%). $\Pr(\Delta X_2)$ = 0.14(1%); 0.11(2%); 0.05(3%); 0.11(4%); 0.11(5%); 0.05(6%); 0.05(7%); 0.10(8%); 0.14(9%); 0.14(10%)].	The hypothetical prior distributions for ΔX_1 and ΔX_2 were produced by using Lotus Excel to generate 100 random numbers with a uniform distribution of successful outcomes for X_{1A} in the range 35% - 45% and for ΔX_{2A} in the range 75% - 85%.
$\Pr(D_n^{\phi} = \Delta X \mid \Delta X)$	All sample size and power calculations were performed using the computer programme SAS assuming a two - tailed test with significance level ϕ = 5% and power = 90% as appropriate. The formula for comparing two binomial proportions without continuity correction and assuming equal allocation was used throughout.
n_t = variable	Sample size is varied for the simulations. The trial is assumed to take two years to design and complete, with subjects enrolled in equal proportions in years t = 0 and t = 1.

2.3.1 Is the proposed trial worth conducting?

Whether or not the hypothetical trial proposed by the clinical development department is worth conducting with regard to profit criteria is considered graphically in Figure 2.1. Figure 2.1 shows how the cost, revenue and profit functions behave with respect to changes in sample size per treatment arm, given the assumed parameter values of the model. As sample size increases, $PTC(n)$ increases at a constant rate reflecting the linear form of the hypothetical cost function. $PTR(n)$ also increases, but at a diminishing rate. This reflects the fact that the expected *additional* difference in ΔX_1 between products A and B (and hence, through the demand function, the expected incremental market share and expected incremental revenue) gets smaller and smaller as the trial sample size is increased. It can be seen that $PTR(n)$ is less than $PTC(n)$ for sample sizes below 50 and greater than 675 subjects per arm. Conversely, $NPV(n)$ is positive for any trial evaluating only endpoint X_1 within that sample size range. This profitable range is determined by the two points at which the parabolic $NPV(n)$ curve intersects the sample size axis.

In the case of the proposed trial with a sample size of 125 per arm, the present value of expected total revenue exceeds the present value of the total cost of conducting the trial, so the trial is worth conducting ($NPV(n) > 0$). It can be seen from Table 2.2 that a profit of $892,524 is to be expected (see Appendix 2.1 for an example calculation). However, the chosen sample size is sub - optimal since $NPV(n)$ continues to increase as sample size increases beyond 125 subjects per arm. The optimal sample size is that which yields the maximum $NPV(n)$. This can be determined graphically by identifying the point at which the marginal change in

$NPV(n)$ for a given change in sample size is declining and equals zero (i.e. $MNPV(n) = 0$). In this example, this occurs where the trial enrols 225 subjects per arm and yields an expected $NPV(n)$ of \$1,093,016. A trial with this sample size would have 90% power to detect an absolute difference of 12% with a two-sided significance test conducted at the 5% level.

2.3.2 What is the optimal trial strategy?

Suppose that the hypothetical firm can conduct only one trial, and that, if both X_1 and X_2 are evaluated, the same sample size must be used. The problem now is to determine the optimal trial design in terms of endpoint measurement and sample size. The optimal trial strategy will depend on whether or not the firm faces constraints imposed by factors such as the available budget or the number of potential trial subjects.

If there are no constraints, there are three options available: measure only X_1, only X_2 or both endpoints simultaneously. The determination of the optimal strategy is shown graphically in Figure 2.2 If only X_1 were to be evaluated, the optimal sample size is 225 subjects per arm, as indicated above. However, it would be preferable to evaluate only X_2 since this yields a maximum expected $NPV(n)$ of \$1,248,345 at the optimal ($NPV(n)$ maximising) sample size of 400 subjects per arm. Note that a single endpoint trial evaluating only X_1 is superior to a trial which measures only X_2 for sample sizes between 50 and 250 subjects per treatment arm. Conversely, a trial to evaluate only X_2 is preferred for sample sizes between 250 and 1000 subjects per treatment arm.

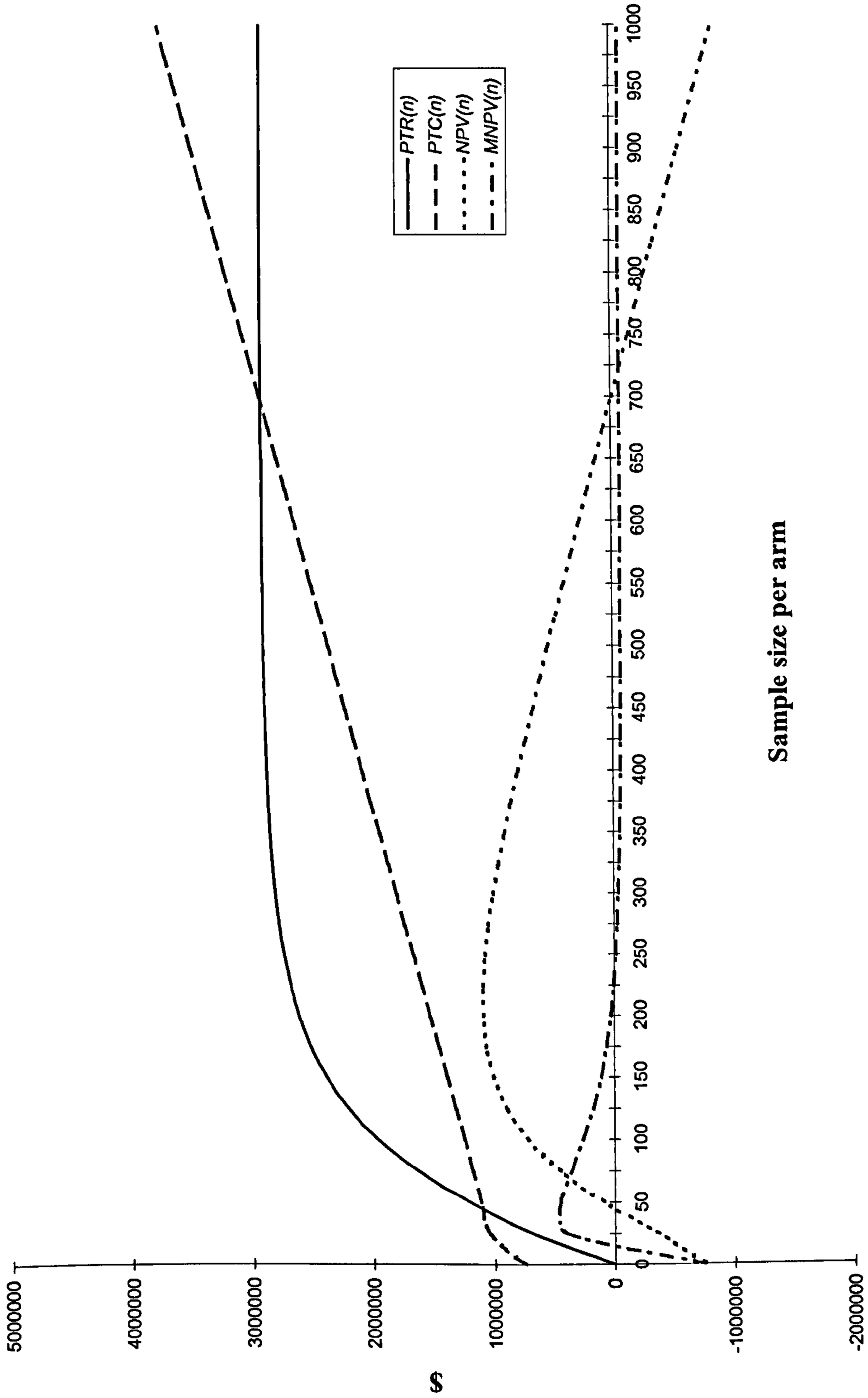


Figure 2.1: Cost and revenue curves for a trial measuring endpoint X_1

Table 2.2: Results from a Hypothetical Analysis

	Trial evaluating variable X_1 only						Trial evaluating variable X_2 only						Trial evaluating variables X_1 and X_2				
n per arm	$E_n(\Delta X_1)$	Q^*	$PTR(n)$	$PTC(n)$	$NPV(n)$	$MNPV(n)$	$E_n(\Delta X_2)$	Q^*	$PTR(n)$	$PTC(n)$	$NPV(n)$	$MNPV(n)$	Q^*	$PTR(n)$	$PTC(n)$	$NPV(n)$	$MNPV(n)$
0	0.00	0.00	0	750000	-750000	-750000	0.00	0.00	0	750000	-750000	-750000	0.00	0	750000	-750000	-750000
25	3.60	0.02	675616	1037500	-361884	388116	0.55	0.01	519773	1055027	-535255	214745	0.03	1195389	1055027	140362	890362
50	6.39	0.03	1198802	1107609	91194	453077	0.89	0.02	834264	1142663	-308399	226856	0.05	2033067	1142663	890404	750042
75	8.70	0.04	1630906	1177717	453188	361995	1.22	0.03	1141766	1230299	-88533	219865	0.07	2772671	1230299	1542372	651969
100	10.47	0.05	1963532	1247826	715706	262518	1.53	0.04	1438441	1317935	120506	209039	0.09	3401973	1317935	2084038	541666
125	11.79	0.06	2210459	1317935	892524	176818	1.83	0.05	1719916	1405571	314346	193840	0.10	3930376	1405571	2524805	440767
150	12.75	0.06	2391154	1388043	1003111	110587	2.12	0.05	1983608	1493207	490401	176055	0.12	4374762	1493207	2881555	356750
175	13.45	0.07	2522977	1458152	1064825	61714	2.38	0.06	2227503	1580842	646661	156260	0.13	4750480	1580842	3169638	288083
200	13.97	0.07	2619710	1528261	1091449	26624	2.61	0.07	2451328	1668478	782850	136189	0.14	5071038	1668478	3402560	232922
225	14.35	0.07	2691386	1598370	1093016	1567	2.83	0.07	2655167	1756114	899053	116203	0.14	5346553	1756114	3590438	187878
250	14.64	0.07	2745018	1668478	1076540	-16476	3.03	0.08	2839842	1843750	996092	97039	0.15	5584860	1843750	3741110	150672
275	14.85	0.07	2785542	1738587	1046955	-29584	3.21	0.08	3006570	1931386	1075184	79092	0.15	5792112	1931386	3860727	119617
300	15.02	0.08	2816489	1808696	1007793	-39162	3.37	0.08	3156413	2019022	1137391	62207	0.16	5972902	2019022	3953880	93154
325	15.14	0.08	2840176	1878804	961372	-46421	3.51	0.09	3291221	2106658	1184563	47172	0.16	6131397	2106658	4024739	70859
350	15.24	0.08	2858628	1948913	909715	-51657	3.64	0.09	3411943	2194293	1217650	33087	0.17	6270571	2194293	4076278	51539
375	15.32	0.08	2872830	2019022	853808	-55907	3.75	0.09	3520431	2281929	1238502	20852	0.17	6393261	2281929	4111331	35053
400	15.38	0.08	2883959	2089130	794828	-58979	3.86	0.10	3617911	2369565	1248345	9844	0.17	6501869	2369565	4132304	20973
425	15.42	0.08	2892717	2159239	733478	-61351	3.95	0.10	3705297	2457201	1248096	-249	0.18	6598014	2457201	4140813	8509
450	15.46	0.08	2899414	2229348	670066	-63412	4.04	0.10	3784271	2544837	1239434	-8662	0.18	6683685	2544837	4138848	-1965
475	15.49	0.08	2904796	2299457	605340	-64726	4.11	0.10	3855240	2632473	1222767	-16667	0.18	6760036	2632473	4127563	-11285
500	15.51	0.08	2908922	2369565	539357	-65982	4.18	0.10	3919524	2720109	1199416	-23351	0.18	6828447	2720109	4108338	-19225
525	15.53	0.08	2912211	2439674	472537	-66820	4.24	0.11	3977893	2807745	1170149	-29267	0.18	6890104	2807745	4082360	-25978
550	15.54	0.08	2914780	2509783	404998	-67540	4.30	0.11	4031008	2895380	1135628	-34521	0.19	6945789	2895380	4050408	-31952
575	15.55	0.08	2916781	2579891	336890	-68108	4.35	0.11	4079474	2983016	1096457	-39170	0.19	6996255	2983016	4013239	-37169
600	15.56	0.08	2918396	2650000	268396	-68494	4.40	0.11	4123751	3070652	1053099	-43359	0.19	7042147	3070652	3971495	-41744
625	15.57	0.08	2919611	2720109	199502	-68894	4.44	0.11	4164531	3158288	1006243	-46856	0.19	7084142	3158288	3925854	-45641
650	15.57	0.08	2920597	2790217	130380	-69122	4.48	0.11	4202214	3245924	956290	-49953	0.19	7122811	3245924	3876887	-48967
675	15.58	0.08	2921359	2860326	61033	-69347	4.52	0.11	4236882	3333560	903322	-52968	0.19	7158241	3333560	3824681	-52206
700	15.58	0.08	2921963	2930435	-8471	-69504	4.55	0.11	4268997	3421196	847802	-55520	0.19	7190961	3421196	3769765	-54916
725	15.58	0.08	2922431	3000543	-78112	-69641	4.58	0.11	4299021	3508832	790189	-57613	0.19	7221452	3508832	3712620	-57145
750	15.58	0.08	2922826	3070652	-147826	-69714	4.61	0.12	4327238	3596467	730770	-59419	0.19	7250064	3596467	3653597	-59024
775	15.58	0.08	2923069	3140761	-217692	-69866	4.64	0.12	4353368	3684103	669265	-61505	0.19	7276437	3684103	3592334	-61263
800	15.59	0.08	2923307	3210870	-287563	-69870	4.67	0.12	4378003	3771739	606264	-63001	0.19	7301310	3771739	3529571	-62763
825	15.59	0.08	2923492	3280978	-357486	-69924	4.69	0.12	4401331	3859375	541956	-64308	0.20	7324823	3859375	3465448	-64123
850	15.59	0.08	2923628	3351087	-427459	-69973	4.72	0.12	4423250	3947011	476239	-65717	0.20	7346878	3947011	3399867	-65581
875	15.59	0.08	2923759	3421196	-497437	-69977	4.74	0.12	4443894	4034647	409247	-66992	0.20	7367653	4034647	3333006	-66860
900	15.59	0.08	2923817	3491304	-567487	-70051	4.76	0.12	4463501	4122283	341218	-68029	0.20	7387318	4122283	3265035	-67971
925	15.59	0.08	2923895	3561413	-637518	-70031	4.78	0.12	4482164	4209918	272245	-68973	0.20	7406058	4209918	3196140	-68895
950	15.59	0.08	2923953	3631522	-707569	-70051	4.80	0.12	4499947	4297554	202393	-69852	0.20	7423900	4297554	3126346	-69794
975	15.59	0.08	2923982	3701630	-777649	-70080	4.82	0.12	4516925	4385190	131735	-70659	0.20	7440906	4385190	3055716	-70630
1000	15.59	0.08	2924010	3771739	-847729	-70080	4.83	0.12	4533046	4472826	60220	-71515	0.20	7457056	4472826	2984230	-71486

Figure 2.2: Determination of the optimal trial design

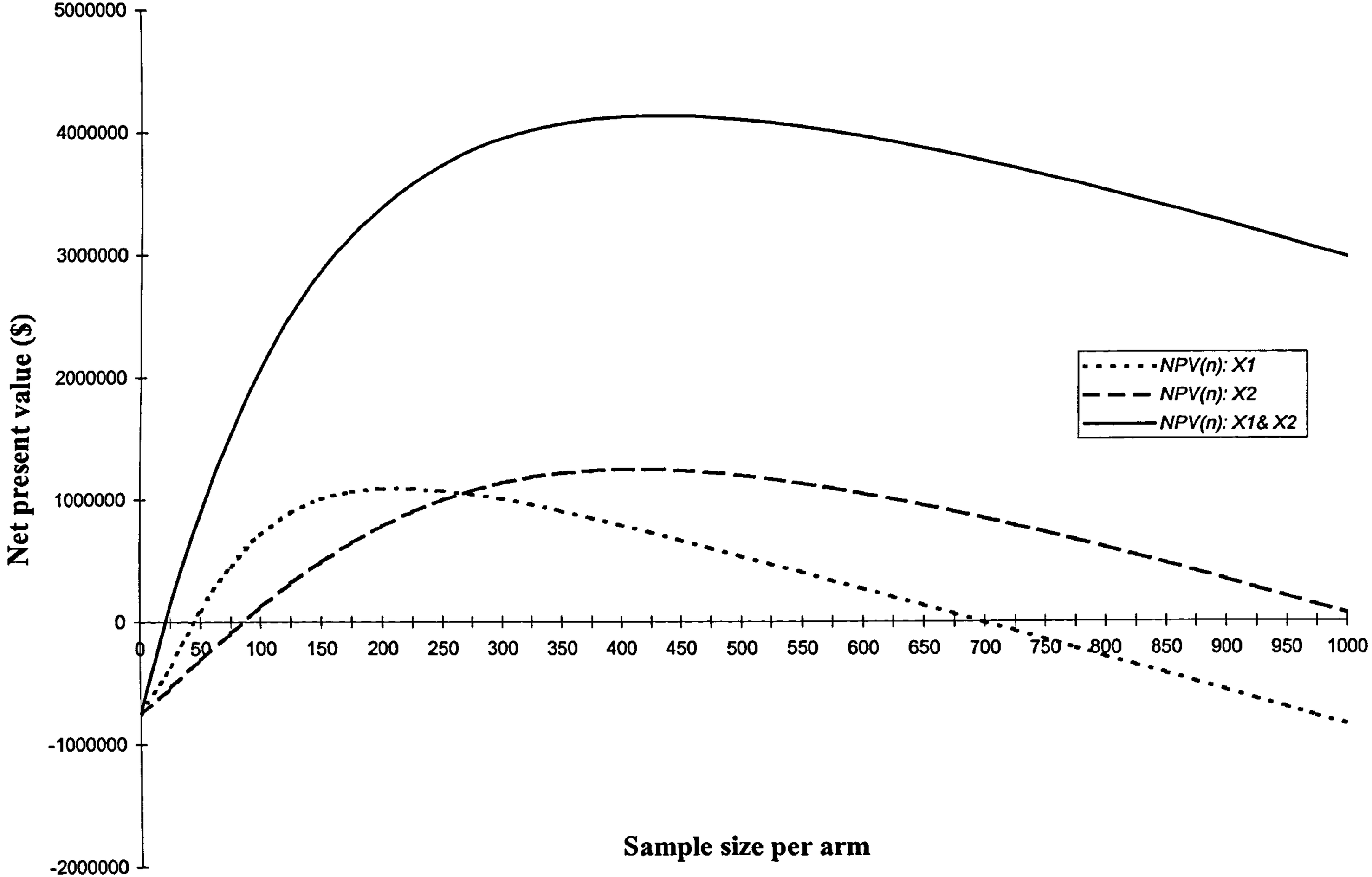


Figure 2.2 shows clearly that no single endpoint trial would be preferable to a trial which simultaneously evaluates both endpoints since the latter always yields an $NPV(n)$ greater than the former. Such trials would produce a positive expected $NPV(n)$ for sample sizes in the range 25 to 1975 subjects per arm. However, the optimal sample size is 425 subjects per arm, which yields an expected profit of \$4,140,813. The firm should therefore conduct a trial which evaluates both X_1 and X_2 at a sample size of 425 subjects per arm. It would have 90% power to detect a difference of 9% on both endpoints X_1 and X_2 with two-sided significance tests conducted at the 5% level.

2.3.3 Constrained optimisation

In practice, the firm might face budget or other constraints that prevent the achievement of the optimal trial design solution identified above. For example, suppose that the hypothetical trial is constrained by the availability of patients for enrollment and that a maximum of 300 subjects can be recruited. The optimal solution can be derived graphically by identifying the trial which gives the highest $NPV(n)$ curve for that sample size. In this case, a trial to evaluate both X_1 and X_2 would be the preferred strategy and would yield an $NPV(n)$ of \$3,953,880 as shown in Table 2.2.

In addition, at a macro level, a company will usually be faced with a capital rationing problem in which a limited investment budget needs to be allocated across competing trials involving different products. In this case, projects should first be ranked based on their profitability indices ($NPV(n)/PTC(n)$) and then selected for funding in

descending order of the index until the available budget is exhausted. Complex capital rationing problems can be solved using integer programming models (when fractional trial projects are not feasible) or linear programming models (where trial projects are divisible).[167]

2.4. DISCUSSION AND CONCLUSION

The model presented in this chapter shows how a pharmaceutical company could use an investment appraisal framework to assist with decisions taken in the design of randomised controlled trials. A hypothetical scenario facing a pharmaceutical company has been used to illustrate how profit criteria might be applied to decide whether a particular RCT is worth conducting, to determine an optimal (NPV maximising) design, and to rank RCTs in terms of their expected NPVs so as to select a portfolio of studies which maximises the return on a given development or trial budget. Whilst the simulations conducted illustrate decisions based on choices about sample size and endpoint measurements, the framework could be applied to choices concerning other trial design parameters, such as which comparators to include, the duration of patient follow-up and what power to use. A number of practical difficulties would need to be overcome in order to utilise the approach. These are highlighted below, together with some directions for future research.

Firstly, the investment appraisal approach to trial design has been illustrated using the NPV method of investment appraisal due to its recognised superiority over rival methods which can lead to incorrect solutions.[166-168] Irrespective of the method used, the approach requires individual RCTs to be viewed as appropriate units to be subjected to investment appraisal. Whilst the impact of a single trial, or small

collection of trials, on the use of a product might be open to question, it is nevertheless appropriate for a firm to evaluate their likely impact at the margin of available evidence. This poses a number of practical difficulties including the choice of time horizon, the determination of a trial's opportunity cost of capital, and how to assign capital, production, distribution, marketing and sales costs to individual trials. Clearly, these issues will depend upon a company's corporate goals, attitudes towards risk, accounting conventions and the specific market circumstances surrounding the development of a product.

Secondly, whilst the investment appraisal approach has been illustrated based on choices relating to the design of a hypothetical Phase IV trial, its potential value is by no means limited to that application. Pharmaceutical companies make their new product investment decisions at key milestones culminating in a decision whether or not to commit significant resources to the full development of a product. If full development proceeds, it involves deciding upon a programme of Phase III registration trials. These trials must demonstrate safety and efficacy only to a level that satisfies the regulatory authorities responsible for granting companies licences to market their products. However, as the debate surrounding the introduction of new drugs for the treatment of Alzheimer's disease highlighted, Phase III trials may fail to provide evidence of sufficient strength or relevance to persuade prescribing clinicians of the benefits of a product's use.[174-178] Thus, if only the minimum information required by regulatory authorities is provided at the time a product is introduced, a slow initial rate of diffusion (and hence low initial sales) might ensue until decision - makers' information needs have adequately been met by additional Phase IV (post marketing approval) trials.

The investment appraisal approach could therefore be used to evaluate the cost - revenue trade-offs associated with alternative late phase development strategies and to identify an NPV maximising portfolio of Phase III and Phase IV trial designs. Such an approach offers a potential improvement to the *ex ante* measurement of commercial "success" as used in a number of methods for assessing the value of pharmaceutical company research and development projects.[163;164] Clearly, an important area for future research is the extent to which Phase III trial designs are constrained by the requirements of regulatory authorities.

Thirdly, a potentially useful application for a company would be to use the approach to evaluate the incremental costs and benefits of incorporating non-clinical objectives into trials, such as quality of life research and economic evaluation studies. Such activities usually require significant additional resources to design and conduct. Within an investment appraisal framework, that additional effort can only be justified if the extra cost is more than offset by the incremental gain in sales revenue. A related important line of future research would be to compare private investment appraisal approaches with those which adopt a societal perspective.[143;149;152;154;155] Of primary interest would be to identify the conditions under which societal and industry objectives and perspectives produce similar recommendations. This will hinge on the relative importance of cost-effectiveness considerations in drug prescribing decision-making about which little is currently known. Within an investment appraisal framework, the latter would be determined through empirical demand analysis which is a central component of the investment appraisal approach.

Fourthly, whilst a dynamic form of demand function as illustrated in this chapter is undoubtedly more realistic than those implied in the models which have

adopted a societal perspective,[143;149;152;154] in practice significant empirical research would be required to define thoroughly the properties of a product's demand function and the associated diffusion process. This would need to be conducted within the context of specific projects, and would need to address questions such as: What is the role of clinicians versus other decision-makers in the product adoption process? What is the relative importance of different types of product differentiation data to different decision - makers? What is the nature of the relationship between differences in product characteristics and prescriber take - up? What factors influencing product diffusion are within the control of a company and which are not? Such research would probably highlight the heterogeneity of decision - making criteria and influences, different attitudes towards risk, the role of price and other treatment cost parameters in the demand function, and the lack of independence of some of the explanatory variables.

Conjoint analysis offers a promising way forward for investigating the properties of demand functions for use within an investment appraisal framework.[179;180] This technique is often used by pharmaceutical companies, although usually for the purposes of product pricing and positioning once Phase III trial results are known. Conjoint analysis could however be applied before RCT designs are finalised in order to assess the relative importance of study design attributes. The results of such research would need to be linked to product diffusion models to produce sales forecasts contingent upon the evidence expected to be provided by different trial designs. Whilst the literature on models concerning the diffusion of innovations is substantial,[181-183] their application to health care technologies has been limited to date.[184-188] This would be a valuable area for future research.

Fifthly, one of the fundamental components of the investment appraisal model as illustrated here is the use of elements of a Bayesian approach to the planning of RCTs.[189] Specifically, the investment appraisal approach necessitates the derivation and use of prior distributions for estimating the expected endpoint outcomes. The results of an investment appraisal analysis will be very sensitive to the choice and reliability of the prior distributions, hence the importance of carefully choosing the method used in their derivation. A good review article describes the different possible approaches, their strengths and limitations, and the sources of evidence for clinical priors.[189] In practice, one would probably adopt a number of approaches to derive a so - called "community of priors",[189] and test the implications for the results of an investment appraisal using sensitivity analysis.

A promising area for further research is to explore how applied demand analysis might be used to further advance the Bayesian approaches to trial design. For example, conjoint analysis could be used not only for evaluating the absolute and relative importance of the attributes entering a product's demand function, but also for eliciting the prior distributions for each of the attributes. Furthermore, such an approach could provide a formal framework for defining "meaningful" effect sizes based on decision - maker preferences for different types and strengths of evidence provided by a trial. Within an investment appraisal framework, a meaningful effect size might be defined as the minimum size of effect required to yield a positive NPV: the *commercially significant* effect size. The extent of pharmaceutical company use of Bayesian approaches is not known, although there is increasing recognition by regulatory authorities and the scientific community of the merits of the approach.[190]

Finally, and notwithstanding the simplifications and limitations of the

illustrative model presented here, a strength of the investment appraisal framework lies in the fact that it provides a rational, transparent and health care decision - maker focused basis for planning and designing RCTs. The framework views trials as an investment in information which will contribute to the nature and extent of product adoption through the strength and relevance of the information they produce. Within this context, the efficiency of a trial or programme of trials is viewed broadly in terms of market adoption and profit. This contrasts with alternative, narrowly focused goals such as the provision of the minimum amount of information necessary to secure marketing approval for a product within the shortest possible time frame. Clearly, the approach could only be utilised effectively if there was agreement between functions within a company to an explicit process for generating the necessary information, conducting the analyses required and acting upon the results. In some companies this might require a fundamental shift in culture away from a largely regulatory (product approval) driven organisation towards an integrated managerial economics approach to drug development decision - making.

APPENDIX 2.1

Example Calculation: NPV for Trial of Endpoint X_1 and Sample Size of 125 per arm

Present value of trial cost

$$PTC_{n=125} = \frac{\$750{,}000 + (125 \times \$1500)}{(1+0.15)^0} + \frac{\$250{,}000 + (125 \times \$1500)}{(1+0.15)^1}$$

$$= \$1{,}317{,}934.78$$

Expected trial outcome

$$E_{n=125}(\Delta X_1) = (.4714 \times .14 \times 11\%) + (.5367 \times .09 \times 12\%) + (.6007 \times .10 \times 13\%) + (.6618 \times .08 \times 14\%) + (.7188 \times .11 \times 15\%) + (.7706 \times .05 \times 16\%) + (.8165 \times .04 \times 17\%) + (.8562 \times .13 \times 18\%) + (.8896 \times .12 \times 19\%) + (.9171 \times .14 \times 20\%)$$

$$= 11.7851\%$$

Expected demand (% market share)

$$Q_{At}^{*} = 1000 \times 11.785\% = 11785\ (5.893\%)$$

$$Q_{A2} = 0.30 \times 11785 + (1 - 0.30) \times 0 = 3535\ (1.769\%)$$

$$Q_{A3} = 0.30 \times 11785 + (1 - 0.30) \times 3535 = 6010\ (3.006\%)$$

$$Q_{A4} = 0.30 \times 11785 + (1 - 0.30) \times 6010 = 7743\ (3.873\%)$$

Present value of expected total revenue

$$PTR_{n=125} = (\$200 \times 3535) / (1+0.15)^2 + (\$200 \times 6010) / (1+0.15)^3 + (\$200 \times 7743) / (1+0.15)^4$$

$$= \$2{,}210{,}335^{*}$$

Expected net present value

$$NPV_{n=125} = PTR_{n=125} - PTC_{n=125} = \$2{,}210{,}335 - \$1{,}317{,}934.78 = \$892{,}400^{*}$$

* There is a small discrepancy with the result reported in Table 2.2 due to rounding.

CHAPTER 3: USE OF RANDOMISED CONTROLLED TRIALS FOR PRODUCING COST EFFECTIVENESS EVIDENCE: POTENTIAL IMPACT OF DESIGN CHOICES ON SAMPLE SIZE AND STUDY DURATION

SUMMARY

A number of approaches to conducting economic evaluations could be adopted. However, some decision-makers have a preference for wholly stochastic cost-effectiveness analyses, particularly if the sampled data are derived from randomised controlled trials (RCTs). Formal requirements for cost-effectiveness evidence have heightened concerns in the pharmaceutical industry that development costs and times might be increased if formal requirements increase the number, duration or costs of RCTs. Whether this proves to be the case or not will depend upon the timing, nature and extent of the cost-effectiveness evidence required. The purpose of this chapter is to illustrate how different requirements for wholly stochastic cost-effectiveness evidence could have a significant impact on two of the major determinants of new drug development costs and times, namely RCT sample size and study duration. Using data collected prospectively in a clinical evaluation, sample sizes were calculated for a number of hypothetical cost-effectiveness study design scenarios and the results compared with a baseline clinical trial design. The sample sizes required for the cost-effectiveness study scenarios were mostly larger than those for the clinical baseline. Circumstances can be such that a wholly stochastic cost-effectiveness analysis might not be a practical proposition even though its clinical counterpart is. In such situations, alternative research methodologies would be required. For

wholly stochastic cost-effectiveness analyses, the importance of prior specification of the different components of study design is emphasised. However, it is doubtful whether all the information necessary for doing this will typically be available when product registration trials are being designed.

3.1 INTRODUCTION

During the last decade or so, economic evaluation has become an increasingly important part of the process of developing and marketing pharmaceutical products. This reflects the emergence, in a number of jurisdictions, of formal requirements for evidence about the cost-effectiveness of a new medicine.[10] The most recent policy move in this area is in the UK, where the National Institute for Clinical Excellence (NICE) has, *inter-alia*, been established to appraise the clinical- and cost-effectiveness of selected health interventions.[191] As O'Brien et al (1994) have pointed out, cost-effectiveness analyses can be wholly deterministic, wholly stochastic or a combination of the two.[31] Whilst economists accept each of these as valid approaches to performing economic evaluations,[39] some decision-makers may regard a wholly stochastic analysis as the preferred approach particularly if the sampled data are derived prospectively from appropriately designed randomised controlled trials (RCTs). For example, NICE previously indicated a preference for an RCT approach to producing cost-effectiveness evidence, and pointed out that pharmaceutical companies may therefore need to modify registration trials for this purpose.[191] More recent guidance from NICE still states a preference for RCTs as the means of generating effectiveness evidence for economic evaluation.[14] However, the current guidance no longer states a preference as to the source of resource use and cost data, probably due to the challenges faced in performing RCTs solely for cost-effectiveness analysis. Recent review papers and guidance to researchers highlight both the challenges involved as well as continued interest in using RCTs as vehicles for economic evaluation.[192-194]

Formal requirements for cost- effectiveness evidence at the time a new

product is introduced have heightened concerns in the pharmaceutical industry that development costs and times might be increased if these requirements increase the number, duration or cost of RCTs. However, whether this proves to be the case or not will depend upon the precise nature and extent of the cost-effectiveness evidence required from RCTs. Specifically, it will depend upon decision-makers' preferences for cost-effectiveness evidence defined in terms of the choices which need to be made about key RCT design attributes: comparators, population, setting, endpoints, effect sizes worth detecting, duration of observation and acceptable probabilities of Type I (α) and Type II (β) errors. Clearly, choices relating to these attributes have to be made at the planning stages of an RCT. Indeed, prior specification of the objectives, design characteristics and statistical properties are essential characteristics of a well conducted experiment.[1;5;170] The purpose of this chapter is to illustrate how different requirements for wholly stochastic cost-effectiveness evidence, defined in terms of different choices about RCT design, could have a significant impact on two of the major determinants of new drug development costs and times, namely RCT sample size and study duration.

In the next section, the choices which need to be made when designing an RCT are summarised together with an assessment of the potential impact which formal requirements for cost-effectiveness evidence generated in this way could have on study sample size and duration of observation. This is followed in section 3.3 by a description of the methods used to calculate sample sizes for a number of hypothetical cost-effectiveness analysis design scenarios. The illustrative scenarios were constructed using a dataset from a clinical evaluation which included the prospective collection of medical care resource utilisation data. The results of the sample size

calculations are presented in section 3.4. They illustrate how different formal requirements for cost – effectiveness evidence, expressed in the form of alternative research design choices, have the potential to increase the size and duration of RCTs significantly compared with those required for a baseline 'clinical' RCT. This in turn means that such requirements have the potential to increase the costs and timelines associated with new product development. In the final section, the implications of the results are discussed.

3.2 RANDOMISED CONTROLLED TRIAL DESIGN ATTRIBUTES AND SAMPLE SIZE CALCULATION

An RCT, whether conducted for the purposes of performing a clinical evaluation, an economic evaluation or both, cannot be designed adequately without explicit choices being made about a number of key study design attributes. The principal attributes requiring prior specification are summarised in the first column of Table 3.1. Choices made in relation to each of these attributes will have implications for both the size and duration of a trial. The second column of Table 3.1 shows some of the potential modifications to clinical trials which health care decision-makers might like to see for the purpose of conducting a cost-effectiveness study. The potential impact which such requirements might have on study sample size and duration of observation, and hence on development costs and timelines, are also shown in the third column of Table 3.1.

The primary concern in this chapter is to illustrate how different requirements for cost-effectiveness evidence might impact the size and duration of an RCT compared with the size and duration of a study conducted solely for the purpose of performing a

clinical evaluation. Given this objective, approaches to sample size determination for both clinical and economic evaluations are considered briefly in turn below.

3.2.1 Sample size calculation for clinical evaluation

In the design of a clinical RCT, it is usual to select one 'primary' endpoint which is used as the principal criterion for comparing the relative merits of the alternative treatments under evaluation. The primary endpoint usually serves as the basis for all subsequent sample size calculations. In order to calculate the required sample size, researchers must choose both the magnitude of the endpoint difference which is deemed to be worth detecting and the acceptable probabilities of Type I (α) and Type II (β) errors. The magnitude of endpoint difference worth detecting requires a decision to be taken about what constitutes a 'clinically meaningful' treatment effect. Typically, the significance level α is set by convention at 5% i.e. a clinical trial is usually designed such that there is a small probability of concluding that there is a difference between the treatments being compared when in fact there is no difference. Similarly, a clinical trial is usually designed with β (the probability of wrongly concluding that there is a difference between treatments) set by convention at either 10% or 20%. The latter is equivalent to setting the statistical power ($1-\beta$) of the test of a trial at 90% or 80% respectively. The precise sample size formula which is appropriate for a given RCT will depend primarily on how the endpoints of interest are to be measured and the methods of statistical inference which are relevant for the data. There is an extensive literature relating to sample size calculations for clinical experiments. Formulae are available for a wide variety of clinical trial designs, types of data and methods of statistical analysis.[170;195;196]

Table 3.1
Trial Design Attributes: Potential Impact on Sample Size, Study Duration and Costs of Modifications for Economic Evaluation

Trial Design Attributes*	Example Modifications to RCTs for Cost-effectiveness Analysis	Potential Implications for Study Sample Size, Duration and Cost
1. Comparators Can be chosen from one or more broad types, including: i) Placebo ii) Most commonly used iii) Most effective iv) Least cost v) Most cost-effective. Specification usually involves the choice of specific product formulations and modes of administration. Most studies compare two treatments although more are possible.	• Replacing a placebo comparator with standard practice • Adding a standard practice arm as well as placebo	• Sample sizes will be larger hence costs will increase • Adding a third treatment arm to a trial would increase total sample size, hence costs will increase • In both cases, enrolment times might be extended depending upon the availability of trial subjects hence development timelines will be longer
2. Population Specification usually involves choices about: i) Age group ii) Sex iii) Ethnic origin iv) Disease stage v) Co-morbidities vi) Previous treatments vii) Concomitant treatments viii) De novo or refractory patients. ix) Sub-group comparisons	• Relaxing trial entry criteria to generate a more typical study population e.g. allowing entry to patients with co-morbidities	• Sample sizes will be larger hence costs will increase • The trial could be more complex to manage hence costs will increase • Enrolment times might be extended depending upon the availability of trial subjects hence development timelines will be longer
3. Setting Specification usually involves choices about: i) Single country, single centre ii) Single country, multi-centre iii) Multinational, single centre iv) Multinational, multi-centre v) Inpatient vi) Outpatient vii) Specialist centre viii) Routine practice centre	• Single setting, centre or country data required	• Enrolment times might be extended depending upon the availability of trial subjects hence timelines are longer

Table 3.1 (continued)

4. Endpoints Specification usually involves choices about: i) Efficacy ii) Effectiveness iii) Side effects iv) Adverse events v) Quality of life vi) Direct costs (NB includes product prices) vii) Indirect costs viii) Resource use ix) Surrogate endpoints	• Adding endpoints for economic analysis	• Increases the amount of data to be collected and analysed which increases costs for a trial of a given sample size
5. Effect sizes Specification usually involves choices about: i) Clinical significance ii) Statistical significance iii) Primary endpoints iv) Secondary endpoints	• Testing cost-effectiveness hypotheses based on predefined levels of willingness to pay for a unit increase in effectiveness	• Increasing the number of endpoints for hypothesis testing increases sample sizes and hence costs
6. Duration of observation Specification usually involves choices about: i) Fixed period of observation ii) Variable (e.g. in sequential designs). Choices are linked closely to the choice of endpoints and the statistical properties of the study.	• Longer periods of follow-up required to evaluate final economic outcomes	• Sample sizes will be larger to ensure sufficient patients complete the study hence costs will increase • Increases the amount of data to be collected and analysed which increases costs for a trial of any sample size • Period of evaluation is longer hence development timelines will be longer
7. Acceptable error rates: α and β Choices are linked closely to the choice of endpoints and the duration of follow-up. Often chosen according to convention and based on the primary endpoint(s) i.e. $\alpha = 5\%$, $\beta = 10\%$ Do not have to be the same for each endpoint (and usually aren't). Used in conjunction with the statistical properties of endpoints, the desired effect sizes and withdrawal rates to determine sample size.	• Conventional probabilities of Type I and Type II error applied to additional (economic) parameters	• Increasing the number of endpoints for hypothesis testing increases sample sizes and hence costs

Table 3.1 (continued)

8. Statistical methods Specification usually involves choices relating to: i) Objectives of the trial ii) Nature of other trial parameters, most notably the disease area and endpoints (type of data) iii) Method of randomisation.	• Hypothesis tests to be performed on all variables and sub-groups of interest at conventional levels of statistical significance	• Sample sizes will be larger hence costs will increase • Enrolment times might be extended depending upon the availability of trial subjects hence timelines will be longer

* Decisions pertaining to these attributes are made by those designing the trial although the choices may be constrained by factors such as regulatory requirements, ethical considerations, patient availability and budget.

3.2.2 Sample size calculation for economic evaluation

In contrast to the design of an RCT conducted for the purpose of performing clinical evaluations, the primary endpoint of interest in an economic evaluation is a ratio. Specifically, economists are interested in estimating the incremental cost-effectiveness ratio (ICER), which is given by:

$$ICER_{AB} = \frac{C_A - C_B}{E_A - E_B} = \frac{\Delta C_{AB}}{\Delta E_{AB}} \qquad (1)$$

where ΔC_{AB} and ΔE_{AB} denote, respectively, the differences in the average cost $(C_A - C_B)$ and the average effectiveness $(E_A - E_B)$ between two treatments, A and B. The economic evaluation analogue to a clinically meaningful difference is a critical threshold value of the ICER (denoted R_C) below which a treatment is deemed to be cost-effective. In other words, treatment A is cost-effective compared against treatment B and should therefore be implemented if:

$$ICER_{AB} < R_C \text{ , for } \Delta E_{AB} > 0 \text{ or } ICER_{AB} > R_C \text{ , for } \Delta E_{AB} < 0 \qquad (2)$$

Whereas the literature pertaining to sample size calculations for clinical experiments is substantial, it is only relatively recently that methods for determining the sample sizes required for performing cost-effectiveness analyses have been proposed.[49;54;65;68;70] Al et al (1998) have used a simulation approach[65], Briggs and Gray (1998) present a parametric formula which ignores covariance[68], and the work of Laska et al (1999) includes formulae based on non-parametric tests.[49] The method used to perform the illustrative sample size calculations reported in this chapter is a parametric approach developed independently by Briggs and Tambour

(1998,2001,2002),[70;197;198] Laska et al (1999)[49] and Gardiner et al (2000)[54] (see formula in the methods section below). Essentially, these authors base their approach on the net benefit (NB) formulation of the cost-effectiveness decision rule set out in equation (2) above due to the advantageous properties of the NB statistic for the analysis of sampled cost-effectiveness data.[66;199] Specifically, the NB approach states that a new treatment is cost-effective and should be implemented if:

$$NBC_{AB} = R_C(\Delta E_{AB}) - \Delta C_{AB} > 0 \qquad (3)$$

or, equivalently,

$$NBE_{AB} = \Delta E_{AB} - \frac{1}{R_C}\Delta C_{AB} > 0 \qquad (4)$$

where NBC_{AB} and NBE_{AB} are the net benefits associated with treatment A expressed in units of money and effectiveness respectively. Using the NB approach, Briggs and Tambour (1998,2001,2002), Laska et al (1999) and Gardiner et al (2000) develop a sample size formula for detecting whether the NB associated with an intervention is positive given acceptable probabilities of Type I (α) and Type II (β) errors.[49;54;70;197;198] It should be noted at this juncture that, unlike clinical evaluations, there is currently no consensus amongst economists about the levels of acceptable errors for use in cost-effectiveness analysis. For the purpose of the analysis presented here, it is assumed that the conventions for clinical RCTs apply.

3.3 METHODS

The primary concern in this chapter is with how different requirements for cost-effectiveness evidence might impact the size and duration of an RCT compared

with the size and duration of a study conducted solely for the purposes of a clinical evaluation. To illustrate these differences, the methods used to calculate sample sizes for a number of hypothetical cost-effectiveness analysis design scenarios are described in this section.

3.3.1 Sources of data

The dataset used for the illustrative analyses presented in this chapter comes from a clinical evaluation which included the prospective collection of medical care resource utilisation data. This dataset has previously been used in other economic evaluation methodology research.[69;86;91;94] The data come from a study which compared the effectiveness of a new drug against placebo for the treatment of a chronic condition for which there is currently no known cure. The primary endpoint was survival, and subjects were followed for four years i.e. long-term survival and medical care resource utilisation histories are available. The medical care resource utilisation variables for which data were collected include: number of days spent in hospital, number of outpatient consultations and number of attendances at a day hospital. For the purpose of this illustrative analysis, the resource use variables have been valued using unit cost data taken from the UK NHS Management Executive database.[200] The actual cost values used are shown in the footnotes to the tables which follow. In addition, data were collected on a clinical variable which is now used as the primary basis for classifying patients in terms of the severity of their disease at diagnosis. The salient descriptive statistics pertaining to the clinical evaluation dataset are presented in Table 3.2. These are presented to highlight the fact that the nature of the dataset permits alternative hypothetical cost-effectiveness study scenarios to be constructed based on the RCT design attributes summarised in Table 3.1. The

construction of the scenarios is discussed below.

3.3.2 RCT design scenarios

To date, researchers investigating the methods for the analysis of sampled economic data and the methods for determining sample sizes for cost-effectiveness studies have ignored the potential importance of a number of key attributes pertaining to the design of an RCT. In order to emphasise the importance of this omission, equation (3) can be re-written as follows:

$$NBC_{ABSt} = R_C \cdot \left(\sum_{t=1}^{H} \frac{E_{ASt} - E_{BSt}}{(1+r_E)^t} \right) - \left(\sum_{t=1}^{H} \sum_{i=1}^{M} \frac{P_i X_{iASt} - P_i X_{iBSt}}{(1+r_C)^t} \right) \qquad (5)$$

where NBC_{ABSt} is the net-benefit, expressed in monetary terms, of treatment A compared with treatment B in population S evaluated over a time-period of t years. In the illustrative calculations presented in this chapter, the effectiveness of the alternative treatments being compared (E_{ASt}, E_{BSt}) is measured in terms of the average years of survival. The cost part of the net-benefit equation has been broken into its two principal components, namely the average quantities of different types of medical resources (denoted by X_i) and their associated unit costs (denoted by P_i). The subscripts E,C on the discount rate r indicate the fact that the effectiveness and cost outcomes can be discounted at different rates. However, whilst differential discounting of costs and effects appears to be favoured in the UK,[191] use of the same rate ($r_E = r_C$) seems to be the most commonly recommended approach.[201]

Table 3.2
Key Features of the Clinical Evaluation Dataset

Variable	All Subjects		A. Advanced Disease		B. Early Disease	
	A n = 81	B n = 67	A n = 52	B n = 46	A n = 29	B n = 21
Costs (£)						
Year 1	16,557 (14,725)	17,288 (13,763)	16,494 (13,693)	19,298 (14,454)	16,671 (16,671)	12,887 (11,197)
Year 2	24,035 (16,049)	20,654 (15,170)	24,754 (16,109)	22,316 (16,298)	22,746 (16,141)	17,012 (11,895)
Year 3	27,398 (18,297)	21,182 (15,192)	27,525 (19,427)	22,586 (16,110)	27,170 (16,402)	18,105 (12,776)
Lifetime	27,640 (18,318)	21,252 (15,186)	27,583 (19,416)	22,667 (16,085)	27,741 (16,495)	18,152 (12,813)
Survival (Years)						
Year 1	0.95 (0.17)	0.72 (0.34)	0.94 (0.20)	0.71 (0.34)	0.97 (0.10)	0.74 (0.33)
Year 2	1.58 (0.50)	1.06 (0.71)	1.52 (0.52)	1.01 (0.69)	1.70 (0.47)	1.16 (0.76)
Year 3	1.91 (0.85)	1.27 (1.03)	1.76 (0.82)	1.21 (1.01)	2.18 (0.85)	1.41 (1.09)
Year 4	1.98 (0.94)	1.34 (1.15)	1.80 (0.88)	1.26 (1.11)	2.29 (0.97)	1.51 (1.24)
Inpatient days						
Year 1	86.60 (85.30)	96.85 (77.85)	85.87 (79.55)	108.35 (81.64)	87.93 (96.21)	71.67 (63.53)
Year 2	123.69 (89.03)	115.36 (85.15)	127.60 (86.42)	125.15 (91.47)	116.69 (94.68)	93.90 (66.34)
Year 3	140.88 (100.31)	118.10 (85.15)	142.10 (104.45)	126.59 (90.47)	138.69 (94.18)	99.52 (70.61)
Year 4	141.98 (100.06)	118.39 (85.10)	142.31 (104.41)	127.00 (90.35)	141.38 (93.52)	99.52 (70.61)
Day patient attendances						
Year 1	1.44 (3.17)	0.30 (1.02)	1.54 (3.29)	0.22 (0.87)	1.28 (2.99)	0.48 (1.29)
Year 2	3.60 (9.33)	0.45 (1.50)	3.71 (10.45)	0.30 (1.03)	3.41 (7.04)	0.76 (2.21)
Year 3	4.12 (10.31)	0.48 (1.68)	4.00 (10.80)	0.30 (1.03)	4.34 (9.54)	0.86 (2.59)
Year 4	4.22 (10.55)	0.49 (1.68)	4.02 (10.81)	0.30 (1.03)	4.59 (10.25)	0.90 (2.59)
Outpatient attendances						
Year 1	11.19 (11.07)	2.51 (3.70)	11.92 (11.59)	2.50 (3.40)	9.86 (10.11)	2.52 (4.38)
Year 2	16.78 (15.16)	3.61 (5.83)	17.17 (15.52)	3.11 (4.37)	16.07 (14.73)	4.71 (8.21)
Year 3	19.02 (17.11)	4.15 (6.75)	18.63 (17.05)	3.35 (4.42)	19.72 (17.50)	5.90 (10.08)
Year 4	19.22 (17.42)	4.39 (7.50)	18.73 (17.18)	3.46 (4.49)	20.10 (18.12)	6.43 (11.57)

Values are sample means with standard deviations shown in parentheses. Mean cost figures are derived using the following unit costs: drug cost = £250 per annum; Inpatient = £176.07 per day; Outpatient = £73.06 per consultation; Day case = £176.07 per attendance.

It should be noted that although the calculations which follow are based on the arithmetic means of effectiveness and costs for each of the time periods for which analyses are performed, the sample averages are not necessarily the appropriate estimators to use in the presence of censored data.[48;76;86;87] However, the use of arithmetic means does not invalidate the main arguments or conclusions in this chapter.

It can be seen from equation (5) that the net benefit demonstrated for a given intervention depends upon the choices made about the various study design attributes summarised in Table 3.1. Consequently, the sample size determined by applying any formula based on the net benefit statistic will also be a function of the choices about RCT design attributes. Despite the existence of a large number of methodological guidelines, these tend to be at a general level.[10] Consequently, there is currently considerable discretion in the choice of design parameters, although arguably the choices would reflect the preferences of the ultimate consumers of the information.

In order to illustrate the impact which different design preferences can have on study sample size and duration, hypothetical RCT design scenarios were constructed based on different assumed choices regarding the following components of equation (5):

- duration of observation, t years
- medical care resource components measured in the RCT, X_i
- medical care resource unit costs, P_i
- discount rates applied to costs (r_C) and effectiveness (r_E)

- study population based on sub-groups of patients with different disease severity at diagnosis, S.

The sample sizes for each of the design scenarios defined on the above variables are compared against the sample size calculated for a hypothetical baseline clinical study of 1 year duration (see below). The impact on sample size of different design choices is evaluated across a range of critical threshold values of the ICER (R_C).

3.3.3 Sample size formulae

The parametric method of sample size determination set out by Briggs & Tambour (1998,2001,2002)[70;197;198] , Laska et al (1999)[49] and Gardiner et al (2000)[54] was used to calculate the sample sizes required for the different RCT design scenarios that were created in the manner described above. Based on the decision rule that a treatment should be implemented if NB is significantly positive, the above authors report formulae for testing the statistical hypothesis that

$$H_0 : NBC_{AB} \leq 0$$

versus the alternative hypothesis

$$H_1 : NBC_{AB} > 0.$$

Following the notation used by Briggs & Tambour (1998,2001,2002), it can be shown that the sample size required to test the above hypothesis is given by:

$$n > \frac{(z_\alpha + z_\beta)^2 \left[R_C^2 (\sigma_{EA}^2 + \sigma_{EB}^2) + (\sigma_{CA}^2 + \sigma_{CB}^2) - 2R_C \rho \sqrt{(\sigma_{EA}^2 + \sigma_{EB}^2)(\sigma_{CA}^2 + \sigma_{CB}^2)} \right]}{(R_C \Delta E_{AB} - \Delta C_{AB})^2} \qquad (6)$$

where n is the number of subjects required per arm of the trial, Z_α and Z_β denote the critical values from the standard normal distribution corresponding to the required significance level and power respectively, $\sigma^2_{EA}, \sigma^2_{EB}, \sigma^2_{CA}, \sigma^2_{CB}$ are the variances of effects (E) and costs (C) for treatments A and B, and ρ is the correlation between ΔC_{AB} and ΔE_{AB}. Note that equation (6) could be rewritten using the notation introduced in equation (5) in order to emphasise the point that sample size is a function of various RCT design attributes. However, this has not been done here in order to simplify the presentation. The formula for estimating ρ from the pilot data (which has unequal sample sizes) is given by:

$$\frac{\frac{\rho_{CEA}.\sigma_{CA}.\sigma_{EA}}{n_A} + \frac{\rho_{CEB}.\sigma_{CB}.\sigma_{EB}}{n_B}}{\sqrt{\frac{\sigma^2_{CA}}{n_A} + \frac{\sigma^2_{CB}}{n_B}} \times \sqrt{\frac{\sigma^2_{EA}}{n_A} + \frac{\sigma^2_{EB}}{n_B}}} \tag{7}$$

where ρ_{CEA} and ρ_{CEB} are the correlation coefficients between costs and effects for treatment groups A and B respectively, and n_A and n_B are the corresponding sample sizes. Readers interested in seeing a detailed derivation of the sample size formula should see Briggs & Tambour (1998,2001,2002), Laska et al (1999) or Gardiner et al (2000).[49;54;70;197;198] A particularly useful aspect of the paper by Gardiner et al (2000) is that they present a formula for determining the ratio of the sample size requirements for tests of hypotheses on the ICER to the requirements for testing differences in effectiveness.[54]

To calculate the sample size required for the baseline clinical study of 1 year duration, the formula for a two sample test of the equality of means given by Rosner (2000)

was used[196]:

$$n_A = \frac{\left(\sigma_{EA}^2 + \sigma_{EB}^2 / r\right)\left(z_{\alpha/2} + z_\beta\right)^2}{E_A - E_B} \text{ and } n_B = rn_A \qquad (8)$$

where $r = 1$ for a trial with equal allocation to each arm. This formula yields the same results as that cited in Gardiner et al (2000).[54]

No allowance is made for withdrawals in any of the sample size calculations although it is common practice to make adjustments for this eventuality.[195] This does not affect the conclusions of the analysis, although it does mean that all sample sizes are probably lower than would be used in practice. All calculations utilise levels of acceptable errors typically used in RCTs, namely $\alpha = 5\%, \beta = 10\%$.

3.4 RESULTS

In order to illustrate how different requirements for wholly stochastic cost-effectiveness evidence might impact the size and duration of an RCT, it is necessary to choose a baseline RCT design where the sample size has been determined based on a 'clinical' primary endpoint. Clearly, any statements about incremental sample sizes and development times attributable to requirements for cost-effectiveness analysis will depend upon the choice of baseline design. For this analysis, the data from the original clinical evaluation are viewed as pilot data assumed to be available for designing an RCT. These were used to calculate the sample size required to compare the two treatments in terms of average survival times after 1 year of follow up for a trial in a population of subjects with mixed disease severity. Using the formula given in equation (8) and assuming equal allocation, 29 subjects per arm would be required

making no allowance for withdrawals from the study.[196] The formula presented by Gardiner et al (2000) yields the same sample size.[54] Thus, a total of 58 subjects and a trial of 1 year duration is used as the baseline design for the comparisons which follow. The sample sizes required for a similar trial of 2, 3 or 4 years duration would be 60, 92 and 114 subjects respectively. Using the formulae presented in equations (6) and (7), sample size calculations were performed for the cost-effectiveness RCT design scenarios described above. The results of these calculations are presented in Tables 3.3 to 3.7, which shows the total sample sizes compared with the clinical baseline. The results are presented in terms of each of the design attributes and are discussed in turn below.

3.4.1 Choice of duration of observation

Rows numbered 2 to 4 of Table 3.3 show the sample size implications of testing economic hypotheses based on different periods of follow-up (1 to 3 years) and for different assumed critical values of the ICER (£5,000, £10,000, £15,000, £20,000 and £25,000). These calculations are based on the study population which includes a mix of patients with both advanced and early disease i.e the same population as for the baseline clinical design. It can be seen that for each economic analysis scenario (1, 2 and 3 - years follow-up), sample sizes are higher than those which would be required for the clinical baseline design (row 1). For the critical ICER values used in Table 3.3, the sample sizes required for economic analysis are a decreasing function of the critical value of the ICER, and the ICER (and hence net benefit) is an increasing function of the duration of subject follow-up. For example, based on an analysis at 1 year and a critical ICER of £10,000 a trial which enrolled a total of 760 subjects (compared with 58 for the baseline clinical analysis) would be required to test the hypothesis that net benefits are statistically significantly positive. In contrast, a 3-year study would require a sample size of 344,229 subjects to test the hypothesis that net benefits are statistically significantly positive.

Table 3.3
Sample Sizes for an RCT when different economic follow-up periods are used

RCT Design Attribute	ρ [a]	ICER[b]	Total Sample Size for Critical ICER of :[c,d]				
			£5,000	£10,000	£15,000	£20,000	£25,000
1. Clinical baseline		- £3,193	58	58	58	58	58
2. Economic: 1 year	0.104	- £3,193	1958	760	410	262	188
3. Economic: 2 years	0.077	£6,447	*	2634	522	248	160
4. Economic: 3 years	0.087	£9,711	*	344229	1322	462	276

a ρ = the correlation between the difference in costs and effects of treatments A and B

b Incremental cost-effectiveness ratio calculated from the data. These values are the hypothesised ICERs for the sample size calculations.

c Probablities of Type I error = 0.05 and Type II error = 0.10 throughout. All sample size calculations take into account the correlation between cost and effectiveness differences using the formula cited in the text.

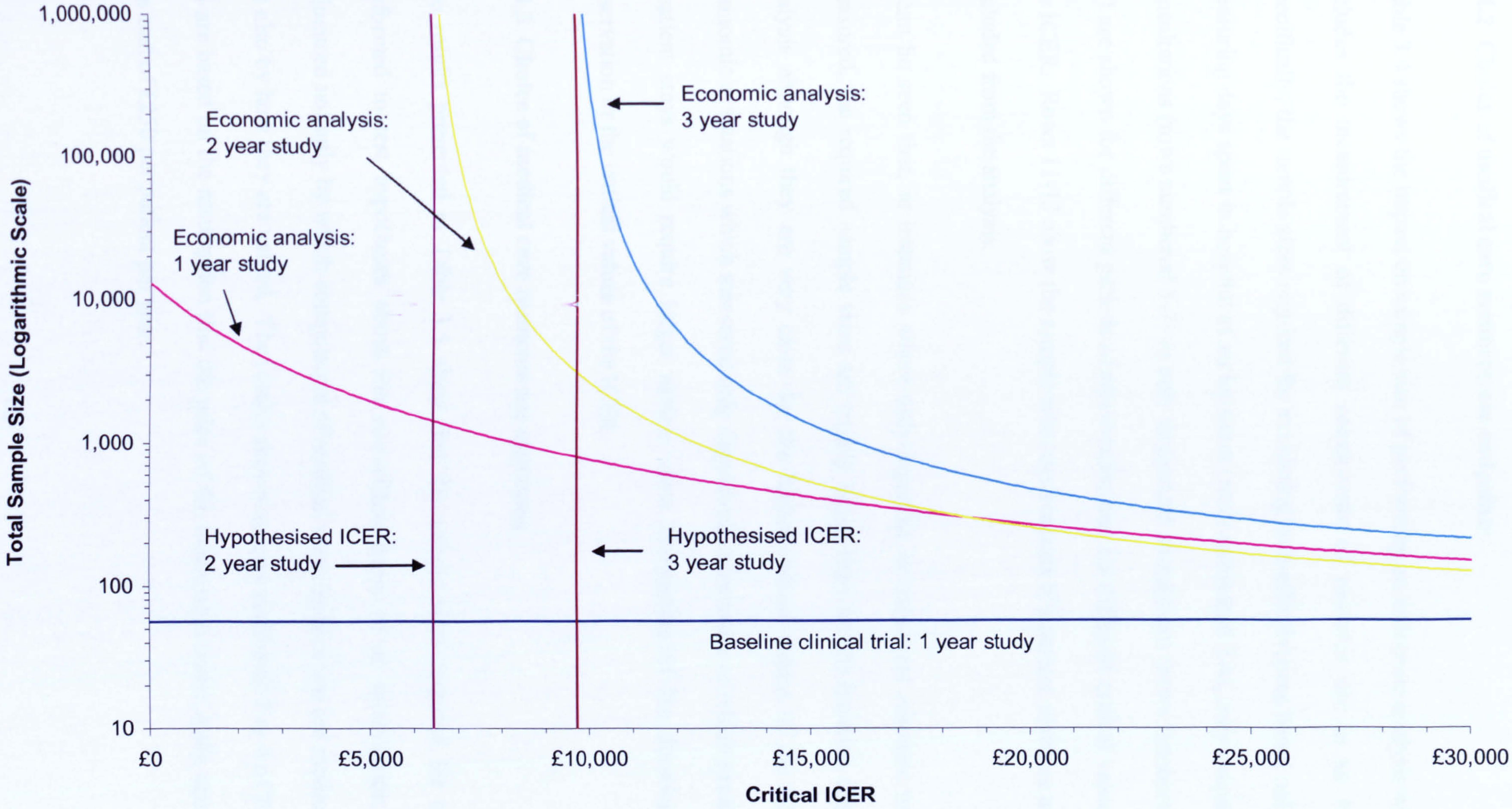
d All calculations are based on undiscounted costs and effects

* Denotes the fact that the sponsor of a technology under investigation would probably not conduct a study to test economic hypotheses if the postulated ICER was greater than the critical value.

The impact on sample size of choosing alternative periods of observation and alternative critical values of the ICER for the purposes of evaluating cost-effectiveness is also illustrated graphically in Figure 3.1. Figure 3.1 shows sample sizes (logarithmic scale) plotted against different critical values of the ICER (£0 to £30000) for the clinical baseline RCT design, and for trial designs which evaluate economic outcomes after 1, 2 and 3 years of observation. The results presented in Figure 3.1 show clearly that sample sizes are larger for each possible economic design scenario (combination of duration and critical values of the ICER) compared with the clinical baseline design, although they converge as the critical value of the ICER increases.

Figure 3.1 shows clearly that study designs which bring the hypothesised ICER and the critical value into close proximity with one another will render a standard frequentist stochastic cost-effectiveness analysis unattainable although in the examples shown here, a standard clinical evaluation would still be a practical proposition. From the pilot data, the hypothesised values for the studies of 2 and 3 years duration are £6,447 and £9,711 respectively. In these cases, the sample sizes required for the cost-effectiveness analyses can be seen to tend to infinity as the critical value of the ICER approaches the hypothesised value.

Figure 3.1: Sample Size Implications of Choosing Different Study Durations



3.4.2 Choice of medical care resource use endpoints

Table 3.4 shows the impact on sample size of performing an economic analysis which includes the measurement of different components of resource use in an RCT. Specifically, the sample sizes required for evaluating cost-effectiveness based only on measuring days spent in hospital as an inpatient (rows numbered 2-4), only outpatient consultations (rows numbered 5-7) or only daypatient attendances (rows numbered 8-10) are shown for different periods of observation and for different critical values of the ICER. Rows 11-13 show the sample size requirements if inpatient resource use is excluded from the analysis.

It can be seen that, in instances where only inpatient or outpatient resource use is measured, the required sample sizes are mostly larger than for the baseline clinical analysis although they are very close for the higher critical values of the ICER. Economic evaluations which measured only daypatient attendances or which excluded inpatient stays would require larger sample sizes irrespective of the duration of observation or the critical values of the ICER.

3.4.3 Choice of medical care resource use unit costs

The results presented in Table 3.5 show that the sample sizes required for trials performed to test hypotheses about the cost-effectiveness of an intervention are influenced not only by which components of medical care resource use are measured, but also by how they are valued. The results shown in rows numbered 2 to 4 of Table 3.3 are based on the assumption that the price of the medication under study equates to a cost of £250 per patient per year.

Table 3.4
Sample Sizes for an RCT when different medical care resources are evaluated

RCT Design Attribute	ρ [a]	ICER[b]	Total Sample Size for Critical ICER of :[c,d] £5,000	£10,000	£15,000	£20,000	£25,000
1. Clinical baseline		- £3,193	58	58	58	58	58
2. Inpatient only: 1 year	0.094	£1,919	128	72	60	56	54
3. Inpatient only: 2 years	0.128	£1,815	122	70	60	56	54
4. Inpatient only: 3 years	0.143	£1,750	176	108	94	88	86
5. Outpatient only: 1 year	0.226	£3,807	820	116	80	68	62
6. Outpatient only: 2 years	0.327	£2,589	182	82	66	60	58
7. Outpatient only: 3 years	0.349	£2,445	256	122	102	94	90
8. Daypatient only: 1 year	0.093	- £6,844	956	476	290	200	150
9. Daypatient only: 2 years	0.057	£3,553	14238	786	290	168	120
10. Daypatient only: 3 years	0.067	£7,010	*	3158	574	290	200
11. Excluding inpatient costs: 1 year	0.102	- £4,074	1600	670	374	244	176
12. Excluding inpatient costs: 2 years	0.072	£5,387	*	1516	404	210	142
13. Excluding inpatient costs: 3 years	0.082	£8,708	*	16794	918	378	242

a ρ = the correlation between the difference in costs and effects of treatments A and B

b Incremental cost-effectiveness ratio calculated from the data. These values are the hypothesised ICERs for the sample size calculations.

c Probablities of Type I error = 0.05 and Type II error = 0.10 throughout. All sample size calculations take into account the correlation between cost and effectiveness differences using the formula cited in the text.

d All calculations are based on undiscounted costs and effects

* Denotes the fact that the sponsor of a technology under investigation would probably not conduct a study to test economic hypotheses if the postulated ICER was greater than the critical value.

Rows numbered 2 to 4 of Table 3.5 show the impact on sample size of doubling the assumed price of the treatment under investigation and rows numbered 5 to 7 show the impact of quadrupling the price. It can be seen that in this case study sample sizes are generally an increasing function of the price of the product under investigation. An identical pattern is observed when the unit costs for inpatient, daypatient and outpatient attendances are simultaneously increased by 50 per cent (rows numbered 8-10) or 100 per cent (rows numbered 11-13) compared with the baseline unit costs assumed for the other calculations. Large increases in sample size requirements are observed in the instances where changing the unit costs brings the hypothesised ICER into very close proximity to the highlighted critical values. For example, when the critical ICER is assumed to be £10,000, the total sample size required for a 2 year cost-effectiveness study increases from 2634 (row number 3 of Table 3.3) to 140408 when unit costs are increased by 50% (row 9 of Table 3.5). Conversely, large sample size reductions are observed when the hypothesised ICER diverges from the critical value as a result of changes in unit costs.

The sensitivity of sample size requirements to the choice of unit costs used to value medical care resource utilisation data is illustrated graphically in Figure 3.2. Figure 3.2 shows sample sizes (logarithmic scale) as a function of critical values of the ICER for the clinical baseline RCT design scenario, and for cost-effectiveness studies performed over a 1-year period using different assumptions about drug price (100% increase) and other resource use (100% increase).

Table 3.5
Sample Sizes for an RCT when different medical care resource costs are used

RCT Design Attribute	ρ [a]	ICER[b]	Total Sample Size for Critical ICER of :[c,d] £5,000	£10,000	£15,000	£20,000	£25,000
1. Clinical baseline		- £3,193	58	58	58	58	58
2. Drug price x 2: 1 year	0.105	- £2,155	2566	896	460	288	202
3. Drug price x 2: 2 years	0.080	£7,202	*	4242	628	278	174
4. Drug price x 2: 3 years	0.092	£10,458	*	*	1786	534	304
5. Drug price x 4: 1 year	0.107	- £80	5092	1302	596	350	236
6. Drug price x 4: 2 years	0.086	£8,712	*	19962	962	356	206
7. Drug price x 4: 3 years	0.101	£11,951	*	*	3942	746	374
8. Resource prices x 1.5: 1 year	0.104	- £5,308	2786	1262	722	472	336
9. Resource prices x 1.5: 2 years	0.076	£9,292	*	140408	2298	714	370
10. Resource prices x 1.5: 3 year	0.085	£14,193	*	*	99512	2262	782
11. Resource prices x 2: 1 year	0.104	- £7,423	3416	1732	1046	704	508
12. Resource prices x 2: 2 years	0.075	£12,138	*	*	15468	2154	858
13. Resource prices x 2: 3 years	0.084	£18,674	*	*	*	65676	3248

a ρ = the correlation between the difference in costs and effects of treatments A and B

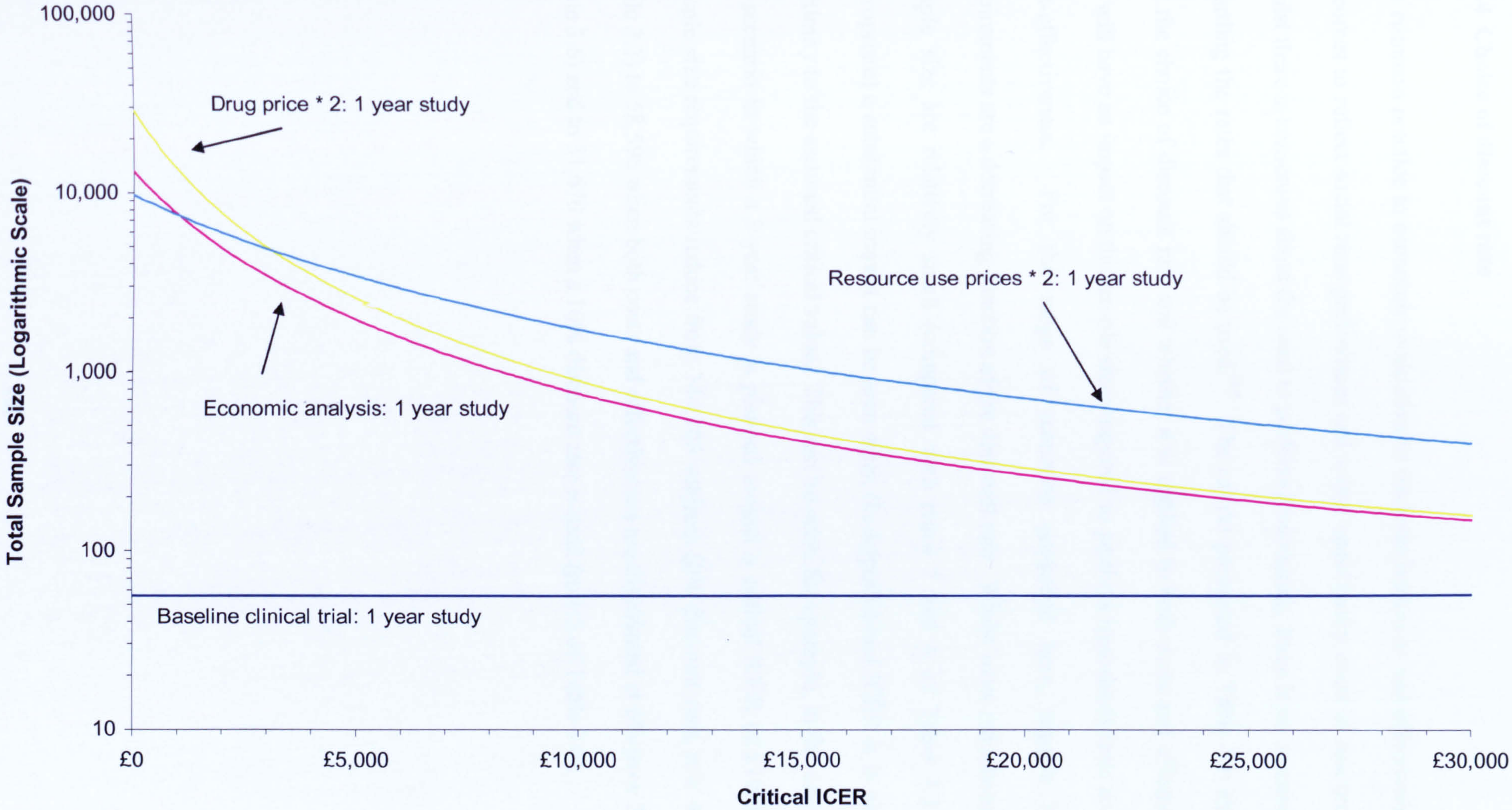
b Incremental cost-effectiveness ratio calculated from the data. These values are the hypothesised ICERs for the sample size calculations.

c Probablities of Type I error = 0.05 and Type II error = 0.10 throughout. All sample size calculations take into account the correlation between cost and effectiveness differences using the formula cited in the text.

d All calculations are based on undiscounted costs and effects

* Denotes the fact that the sponsor of a technology under investigation would probably not conduct a study to test economic hypotheses if the postulated ICER was greater than the critical value.

Figure 3.2: Sample Size Implications of Choosing Different Resource Use Prices



3.4.4 Choice of discount rate

It is common practice in economic evaluation to discount both cost and effectiveness outcomes to reflect social time preferences and social opportunity costs of resources. Whilst there is consensus about the need to perform discounting, there is no consensus regarding the rates that should be used.[201] The results presented in Table 3.6 show that the choice of discount rate and whether it is applied to both costs and effects or not will have an impact on the sample sizes required to perform hypothesis tests about cost-effectiveness. For the range of scenarios presented here, sample size requirements are a decreasing function of the discount rate. Whilst most reductions in sample size are relatively small (compared with rows 3 and 4 of Table 3.3 as appropriate) a substantial impact can be seen when the hypothesised ICER is in close proximity to the assumed critical value. This can be seen, for example, in the case of the scenario in which a 3-year study is planned around a critical ICER of £10,000. Sample size requirements reduce from 344,106 subjects (0% discount rate, row 4 of Table 3.3) to 58,596 when both costs and effectiveness are discounted at 6% (row 3 of Table 3.6) and to 31,670 when a 10% discount rate is used (row 5 of Table 3.6).

Table 3.6
Sample Sizes for an RCT when different discount rates are used

RCT Design Attribute	ρ [a]	ICER[b]	Total Sample Size for Critical ICER of :[c,d] £5,000	£10,000	£15,000	£20,000	£25,000
1. Clinical baseline		- £3,193	58	58	58	58	58
2. Discount rate: 6% 2 years	0.070	£6,201	*	2394	510	246	160
3. Discount rate: 6% 3 years	0.075	£9,287	*	58596	1156	430	260
4. Discount rate: 10% 2 years	0.066	£6,044	*	2262	504	246	160
5. Discount rate: 10% 3 years	0.068	£9,020	*	31670	1070	412	252
6. Discount costs only: 6% 2 years	0.065	£6,003	*	2052	468	232	154
7. Discount costs only: 6% 3 years	0.068	£8,860	*	21614	970	392	248

a ρ = the correlation between the difference in costs and effects of treatments A and B

b Incremental cost-effectiveness ratio calculated from the data. These values are the hypothesised ICERs for the sample size calculations.

c Probablities of Type I error = 0.05 and Type II error = 0.10 throughout. All sample size calculations take into account the correlation between cost and effectiveness differences using the formula cited in the text.

d All calculations are based on discounted costs and effects using the rates shown in column 1.

* Denotes the fact that the sponsor of a technology under investigation would probably not conduct a study to test economic hypotheses if the postulated ICER was greater than the critical value.

3.4.5 Choice of study population

The results presented so far pertain to a study population which contains patients with different levels of disease severity at diagnosis. However, it is likely that the cost-effectiveness of an intervention will vary amongst sub-groups within a population of people with a disease. Therefore a diagnostic variable was used to partition the dataset into two groups of patients. For illustrative purposes, these are referred to as sub-group A (advanced disease at diagnosis) and sub-group B (early disease at diagnosis). Table 3.7 shows the sample size requirements for performing separate trials to test the hypothesis that the net benefit of the intervention is statistically significantly positive for sub-groups A and B. In both cases, the sample sizes required are much larger than those necessary for the baseline clinical trial design involving a mix of patients (row 1). Whilst the sample sizes for sub-group A (rows 2-4 of Table 3.7) are mostly lower than the corresponding requirements for an economic study involving a mix of patient types (rows 2-4 of Table 3.3), the sample size requirements for sub-group B are predominantly larger (rows 5-7 of Table 3.7). Once again, the largest changes in sample sizes occur when the hypothesised ICER is in close proximity to the assumed critical value.

The sensitivity of sample size requirements to the choice of study population is illustrated graphically in Figure 3.3. Figure 3.3 shows sample sizes (logarithmic scale) as a function of critical values of the ICER for the clinical baseline RCT design scenario, and for cost-effectiveness studies performed over a 1-year period using the three different study populations.

Table 3.7
Sample Sizes for an RCT when different study populations are used

RCT Design Attribute	ρ [a]	ICER[b]	Total Sample Size for Critical ICER of :[c,d] £5,000	£10,000	£15,000	£20,000	£25,000
1. Clinical baseline		- £3,193	58	58	58	58	58
2. Sub – group A: 1 year	0.134	- £12,429	432	262	180	134	106
3. Sub – group A: 2 years	0.092	£4,800	875920	1384	408	216	146
4. Sub – group A: 3 years	0.060	£8,893	*	34794	1430	556	342
5. Sub – group B: 1 year	0.052	£16,390	*	*	*	10764	1994
6. Sub – group B: 2 years	0.115	£10,710	*	*	1690	444	234
7. Sub – group B: 3 years	0.208	£11,827	*	*	1974	416	222

a ρ = the correlation between the difference in costs and effects of treatments A and B

b Incremental cost-effectiveness ratio calculated from the data. These values are the hypothesised ICERs for the sample size calculations.

c Probablities of Type I error = 0.05 and Type II error = 0.10 throughout. All sample size calculations take into account the correlation between cost and effectiveness differences using the formula cited in the text.

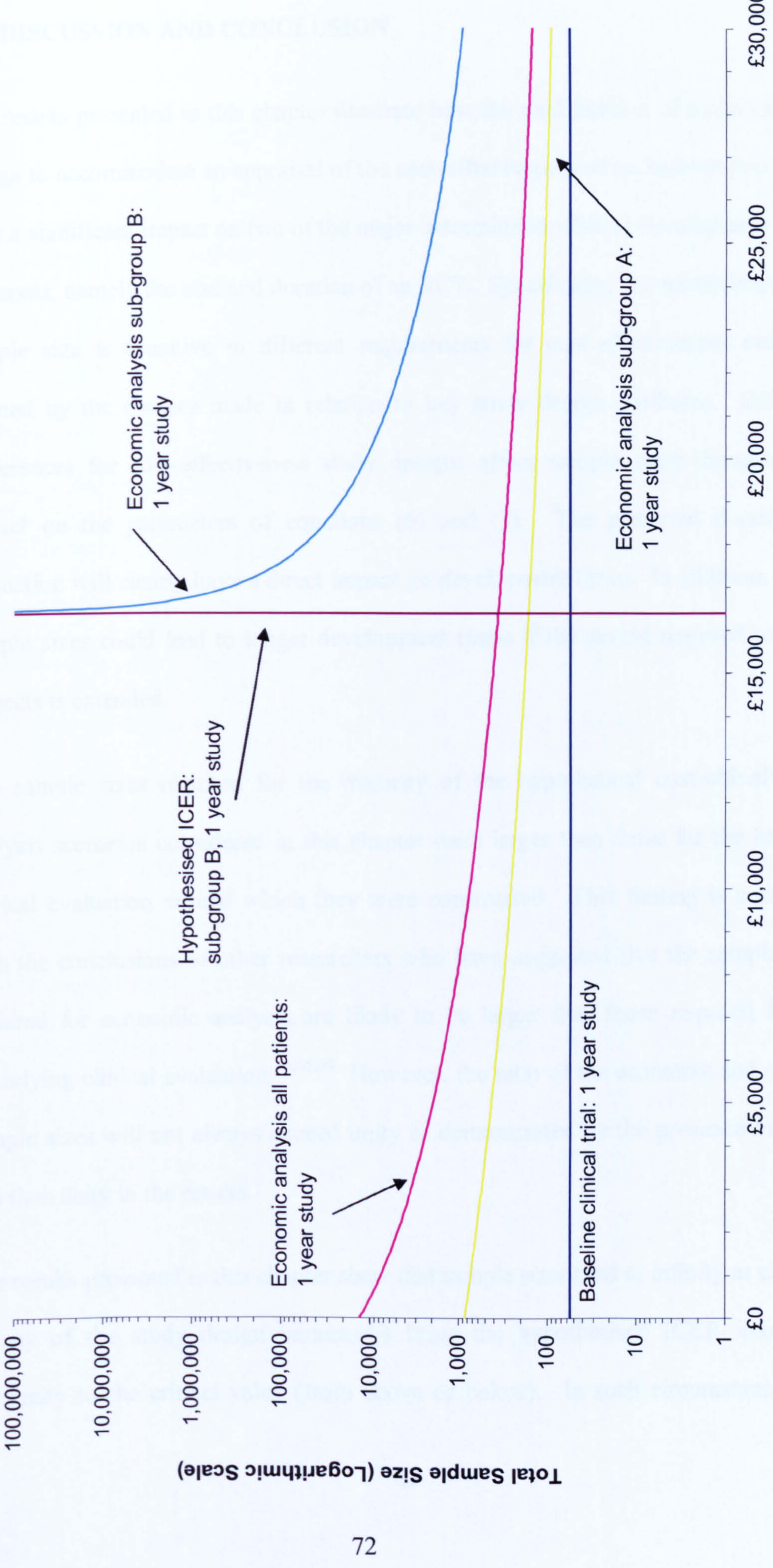
d All calculations are based on discounted costs and effects using the rates shown in column 1.

* Denotes the fact that the sponsor of a technology under investigation would probably not conduct a study to test economic hypotheses if the postulated ICER was greater than the critical value.

A Sub-group A = advanced disease at diagnosis.

B Sub-group B = early disease at diagnosis.

Figure 3.3: Sample Size Implications of Choosing Different Study Population Sub-Groups and Study Durations for Economic Evaluations



3.5 DISCUSSION AND CONCLUSION

The results presented in this chapter illustrate how the modification of a clinical trial design to accommodate an appraisal of the cost-effectiveness of an intervention could have a significant impact on two of the major determinants of drug development times and costs, namely the size and duration of an RCT. Specifically, the results show that sample size is sensitive to different requirements for cost-effectiveness evidence defined by the choices made in relation to key study design attributes. Different preferences for cost-effectiveness study designs affect sample sizes through their impact on the parameters of equations (6) and (7). The preferred duration of evaluation will clearly have a direct impact on development times. In addition, larger sample sizes could lead to longer development times if the period required to enrol subjects is extended.

The sample sizes required for the majority of the hypothetical cost-effectiveness analysis scenarios considered in this chapter were larger than those for the baseline clinical evaluation around which they were constructed. This finding is consistent with the conclusions of other researchers who have suggested that the sample sizes required for economic analysis are likely to be larger than those required for the underlying clinical evaluation.[54;65;69] However, the ratio of the economic and clinical sample sizes will not always exceed unity as demonstrated by the presence of ratios less than unity in the results.

The results presented in this chapter show that sample sizes tend to infinity as changes to any of the study design parameters bring the hypothesised ICER into close proximity to the critical value (from above or below). In such circumstances, the

required sample sizes would likely prohibit a conventional frequentist stochastic approach to cost-effectiveness analysis e.g. one can envisage extreme situations where the epidemiology of a disease is such that it could take hundreds of years to enrol a sufficient number of patients into a trial. Therefore one of the practical implications of this finding is that alternative approaches to the evaluation of the cost-effectiveness of an intervention will be necessary in such situations. Moreover, during the planning stages of an RCT, researchers would be unable to identify the potential for such a situation unless, at that time, the critical value of the ICER was known and unless pilot data were available to formulate a hypothesis about the cost-effectiveness of a new treatment. Typically, such data are unavailable to the sponsor of a new technology prior to its introduction to the market. This has led some researchers to question the practical value of being able to perform sample size calculations for an economic evaluation even though methods for doing so are available.[65]

This chapter illustrates the potential impact of a requirement to produce wholly stochastic cost-effectiveness evidence within the framework of an RCT in which the statistical hypothesis tested is that the net-benefit associated with an intervention is positive at conventional probabilities of Type I (α) and Type II (β) errors. Clearly, this is not to suggest that this is the only research approach which could be adopted for producing cost-effectiveness evidence. Rather the goal of the chapter is to shed some light on the potential implications *if* such an approach were to be implemented in practice. Indeed, alternative modelling approaches (deterministic analysis) are frequently applied,[31] and some researchers have argued that modelling may be a necessary complement to RCT based evidence (partially stochastic analysis).[39;52] From a practical standpoint, the need for such alternative research approaches seems

inevitable for the extreme situations referred to above.

Another way of avoiding the potential sample size problems of stochastic cost-effectiveness analysis would be to accept the probabilities of Type I and Type II errors resulting from the sample sizes required for the clinical evaluation, rather than insisting that the conventions used in medical research be applied to economic evaluation. Emphasis would then be placed on reporting confidence intervals or cost-effectiveness acceptability curves rather than on formal hypothesis testing. With this approach, any additional costs associated with a requirement for RCT based cost-effectiveness evidence would then be driven by the administrative consequences of collecting additional data and by any requirement to extend the duration of a trial. Alternatively, it has recently been argued that decisions about cost-effectiveness based on statistical inference are irrelevant and that a Bayesian decision theoretic approach to the stochastic evaluation of health care technologies should be used.[157] Whether this alternative approach will have more or less of an impact on pharmaceutical company development costs and timelines is unknown. This will depend upon a number of factors, such as whether and how the costs borne by the sponsor of the technology are incorporated into the analysis.

The results of this study highlight that the sample size requirements for cost-effectiveness analyses can be sensitive to the choices made about a number of RCT design attributes. The degree of uncertainty surrounding the results of any cost-effectiveness analysis will also be sensitive to these choices. One of the practical implications of this study, therefore, is that it is difficult to see how sponsors of technologies could conduct adequately designed trials without a clear prior specification of the key RCT design attributes referred to in this chapter,

namely: comparators, study population, endpoint measurements, unit costs, duration of observation, discount rates, critical value of the ICER, and acceptable probabilities of Type I and Type II errors. Failure to be explicit about each of these factors prior to the conduct of an RCT would be tantamount to bad scientific practice. Moreover, being clear on these design issues is a necessary condition for evaluating the practicality and feasibility of conducting a wholly stochastic cost-effectiveness analysis, as referred to above.

A primary concern amongst companies developing new health care technologies is how a requirement to produce evidence of cost-effectiveness based on RCTs will impact development costs and development times. Clearly, the direction and magnitude of the impact will depend upon the precise nature and extent of the cost-effectiveness evidence required by decision-makers. With respect to development times, the impact of cost-effectiveness evidence requirements will depend upon the desired duration of the evaluation of cost and effectiveness outcomes compared with the period of follow-up required purely for the purpose of conducting a clinical evaluation. For example, a requirement to follow subjects for 3 years as opposed to 1-year will extend product development times by at least 2 years. Development times can be extended indirectly if larger sample sizes affect enrolment rates. Sponsors of technologies will also be concerned about the revenue implications of extended development times. Studies of longer duration will erode the effective patent life of a product and could therefore have a negative effect on cumulative revenue. On the other hand, stronger cost-effectiveness evidence might lead to wider diffusion once the product is introduced i.e. the gains in revenue associated with a delayed introduction with stronger evidence might exceed the losses incurred as a result of that

delay.[202] The potential gains and losses associated with different strategies for producing cost-effectiveness evidence would be a worthwhile area for future research.

The impact of cost-effectiveness evidence requirements on the scale of investment required for product development will depend primarily upon the relationship between trial sample size and the cost of conducting an RCT. This relationship will be known by a company, who will be able to convert the sample size calculations into cost functions.[172] If, for example, the total cost of conducting a trial was simply a function of sample size based on a fixed cost per subject enrolled, then Figures 3.1 to 3.3 could be interpreted as cost functions with an appropriate adjustment made to the sample size scale representing the fixed cost per subject. The figures would currently equate to a cost function where the cost per subject enrolled is £1.00. A cost per subject enrolled of £1,000 would require the sample sizes in Figures 3.1 to 3.3 to be multiplied by 1,000 in order to transform the figures into cost functions. Little is published about the behaviour of RCT cost functions. This would be a necessary area for future research if the economic consequences of alternative requirements for cost-effectiveness evidence are to be assessed. Clearly, the cost implications will be sensitive to choices made about study design attributes in the same way as sample sizes are. Sponsors of technologies will therefore have to decide whether the investment in the production of wholly stochastic cost-effectiveness evidence is worth making. This will require companies to compare the costs of alternative requirements with the expected revenue associated with meeting them. An approach for doing this has previously been proposed.[203]

In conclusion, formal requirements for wholly stochastic cost-effectiveness evidence based on the standard frequentist paradigm have the potential to increase

the size, duration and number of RCTs significantly and hence the costs and timelines associated with new product development. Moreover, it is possible to envisage situations where such an approach would be impossible to adopt. Clearly, further research is required into the issue of how to appraise the economic consequences of alternative economic evaluation research strategies. Ultimately, the results of such research could be used to inform the development of economic evaluation guidelines, specifically relating to the choice of research method appropriate to different product circumstances. In situations where wholly stochastic cost-effectiveness evidence is required from RCTs, the importance of prior specification of the key economic evaluation design attributes should not be underestimated.

CHAPTER 4: THE USE OF DISCRETE CHOICE ANALYSIS IN THE DESIGN OF RANDOMISED CONTROLLED TRIALS

SUMMARY

Randomised controlled trials (RCTs) are the primary means by which pharmaceutical companies evaluate the therapeutic benefits of their products. The strength and relevance of the evidence provided from RCTs will determine whether a product can be marketed or not and the subsequent extent of its use. In order to gain access to a market, pharmaceutical companies must perform RCTs to produce safety and efficacy evidence to a level which satisfies the regulatory bodies responsible for granting product licences. However, the safety and efficacy evidence produced for that purpose may not be sufficient to ensure that a product is reimbursed and actually used in clinical practice. Health technology assessment and appraisal bodies, such as the National Institute for Clinical Excellence (NICE) and Hospital Drugs and Therapeutic Committees, critically appraise the nature and relevance of RCT evidence in order to make recommendations about the extent to which a product should be used. Individual clinicians will make treatment decisions based on their own assessment of the evidence, as well as taking into account the reviews performed by advisory bodies. Thus, those involved with product adoption decisions will have preferences for the types of evidence they want to see and, consequently, the extent to which these preferences are satisfied will influence the nature and extent of a treatment's use. It is therefore important for sponsors of drugs to consider decision-makers' preferences for RCT designs when planning their studies. The primary objective of this chapter is to illustrate how discrete choice analysis (DCA) could be used for that purpose. The

approach is illustrated using, as a case study, the design of trials to evaluate adjuvant bisphosphonates in the management of patients with primary operable breast cancer. Clinicians' preferences for evidence are determined and then used to identify a trial design likely to lead to the highest probability of prescribing the product (market share). However, evidence generation has a cost attached to it. Therefore the chapter goes on to look at how physician preferences for evidence and the resulting predicted impact on product use can be combined with trial design costs in an overall investment appraisal framework. Within such a framework, it is shown how a company producing a technology could identify the profit maximising RCT strategy. Finally, a number of issues for consideration in future research are briefly discussed, including the circumstances under which private and public sector perspectives are likely to be aligned.

4.1 INTRODUCTION

Discrete choice analysis (DCA) is the name given to a set of multivariate data analysis techniques which can be used to predict decision-makers' choices between alternative products or services.[204-206] The techniques have been widely applied to assist with product design and marketing decisions in a number of industries.[180;207;208] In the commercial context, the primary interest has been to use DCA to estimate the probability that a decision-maker will choose a given product or service from the set of available alternatives. Since the probability of choosing a given product is assumed to depend upon the utility derived from its attributes compared with that of its alternatives, it is possible to use DCA to estimate the demand for both new and existing products given different defining characteristics.

In contrast to the commercial applications of DCA, where the primary interest is in modelling product demand, health economists have recently begun to use the technique for estimating the value of treatment processes and outcomes in preference, utility or monetary terms. The literature on health applications of DCA is now extensive.[209-228;228-260] But, to date, DCA has not been applied to assist with the design of randomised controlled trials (RCTs). However, the purpose of RCTs is such that DCA is likely to be of value in the RCT planning and design context because there is a relationship between the decision to adopt a health care intervention (demand for the intervention) and the design characteristics of the RCTs used to evaluate its benefits.

Randomised controlled trials (RCTs) are the primary means by which pharmaceutical companies evaluate the therapeutic benefits of their products.[1-5] The strength and

relevance of the evidence provided from RCTs will determine whether a product can be marketed or not and the subsequent extent of its use. In order to gain access to a market, pharmaceutical companies must perform RCTs to produce safety and efficacy evidence to a level which satisfies the regulatory bodies responsible for granting product licences. However, the safety and efficacy evidence produced for that purpose may not be sufficient to ensure that a product is reimbursed and actually used in clinical practice. Health technology assessment and appraisal bodies, such as the National Institute for Clinical Excellence (NICE) and Hospital Drugs and Therapeutic Committees, critically appraise the nature and relevance of RCT evidence in order to make recommendations about the extent to which a product should be used.[191] Individual clinicians will make treatment decisions based on their own assessment of the evidence, as well as taking into account the reviews performed by advisory bodies. Thus, those involved with product adoption decisions will have preferences for the types of evidence they want to see and, consequently, the extent to which these preferences are satisfied will influence the nature and extent of a treatment's use. It is therefore important for sponsors of drugs to consider decision-makers' preferences for RCT designs when planning their studies. The primary objective of this chapter is to illustrate how discrete choice analysis (DCA) could be used for that purpose. The approach is illustrated using a discrete choice stated preference (SP) survey concerned with the design of trials to evaluate adjuvant bisphosphonates in the management of patients with primary operable breast cancer.

The remainder of the chapter is divided into seven sections. In the following section, a discrete choice modelling approach to drug prescribing behaviour is set out in general form. This is followed in section 4.3 by an overview of the key components

of a discrete choice SP survey. In section 4.4, the design of the adjuvant bisphosphonates case-study survey is presented. The results for the non-choice question components of the survey are presented in section 4.5. The results pertaining to the estimation of the parameters of a binary discrete choice model are presented in section 4.6 where consideration is given to the qualitative and quantitative effects. Section 4.7 focuses on using the discrete choice model results for the design of RCTs. Specifically, the use of the results to determine designs which optimise the probability of product adoption and to operationalise an investment appraisal approach to RCT design are illustrated. The final section includes a discussion of the results and the implications for future research.

4.2 A DISCRETE CHOICE MODEL OF DRUG DEMAND

In this section, a discrete choice model of drug demand is set out in general form. A specific binary choice formulation of this model is used in the applied example which follows later in the chapter.

4.2.1 Random utility theory of drug choice behaviour

Discrete choice models derive from random utility theory of choice behaviour.[204;206] Under this theory, the probability that a clinician will choose drug i from the set of alternative treatments available, J, is given by:

$$Pr(i \mid \mathrm{J}) = Pr(U_i > U_j) \quad \forall j \in J, j \neq i \qquad (1)$$

where U_i and U_j denote the utility which a clinician derives from using the different products, i and j. A clinician is assumed to choose the treatment option which

maximises his or her utility. Assuming that the clinician is behaving as a perfect agent, this should also be the choice which maximises the utility of the patient receiving the treatment. Since the utility of a treatment is assumed to be derived from the characteristics that define it, equation (1) can be re-written as:

$$Pr(i \mid \mathrm{J}) = Pr\left(U_i(Z_i) > U_j(Z_j)\right) \quad \forall j \in J, j \neq i \qquad (2)$$

where Z_i and Z_j denote vectors of characteristics the levels of which define the treatment alternatives, i and j. Note that the vectors of attributes can include characteristics of the clinician e.g. the preferences of primary care physicians might differ from those of hospital specialists.

If a clinician's utility function was known and if all the relevant characteristics were observed, then perfect predictions could be made about a clinician's choice of treatment. Since this is not the case in practice, a discrete choice model of behaviour can be constructed based on the following identity:

$$Pr(i \mid \mathrm{J}) = Pr\left(U_i(Z_i) > U_j(Z_j)\right) = Pr\left(V_i(X_i) + \varepsilon_i > V_j(X_j) + \varepsilon_j\right) \quad \forall j \in J, j \neq i \qquad (3)$$

where V_i and V_j denote the observable components of utility, X_i and X_j are vectors of observable treatment characteristics and ε_i and ε_j are the unobserved random components of utility for products i and j respectively. The latter takes into account the difference between the true, U, and observed, V, utility. The right hand side of equation (3) can be re-arranged to give the following general (multinomial) expression for a random utility model of drug prescribing behaviour:

$$Pr(i \mid \mathrm{J}) = Pr(\varepsilon_j - \varepsilon_i < V_i(X_i) - V_j(X_j)) \quad \forall j \in J, j \neq i \tag{4}$$

4.2.2 Discrete choice model formulations

In order to operationalise the above model, it is necessary to specify functional forms for both the observable and unobservable components of utility. For the deterministic component of utility, it is common practice in discrete choice models to specify V as a function which is linear in the vector of unknown parameters, β', such that:

$$V_i = \beta' X_i = \beta_1 X_{1i} + \ldots\ldots\ldots + \beta_l X_{li} \tag{5}$$

$$V_j = \beta' X_j = \beta_1 X_{1j} + \ldots\ldots\ldots + \beta_l X_{lj} \quad \forall j \in J, j \neq i$$

where $\beta_1 \ldots\ldots\beta_l$ are the coefficients to be estimated for each of the l attributes included in the model. In practice it has been observed that the linear additive model of equation (5) works well in most applied situations[204] and is a formulation that has been used frequently in recent health economics applications. This functional form for the observable component of utility will therefore be used in the analyses which follow.

For the unobservable component of utility, the disturbances $(\varepsilon_i, \varepsilon_j)$ are assumed to be distributed randomly (hence the name random utility model). A number of alternative distributions can be assumed which give rise to different discrete choice model formulations.[204;205] The most frequently used approaches are the logit and probit models. It has been noted that in practice there is little difference between the results derived from those two approaches.[205] Since the probit model is used in the case-study which follows, it is described in more detail here. In the case of probit

models, the unobserved components of utility are assumed to be distributed jointly normal. Using the probit discrete choice model formulation, the probability that a prescribing clinician will choose drug i from the set of alternative treatments available, J, is given by:

$$Pr(i \mid \mathrm{J}) = \int_{\varepsilon_i=-\infty}^{\infty} \int_{\varepsilon_1=-\infty}^{\varepsilon_i+V_i-V_1} \int_{\varepsilon_2=-\infty}^{\varepsilon_i+V_i-V_2} \ldots\ldots \int_{\varepsilon_j=-\infty}^{\varepsilon_i+V_i-V_j} \Phi\left(\tilde{\varepsilon}\right) de_j \ldots de_2 de_1 de_i, \qquad (6)$$

where $\Phi(.)$ denotes the standardised cumulative normal distribution, $\tilde{\varepsilon}$ is a vector composed of each disturbance ε_i for all i in J and there are j alternatives in J. The probit model is estimated using maximum likelihood techniques. This gives rise to estimates for $\beta_1 \ldots\ldots \beta_l$ and consequently, through equation (6), the probabilities of choosing alternative products can be derived.

4.3 DISCRETE CHOICE STATED PREFERENCE SURVEYS

The parameters of a discrete choice model, such as that set out in equation (6), can be estimated using data pertaining to observed choice behaviour (revealed preference data), simulated choice behaviour (stated preference data) or a combination of the two.[204;206] Regardless of the source of data, the dataset needs to contain, for each alternative in the choice set, an indicator of the choice made together with the defining characteristics of the alternatives (the Xs in the above equations).

To date, the approach typically adopted by health economists has been to use simulated choice data obtained from discrete choice stated preference surveys (often referred to as conjoint analysis).[179;250] Since the required data on actual drug choice behaviour would be difficult to obtain and, by definition, is not available for

new products in development, the approach adopted in this chapter is to use data generated from a discrete choice stated preference (SP) survey. The design stages of such surveys have been enumerated in detail elsewhere,[179;206] but generally they include the following components:

1) Determination of attributes, levels and scenarios

2) Elicitation of preferences

3) Data analysis and interpretation

These stages are discussed briefly in turn below.

4.3.1 Determination of attributes, levels and scenarios

When designing an SP survey, the attributes (characteristics) of interest need to be defined and levels (values) need to be assigned to them. A number of approaches to doing this have been identified, including the use of literature reviews, interviews and selection based on a specific research question.[179;206] The various approaches are not mutually exclusive and, in practice, a combination of them is often used. At a general level, it is postulated here that the probability of choosing a given health care intervention is a function not only of the demonstrated benefits, but also of the 'design' characteristics of the RCTs from which the evidence of product benefit is derived. Consequently, the design problem in the current context involves selecting attributes and levels from the set of RCT design characteristics enumerated in Table 4.1. In order to ensure that an SP survey is realistic, the literature suggests that attribute levels should be plausible and capable of being traded.[179]

Once the attributes and their levels have been determined, they are combined into

scenarios or profiles to present to survey participants for evaluation. A scenario is a combination of attributes and levels that characterize the choice object of interest in the study, in this case RCTs. The number of possible scenarios (the full factorial design) defined by the chosen number of attributes and levels can be very large and is given by:

$$S = \prod L_i^{A_i}$$

where S denotes all possible combinations of attribute levels and A_i denotes the number of attributes possessing the number of levels L_i. An SP survey with a very large number of scenarios would be impractical due to the cognitive burden which the presentation of a large number of scenarios would place on survey participants. Therefore a practical problem to overcome is how to reduce the number of scenarios whilst ensuring that the parameters of the model can be reliably estimated. A common approach to reducing the number of scenarios is to identify an orthogonal fraction using experimental design catalogues such as those available in computer programmes like SPEED.[261] Orthogonal arrays of scenarios are such that each attribute level appears an equal number of times and the attributes are uncorrelated.[206]

Table 4.1
Trial Design Attributes

1. Comparators
Can be chosen from one or more broad types, including:

i) Placebo
ii) Most commonly used
iii) Most effective
iv) Least cost
v) Most cost-effective.

Specification usually involves the choice of specific product formulations and modes of administration.

Most studies compare two treatments although more are possible.

2. Population
Specification usually involves choices about:

i) Age group
ii) Sex
iii) Ethnic origin
iv) Disease stage
v) Co-morbidities
vi) Previous treatments
vii) Concomitant treatments
viii) De novo or refractory patients.
ix) Sub-group comparisons

3. Setting
Specification usually involves choices about:

i) Single country, single centre
ii) Single country, multi-centre
iii) Multinational, single centre
iv) Multinational, multi-centre
v) Inpatient
vi) Outpatient
vii) Specialist centre
viii) Routine practice centre

4. Endpoints
Specification usually involves choices about:

i) Efficacy
ii) Effectiveness
iii) Side effects
iv) Adverse events
v) Quality of life
vi) Direct costs (NB includes product prices)
vii) Indirect costs
viii) Resource use
ix) Surrogate endpoints

5. Effect sizes
Specification usually involves choices about:

i) Clinical significance
ii) Statistical significance
iii) Primary endpoints
iv) Secondary endpoints

Table 4.1 (continued)

6. Duration of observation

Specification usually involves choices about:

i) Fixed period of observation
ii) Variable (e.g. in sequential designs).

Choices are linked closely to the choice of endpoints and the statistical properties of the study.

7. Acceptable error rates: α and β

Choices are linked closely to the choice of endpoints and the duration of follow-up.

Often chosen according to convention and based on the primary endpoint(s) i.e. $\alpha = 5\%$, $\beta = 10\%$

Do not have to be the same for each endpoint (and usually aren't).

Used in conjunction with the statistical properties of endpoints, the desired effect sizes and withdrawal rates to determine sample size.

8. Statistical methods

Specification usually involves choices relating to:

i) Objectives of the trial
ii) Nature of other trial parameters, most notably the disease area and endpoints (type of data)
iii) Method of randomisation.

4.3.2 Elicitation of preferences

Preferences are elicited by presenting the scenarios to respondents who are asked to rank or rate each of them, or to indicate their preference (choice) from sets of two or more profiles presented alongside each other (the discrete choice format). The preference elicitation approach preferred by health economists to date has been the discrete choice format since it reflects the random utility theory of choice behaviour (see above).[179;204;206] Further, health economists have tended to elicit preferences using binary (pairwise) choice tasks in which respondents select their preferred scenario from each of a number of pairs (the choice set). Typically, choice sets have been generated by randomly pairing (without replacement) the scenarios in the orthogonal array, although alternative approaches could be employed.

The choice sets usually incorporate some pairwise comparisons that form the basis of tests to identify inconsistent respondents. Inconsistent respondents are traditionally defined as those who do not make the choices one would expect them to make given the researcher's prior expectations about a positive or negative relationship between the attribute values and utility. Thus, to test for inconsistency defined in this way, the design needs to contain some pairwise combinations of scenarios for which the preferred scenario might be predicted *a priori*. These can fall naturally from the random generation of the choice sets or be generated manually. Using such tests, inconsistent respondents can then be identified at the analysis stage and dropped from the analysis along with non-trading subjects (see below).

As far as the author is aware, there is no formula for estimating the sample sizes required for binary choice SP surveys. Consequently, there is no firm statistical basis

for the sample sizes used in previously reported studies. However, a notable feature of discrete choice surveys is that each respondent can provide as many as *n* observations to the dataset, where *n* is the number of choice sets included in the survey. Thus, a relatively small number of respondents can provide a sufficiently large number of observations for valid statistical analyses to be performed. Finally, preference elicitation surveys have been administered to respondents in a variety of ways including the use of mail, phone, web and interactive computer elicitation techniques with adequate responses having been reported for each.[179]

4.3.3 Data analysis and interpretation

The statistical method used to analyse SP data depends upon the approach used to elicit preferences. For the discrete choice approach, which is of primary interest here, probit regression has been widely used by health economists for estimating the parameters of discrete choice models. A number of probit estimators are available in statistics programmes such as Stata Version 7.0 (Stata).[262] However, in previously published studies, researchers have tended not to specify the statistics programmes or the precise estimation commands they have used.

A standard probit estimator relies on the assumption that the explanatory variables and the error term are independently and identically distributed and that they are uncorrelated. These assumptions are likely to be inappropriate in the case of discrete choice data obtained from SP surveys since multiple observations are obtained from each respondent. Stata provides two alternative probit estimation commands which are appropriate for such repeated measurement panel data: probit (cluster) and xtprobit (pa, robust).[262] Both estimators, which are essentially equivalent, take into

account the potential for a respondent's responses to be correlated. Both approaches also generate robust standard errors.

An estimation issue which arises in the literature is whether or not discrete choice regression models should be specified with a constant term. Examples of both approaches can be found. The answer to this issue appears to lie in the way in which the choice exercise is framed. For example, in a study which looked at preferences for miscarriage management, the scenarios presented to respondents were labeled "surgical treatment" and "medical treatment."[245] Since these labels convey information which might be used by respondents to decide which option was preferred, the authors estimated a model with a constant term. The authors interpreted the negative constant as indicating a general preference for surgical over medical management when all the attributes for the two interventions are the same.

In contrast, models have been estimated without constants where the labeling of the choices conveys no properties of the alternatives. For example, in a study looking at preferences for in vitro fertilization services, the choice alternatives were labeled "clinic A" and "clinic B" and the authors estimated a model without a constant term.[263] Thus, contemporary practice is to omit constants when the choice task involves generically labeled alternatives and vice-versa.

A linear additive form of the utility function has typically been assumed by health economists on the grounds that research has shown that alternative models seldom result in a better fit than the linear additive model.[264] It has recently been pointed out that a simple regression error specification test (RESET) could be applied to determine whether there are problems associated with the linear functional form of

discrete choice models.[265] However, in health economics applications, only one study could be found that reports a test for model mis-specification.[238]

In the health economics literature, it has become common practice to estimate models based only on a subset of respondents who are deemed to be consistent traders. Inconsistent respondents (as defined above) and non-trading respondents are typically dropped from the analysis and the results obtained from the full sample are not usually reported. A non-trading respondent is defined as one who always selects a choice scenario with a higher level of a particular (dominant) attribute irrespective of the levels of the remaining attributes. Such respondents are identified at the analysis stage by looking for choice patterns consistent with this behaviour. However, it has been noted that whether or not analyses should be performed on the full sample or only on consistent traders depends upon the objectives of the study.[210] In this study, analyses are reported for both the full sample as well as subsets of consistent traders.

Finally, health economics researchers have primarily been interested in using the regression coefficients for deriving utility scores and, where a cost attribute is included, estimates of willingness-to-pay (WTP).[179] This has enabled, for example, alternative service configurations to be ranked in terms of their utility scores. To date, health economists applying discrete choice stated preference surveys have not derived predicted choice probabilities from their models, although these are the primary interest here.

4.4 CASE STUDY OF ADJUVANT BISPHOSPHONATES

4.4.1 Adjuvant bisphosphonates in the management of breast cancer

Breast cancer is the most common form of female cancer in England and Wales where, in 1998, there were 34,822 newly diagnosed cases representing an incidence rate of 130.83 per 100,000 females. The incidence of breast cancer increases sharply with age and, overall, has been rising since the early 1970s. During the same period, mortality from breast cancer has fallen. Currently, the survival rate at 5 years post-diagnosis is 75.9%. In 2000, there were 11,340 deaths from breast cancer in England and Wales.[266]

National guidance exists for the management of patients with breast cancer.[266] Management is centred on multidisciplinary teams composed of breast surgeons, oncologists (clinical and medical), radiologists, pathologists and breast care nurses. The precise nature of initial treatment depends upon the clinical staging of the disease at diagnosis, but typically involves a combination of surgery, chemotherapy, radiotherapy and hormone replacement therapy. After completion of initial treatment, patients are monitored on an ongoing basis to ensure early detection of disease recurrence (relapse). In patients who relapse, most have metastatic (distant) disease which often affects both organs (visceral metastases) and bone (osseous metastases). The prognosis for patients with metastatic disease is poor, with the aim of treatment being palliative rather than curative.

The case study presented here is concerned with the preventive use of a class of drugs known as bisphosphonates which inhibit bone resorption (destruction). Clinical research has shown that bisphosphonates reduce the incidence of hypercalcaemia

and pathological bone fractures in patients with established bone metastases from breast cancer. Moreover, bisphosphonates have been shown to reduce the risk of bone metastases in patients with relapsed breast cancer without obvious bone involvement. In view of these proven benefits, the National Institute for Clinical Excellence in England and Wales (NICE) has recently recommended that bisphosphonates be used in the management of patients with bone metastases.[266] It has been estimated that currently about one third of patients with bone metastases receive bisphosphonate treatment at an approximate annual cost in England and Wales of £3.9 million. The annual cost could rise to as much as £25.6 million per annum if there is adherence to the recent guidance. The annual cost per patient for one of the more researched oral bisphosphonates (sodium clodronate) is about £2,200 and, once initiated, is recommended to be continued as long as skeletal disease remains an important problem.[267]

Whilst bisphosphonates have been recommended by NICE as a *treatment* for patients with bone metastases[266], the benefits of adjuvant bisphosphonates as a therapeutic strategy for the prevention of metastatic bone disease in patients with primary operable breast cancer has yet to be definitively established. A trial performed by Diel et al (1998) showed that, after 2 years treatment and 3 years of follow-up, the incidence of both osseous and visceral metastases was significantly lower for patients treated with the oral bisphosphonate clodronate compared with the control group.[268] Moreover, a statistically significant reduction in all cause mortality was observed. More recently, a larger and more representative prevention trial has demonstrated similar benefits.[269] Specifically, patients treated with clodronate experienced a statistically significantly lower rate of bone metastases compared with the placebo

controls during a 2 year treatment period. This trend was observed at the end of a 5.5 years follow-up period although the difference was not statistically significant. A significant reduction in all cause mortality was observed at the end of the long-term follow-up period. Whilst the evidence is suggestive of benefits associated with early bisphosphonate use, the indication remains under investigation and a further large trial of adjuvant clodronate is currently being conducted.[270]

Given that the early (preventive) use of bisphosphonates is a new indication, it was felt that it would make a practical case study for assessing the potential use of discrete choice analysis in the design of RCTs. This is primarily because it permits the use of a binary choice model formulation (see below). Consequently, a stated preference experiment has been designed to generate choice data taking the potential preventive use of adjuvant bisphosphonate therapy in primary operable breast cancer as an applied case study. However, it is important to note that the analyses which follow are exploratory and illustrative i.e. they are not intended as a definitive application of the method in this disease area.

4.4.2 Binary choice model formulation

In the case study which follows, a clinician is assumed to be faced with a binary choice situation in which he or she has to decide between two alternative bisphosphonate prevention regimens, i or j. Such binary choice behaviour is a special case of the multinomial choice situation described above since decision makers are assumed to be faced with exactly two alternative courses of action: $J = \{i, j\}$. Thus, for the binary choice probit model, equation (6) becomes:

$$Pr(i \mid \mathrm{J}) = \Phi(V_i - V_j) = \Phi(\beta'(X_i - X_j)) \qquad (7)$$

which is estimated using maximum likelihood techniques.

Thus, in this case study we are interested in predicting the probability of product adoption given different product benefit and trial design characteristics, X_i , X_j associated with the use of adjuvant bisphosphonates. The approach uses stated preference data generated from a discrete choice experiment the key design components of which are described below.

4.4.3 Determination of attributes and attribute levels

The design problem involves selecting, from the generic RCT design characteristics previously enumerated in Table 4.1, attributes and levels of specific relevance to the bisphosphonates case-study. These were determined by reviewing adjuvant bisphosphonate RCT publications[268-270] and discussing a preliminary (pilot) survey design with physicians with specialist knowledge of breast cancer management. The specific attributes and levels chosen for the analysis and how they relate to the generic characteristics in Table 4.1 are discussed in turn below and are summarised in Table 4.2.

Table 4.2
Attributes and Levels for the Stated Preference Survey

ATTRIBUTES X_k	LEVELS
Endpoint	
The primary measure of effectiveness used in the trial	Patients without metastatic bone disease[1] Patients alive without disease recurrence[2]
Effectiveness	
Difference in % of patients achieving primary endpoint at the end of the trial: (bisphosphonate *minus* current practice)	1% 10% 25% 40%
Uncertainty	
Width of 95% confidence interval for the effectiveness outcome	Level 1 : ± 0.01 x % Effectiveness Level 2 : ± 0.25 x % Effectiveness Level 3 : ± 0.75 x % Effectiveness Level 4 : ± 0.99 x % Effectiveness
Duration	
The duration of observation of patients enrolled in the trial	2 years 4 years 8 years 10 years
Population	
Disease stage at diagnosis for patients enrolled in the trial	Stage III only[1] Stages I, II and III[2]
Cost	
Additional cost of using adjuvant bisphosphonate prevention (compared with current practice) per 100 patients treated	£0 £450,000 £900,000 £1,800,000

Notes.

1. Binary variable coded 0 for analysis.
2. Binary variable coded 1 for analysis.

Endpoint. A large number of outcome measurements (endpoints) are usually made in RCTs. However, it is usual practice to select one outcome measure (the primary endpoint) which is used as the primary basis for discriminating between treatments under investigation and for determining the sample sizes required for the study. In order to explore the impact of the choice of primary endpoint on the decision to use adjuvant bisphosphonates, a categorical attribute with two levels was used. The first level, 'patients without metastatic bone disease', was chosen to reflect a primary hypothesis relating to adjuvant bisphosphonates, namely that the incidence of bone metastases is reduced as a result of their use.[268;269] The second level, 'patients alive without disease recurrence' was chosen to reflect the fact that disease free survival is arguably a more relevant primary endpoint, as reflected in the protocol of a recently designed and ongoing trial.[270] The first level was coded '0' for analysis and the second level was coded '1' for analysis.

Effectiveness. A challenge in designing this survey was to choose levels for the effectiveness attribute that would be plausible when combined with the levels of the endpoint, study population and duration attributes. It was also necessary to ensure that trading would take place (by not choosing attribute levels too close together) and that predictions of product adoption could encompass possible improved effectiveness of future treatments (by not restricting the levels to previously observed ranges). The effectiveness attribute was included as a continuous variable representing the absolute difference in the percentage of patients achieving the primary endpoint at the end of the trial (% effectiveness for bisphosphonate minus % effectiveness for current treatment practice). Four positive levels were chosen (1%, 10%, 25% and 40%) which means that only statistically significant improvements in effectiveness in

favour of bisphosphonate prevention are considered in the analysis.

Allowing for the considerations mentioned above, the choice of attribute levels was informed by interpolating, for different annual time points, effectiveness outcomes from results reported for two recent trials[268;269] (see Appendix 4.1). From this interpolation, the smallest statistically significant difference observed was 2% (95% confidence interval: 0.33% to 3.67%) and the largest was 18% (95% confidence interval: 11.75% to 24.25%). The highest upper limit of the 95% confidence interval was 27.05% and the smallest was 0.23%. In order to facilitate respondents' interpretation, the effectiveness attribute values were also presented as 'number needed to treat' (NNT).

Uncertainty. A continuous variable attribute was included to assess the impact, on the adoption decision, of the degree of precision surrounding the point estimate of effectiveness for the primary endpoint. This was achieved by presenting 95% confidence intervals for the effectiveness outcomes which were calculated using the formula:

$$95\%\ CI = \pm P\left(Effectiveness\ (\%)\right)$$

where P , which denotes 'proportion', took on four values: 0.01, 0.25, 0.75 and 0.99. These levels of precision were the values used for the uncertainty variable at the analysis stage. The upper value was chosen to ensure that a high degree of uncertainty could be accommodated in the design without violating the assumption about the statistical significance of the results (see below). In otherwords, the constraint that P could not exceed unity ensured that the 95% confidence intervals did not straddle zero. The lower limit was chosen to accommodate a very low

degree of uncertainty. The selection of the intermediate values was arbitrary, being equidistant from the upper and lower values. In addition to presenting respondents with the 95% confidence intervals expressed as percentages, they were also presented in terms of NNT for the reasons stated for the effectiveness attribute above.

Duration. In order to assess the impact of duration of subject follow-up on the decision to use adjuvant bisphosphonates, duration of observation was included in the design as a continuous variable attribute with four levels: 2, 4, 8 and 10 years. These values were chosen to ensure that the range encompassed the periods of observation in two reported trials. Diel et al reported a median period of follow-up of 3 years, although some subjects were observed for as long as 7 years.[268] In the Powles et al trial, the median period of follow-up was 5 years with a maximum of 9.5 years.[269] The lowest level was chosen because it represents the duration of bisphosphonate prevention medication given in both trials.

Population. Patients with primary operable breast cancer can be classified into three stages of disease at diagnosis (Stages I, II and III) which reflect how advanced the disease is at presentation. Recent studies permitted the enrolment of patients from each of these three stages, although one study enrolled only subjects who were deemed to be at high risk of developing bone metastases.[268] In order to explore the impact of choice of study population on the decision to use adjuvant bisphosphonates, a categorical attribute with two levels was included in the survey design. The first level represented contemporary trial design practice of enrolling any patient with primary operable breast cancer i.e. with Stages I, II or III disease at diagnosis. The second level was chosen to depict an arguably less representative trial in which only patients with Stage III disease at diagnosis were enrolled. Since this

population has more advanced disease, such a trial would depict a desire on the part of a study sponsor to demonstrate a therapeutic benefit in a shorter period of time. *A priori*, one would expect respondents to prefer a trial which is more representative of the actual population being treated hence the first level was coded '1' for analysis and the second level was coded '0'.

Cost. In order to assess the impact of the cost of using bisphosphonates on the decision to use them, cost was included in the design as a continuous variable attribute with four levels: £0, £450,000, £900,000 and £1,800,000. For consistency with the measurement of effectiveness, the levels were defined as the additional cost of using adjuvant bisphosphonate prevention (compared with current practice) per 100 patients treated. Moreover, in the introduction to the discrete choice task, it was pointed out that the cost related to the period of the trial and that the value could reflect different product formulations and durations of medication. The level £450,000 reflects the approximate UK price for the oral clodronate dosing regimen used in the Powles et al trial.[267;269] The other levels were chosen to provide a wide range of cost possibilities which could reflect, for example, different pricing policies, dosing regimens or duration of bisphosphonate prevention medication.

Other RCT design attribute considerations. Not all the RCT design characteristics presented in Table 4.1 appeared explicitly as attributes in the case-study survey design although all but study setting were covered in the survey questionnaire in some way. Those that were not included as attributes are considered briefly in turn below.

Comparators. The choice of comparator is an important aspect of RCT design. Since adjuvant bisphosphonate prevention is not currently standard practice, the issue of

comparing explicitly against an alternative prevention regimen does not arise. The choice of comparator was not therefore included as an attribute in the discrete choice survey. Instead, in the introduction to the discrete choice tasks, respondents were asked to assume that the evidence presented came from trials where standard practice was permitted in both arms of the trial, including the use of bisphosphonates, as appropriate, in the event of relapse (see questionnaire in Appendix 4.2). These assumptions reflect the practice actually adopted in recent trials. Thus, the comparators were assumed to be standard practice plus placebo versus standard practice plus bisphosphonate prevention.

Statistical properties. The survey did not include any attributes pertaining to the statistical properties of the hypothetical RCT designs, such as sample sizes, probabilities of type I and type II errors or the statistical methods used to analyse the data. However, in the introduction to the discrete choice tasks, respondents were asked to assume that the results presented to them were statistically significant at the conventional 5% level and, more generally, that the trials were well conducted (see survey questionnaire in Appendix 4.2).

Setting. No reference was made to the setting of the study, such as whether the trial was conducted in a number of centres or in a number of countries. Since the respondents are likely to be familiar with RCTs conducted in this context, it is reasonable to suppose that they would expect such studies to be multinational, multicentre trials.

Thus, the final design was based on the six attributes as decribed above. These were used to produce hypothetical RCT design scenarios using the method described

below.

4.4.4 Generation of the discrete choice RCT design scenarios

The number of RCT design scenarios which can be defined given the attributes and levels shown in Table 4.2 is $4^4 \times 2^2 = 1024$ (the full factorial). In order to construct a cognitively manageable number of binary choice questions, an orthogonal fraction was obtained using SPEED experimental design software.[261] This resulted in a set of 16 RCT profiles as shown in Table 4.3. A key property of this fraction of profiles is that the attributes are not correlated, and that the levels appear the same number of times.

In order to generate the binary choice questions to present to respondents, a method described by Louviere, Hensher and Swait (2000) was used.[206] This involves pairing each of the 16 RCT profiles shown in Table 4.3 with a different RCT profile randomly selected from a duplicate set. This process resulted in 16 choice sets. The differences between the attribute levels for each choice set are shown in Table 4.4. In order to minimise the problem of multicollinearity, the differences in attribute levels must not be significantly correlated. The absence of statistically significant correlations at conventional levels confirms that the resulting experimental design is reasonably orthogonal.

Table 4.3
Orthogonal 16 Profile Fraction of the Full Factorial Design

	ENDPOINT	EFFECTIVENESS	UNCERTAINTY	DURATION	POPULATION	COST
Profile from fractional Design	**Primary endpoint:**	**Difference in % achieving primary endpoint at the end of the trial**	**Uncertainty (95% CI = ± value in cell x X_2)**	**Duration of observation (years)**	**Disease stage at diagnosis**	**Additional cost of preventive strategy (£ per 100 patients)**
	X_1	X_2	X_3	X_4	X_5	X_6
1	DFS	10	0.75	10	Stage III	1,800,000
2	MBD	10	0.01	8	Stage III	900,000
3	DFS	10	0.25	4	Stage I,II & III	450,000
4	MBD	10	0.99	2	Stage I,II & III	0
5	DFS	25	0.75	8	Stage I,II & III	0
6	MBD	25	0.01	10	Stage I,II & III	450,000
7	DFS	25	0.25	2	Stage III	900,000
8	MBD	25	0.99	4	Stage III	1,800,000
9	MBD	40	0.75	4	Stage I,II & III	900,000
10	DFS	40	0.01	2	Stage I,II & III	1,800,000
11	MBD	40	0.25	10	Stage III	0
12	DFS	40	0.99	8	Stage III	450,000
13	MBD	1	0.75	2	Stage III	450,000
14	DFS	1	0.01	4	Stage III	0
15	MBD	1	0.25	8	Stage I,II & III	1,800,000
16	DFS	1	0.99	10	Stage I,II & III	900,000

DFS = patients alive without disease recurrence.
MBD = patients without metastatic bone disease.

Table 4.4
16 Choice Sets for the Stated Preference Survey

Choice set	Endpoint X_1	Effectiveness X_2	Uncertainty X_3	Duration X_4	Population X_5	Cost X_6
1. Profile 1 versus Profile 15	1	9	0.50	2	-1	0
2. Profile 2 versus Profile 5	-1	-15	-0.74	0	-1	900000
3. Profile 3 versus Profile 16	0	9	-0.74	-6	0	-450000
4. Profile 4 versus Profile 14	-1	9	0.98	-2	1	0
5. Profile 5 versus Profile 7	0	0	0.50	6	1	-900000
6. Profile 6 versus Profile 9	0	-15	-0.74	6	0	-450000
7. Profile 7 versus Profile 1	0	15	-0.50	-8	0	-900000
8. Profile 8 versus Profile 11*	0	-15	0.74	-6	0	1800000
9. Profile 9* versus Profile 8	0	15	-0.24	0	1	-900000
10. Profile 10 versus Profile 12	0	0	-0.98	-6	1	1350000
11. Profile 11 versus Profile 3	-1	30	0.00	6	-1	-450000
12. Profile 12 versus Profile 10	0	0	0.98	6	-1	-1350000
13. Profile 13 versus Profile 6*	0	-24	0.74	-8	-1	0
14. Profile 14* versus Profile 13	1	0	-0.74	2	0	-450000
15. Profile 15 versus Profile 4	0	-9	-0.74	6	0	1800000
16. Profile 16 versus Profile 2	1	-9	0.98	2	1	0

Notes.

Left option in choice set possibly dominant if signs in the cells are: $X_1 = +$; $X_2 = +$; $X_3 = -$; $X_4 = +$; $X_5 = +$; $X_6 = -$.

Right option in choice set possibly dominant if signs in the cells are: $X_1 = -$; $X_2 = -$; $X_3 = +$; $X_4 = -$; $X_5 = -$; $X_6 = +$.

* Denotes the choices a respondent would make if behaving in line with prior expectations. The choice sets in which they appear, together with an expectation that a respondent should choose the same option from the identical choice sets 10 and 12, are used in the construction of tests for inconsistent respondents as described in the text.

Figure 4.1

Example Binary Choice Question

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients alive without disease recurrence	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	10% [NNT = 10]	1% [NNT = 100]
95% confidence interval on the primary endpoint	2.50% to 17.50% [NNT = 5.71 to 40.00]	0.75% to 1.25% [NNT = 80.00 to 133.33]
Duration of observation	10 years	8 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 1,800,000	£ 1,800,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B
	✓	

In the study questionnaire presented to respondents, the choice sets were formed into 16 binary choice questions. Respondents were asked to consider each choice and then indicate which alternative they would prefer based on the profile descriptions presented. Figure 4.1 provides an example of one of the choice questions presented to physicians in the stated preference survey. The study questionnaire is presented in full in Appendix 4.2.

4.4.5 Formulation of non-choice questions

A number of additional (non-choice) questions were included in the survey instrument. These covered a number of factors including:

i) Characteristics of the respondent: specialty, grade and budget responsibilities

ii) Patient caseload by stage of disease at diagnosis

iii) Respondent views on the relative importance of different decision-makers in the product adoption decision (to assess the extent to which the sample covered the important decision-makers)

iv) The importance of different trial design characteristics (to assess whether or not important attributes were omitted from the design)

v) The importance of different endpoints (to assess whether or not important endpoints were omitted from the design)

vi) The difficulty of completing the questionnaire and the time taken by both the respondent and interviewer (to assess the practicality of the

survey).

In addition to the above, respondents were invited to make comments on any aspect of the survey. The format of the non-choice questions can be seen in the full survey instrument which is presented as Appendix 4.2.

4.4.6 Sample selection and survey administration

A priori, a number of decision-makers and other influences can be hypothesized to affect the product adoption decision. One issue is whether to sample individuals or a collective decision-making unit. In this study, it was decided to focus on a sample of senior physicians selected primarily from the specialties which, *a priori*, are the most actively involved in the management of patients with this condition and, consequently, the choice of adopting the new treatment regimen or not.

The population from which the sample was selected was identified from a proprietary database containing details of UK physicians including their specialty and contact details[271] A search of the database was performed in order to identify clinicians involved in the management of breast cancer. The results of the search are summarised in Table 4.5.

Table 4.5
Stated Preference Survey Sample

Specialty[1]	Population[2]	Invited to participate	Agreed to participate[3]	Completed the questionnaire[3]
Medical oncologist	32	21	18 (85.71%)	14 (66.67%)
Clinical oncologist	284	58	18 (31.03%)	16 (27.59%)
Surgical oncologist	28	22	19 (86.36%)	17 (77.27%)
Radiologist	33	9	5 (55.56%)	2 (22.22%)
Other	17	15	10 (66.67%)	5 (33.33%)
Total	394	125	70 (56.00%)	54 (43.20%)

Notes.

1. Specialists were senior registrar grade or higher and actively involved in breast cancer management.
2. Identified from The Medical Directory, FT Business Ltd, 1999.[271]
3. Percentages are the response rates relative to the number invited to participate.

In this survey, the questionnaire was administered using a telephone-mail-telephone technique which involved the following three steps:

i) Calling potential respondents to enlist involvement and, if willing to participate, to arrange a telephone follow-up interview

ii) Mailing the questionnaire to participants to review the materials and complete the responses, and

iii) Follow-up telephone interviews to record responses on paper.

A professional market research agency was commissioned to implement the survey, although they were not involved with the design of the questionnaire, the processing of the data, the statistical analysis or the interpretation of the results. All completed questionnaires were mailed to the author who processed and analysed the data.

Prior to the implementation of the full survey, simulations based on pilot data (4 completed questionnaires) were used to assess the results with different sample sizes. In this way, a sample in excess of 25 respondents was deemed necessary although a target of 100 was set within the data collection budget available. A final sample of 54 was achieved (see Table 4.5 and the results section below).

4.4.7 Model specification and estimation

In order to estimate a binary choice probit model of the demand for bisphosphonates as postulated in equation (7) above, the following linear additive utility function was assumed:

$$\Delta V = \beta_1\ \Delta ENDPOINT + \beta_2\ \Delta EFFECTIVENESS + \beta_3\ \Delta UNCERTAINTY + \beta_4\ \Delta DURATION + \beta_5\ \Delta POPULATION + \beta_6\ \Delta COST + \varepsilon \quad (8)$$

where ΔV is the difference in utility between the two bisphosphonate prevention regimens, $\Delta ENDPOINT$ is the difference in the primary endpoint, $\Delta EFFECTIVENESS$ is the difference in the effect size demonstrated, $\Delta UNCERTAINTY$ is the difference in the degree of uncertainty surrounding the demonstrated effectiveness, $\Delta DURATION$ is the difference in the duration of observation, $\Delta POPULATION$ is the difference in study population and $\Delta COST$ is the difference between the incremental cost associated with bisphosphonate use. $\beta_1 - \beta_6$ are the parameters to be estimated, and ε is the unobservable error term for the model which reflects the unobservable factors in the utility function. Given that the choice alternatives presented to respondents are couched in 'generic' terms (i.e. Prevention A and Prevention B), models were estimated without a constant.

The explanatory variables are measured as the differences between the levels of the attributes appearing in the 16 choice questions (prevention A minus prevention B) as shown in Table 4.4. ΔV is measured as a binary variable which takes on the value '1' if prevention A is chosen (the left hand side of the choice sets) and '0' if prevention B is chosen (the right hand side choice).

Models were estimated using the probit (cluster) command in Stata version 7 (Stata).[262] A regression error specification test (RESET) was applied to each model in order to determine whether there were problems associated with the functional form of the model.[265] Any model failing the RESET test at conventional levels of significance ($p < 0.05$) would be regarded as being mis-specified.

Models were estimated for the full sample of respondents and for two sub-groups of 'consistent traders' identified using the definitions of inconsistent and non-trading respondents given below.

Consistent traders sub-group A. In this survey, a test for consistency fell naturally from the random pairing of the choice scenarios since two choice sets contained the same profiles (see choice sets 10 and 12 in Table 4.4 and Appendix 4.2). One would expect a respondent who is consistent with their answers to select the same scenario for both of these choices. An advantage of this definition over the conventional approach described below is that it is not necessary to have prior expectations about the qualitative effects to perform this test. In this study it is therefore regarded as the primary test of consistency. Respondents who failed to choose the same scenario for choice sets 10 and 12 were dropped for this sub-group analysis together with non-trading respondents. Non-trading respondents were identified at the analysis stage by examining those individuals who exhibited any one of the choice patterns shown in Table 4.6.

Consistent traders sub-group B. A sub-group analysis was also performed based on a conventional test of consistency. Table 4.4 shows four choice sets for which the preferred scenarios might be predicted given the expected signs of the coefficients.

Respondents who failed to make choices in line with those that might be expected for choice sets 8, 9, 13 or 14 were dropped for this sub-group analysis together with non-trading respondents who were identified in the same way as for sub-group A above.

No other sub-group analyses were performed (e.g. separate analyses by specialty) due to the relatively small sample sizes to which such analyses would give rise.

Table 4.6

Choice patterns used to define non-trading respondents[1]

Choice set	Endpoint	Effectiveness	Uncertainty	Duration	Population	Cost
1	1	1	1	1	0	
2	0	0	0		0	1
3		1	0	0		0
4	0	1	1	0	1	
5			1	1	1	0
6		0	0	1		0
7		1	0	0		0
8		0	1	0		1
9		1	0		1	0
10			0	0	1	1
11	0	1		1	0	0
12			1	1	0	0
13		0	1	0	0	
14	1		0	1		0
15		0	0	1		1
16	1	0	1	1	1	

1. Non-trading respondents were deemed to be those who exhibited the choice patterns specified in the columns of Table 4.6. A '1' in a column indicates that bisphosphonate prevention option A was chosen and '0' indicates the choice of option B.

4.5 RESULTS: NON-CHOICE QUESTIONS

In this section, the results of the non-choice question components of the stated preference survey are presented. Many of the results tables referred to in this section can be found in Appendix 4.3. Such tables are denoted Table A4.3.1, A4.3.2 etc.

4.5.1 Study population and sample

The survey response rate, by specialty, is shown in Table 4.5. A total of 394 specialists were identified of which 125 (31.73%) were invited to participate in the survey. Of those invited to participate, 55 (44.00%) refused and 70 (56%) accepted. Of those agreeing to participate, questionnaires were obtained from 54 providing an overall response rate of 43.20%. Therefore a sample of 54 questionnaires was obtained within the budget constraint and the completion rate for those who responded was 100%.

4.5.2 Respondent characteristics

The composition of the 54 respondents in terms of their specialty and title are shown in Table A4.3.1. The sample included 17 surgical oncologists (31.48%), 14 medical oncologists (25.93%) and 16 clinical oncologists (29.63%). Forty respondents (74.07%) were senior registrar grade or higher. At the time the survey was conducted, only one of the respondents was not involved in the day-to-day management of patients with breast cancer. The annual number of new cases of breast cancer seen by the respondents is summarised, by specialty, in Table A4.3.2. For the sample as a whole, the average number of new cases seen each year is 176.37 (SD = 142.42). The estimated distribution of new cases by stage of disease at diagnosis is shown in Table

A4.3.3. Only 14 respondents (25.93%) indicated having any involvement with the management of budgets related to the treatment of patients with breast cancer (Table A4.3.4). The nature of that responsibility, exactly as articulated by the respondents, can be found in Table A4.3.15.

4.5.3 Survey completion

Apart from the optional open-ended questions, there were no missing responses. It can be seen from Table A4.3.5 that only 2 respondents found the questionnaire "very difficult" to complete. 22 respondents (40.74%) found the questionnaire "moderately difficult" to complete, 18 (33.33%) found it "slightly difficult" to complete and 12 (22.22%) found it "not difficult" to complete. Respondents spent an average of 26.76 minutes (SD = 13.11) reviewing the materials and preparing their responses for the telephone interview (Table A4.3.6). The telephone interviews lasted an average of 11.41 minutes (SD = 4.56). Therefore in total respondents spent an average of 38.17 minutes (SD = 13.32) participating in this survey.

Of the 54 respondents, 48 (88.89%) indicated a willingness to participate in future research. This required their personal details to be disclosed (Table A4.3.7). The same number indicated that they would like to see the results of the study (Table A4.3.8). Finally, 30 respondents (55.56%) provided comments on the questionnaire (Table A4.3.9). The comments, exactly as articulated by the respondents, can be found in Table A4.3.16.

4.5.4 Influences on the decision to use adjuvant bisphosphonates

Respondents were asked to rate the importance of a predetermined list of specialties

on a 3 point ordinal scale:

i) High degree of influence on the decision to adopt bisphosphonates (coded 1 for analysis)

ii) Some influence on the decision to adopt bisphosphonates (coded 2 for analysis)

iii) No influence on the decision to adopt bisphosphonates (coded 3 for analysis).

The results of the analysis of the responses to this question using the above coding are shown in Table A4.3.10. Medical oncologists (mean rating 1.15, SD = 0.41), radiotherapists (1.35, SD = 0.55) and surgical oncologists (1.80, SD = 0.59) were viewed as the specialties with the highest degree of influence on the decision to use bisphosphonates.

36 respondents (66.67%) indicated that important influences on the decision to use bisphosphonates were missing from the list of specialties provided (Table A4.3.11). These are shown, exactly as articulated by the respondents, in Table A4.3.17. The missing influences cited were other specialties (16 citations, 33.33% of all citations), nurses (14, 29.17%), patients / relatives / patient support groups (11, 22.92%), managers / policy makers (6, 12.50%) and the media (1, 2.08%).

4.5.5 Importance of adjuvant bisphosphonate trial design characteristics

Respondents were asked to rate the importance of a predetermined list of trial design characteristics on a 4 point ordinal scale:

i) Very important characteristic (coded 1 for analysis)

ii) Quite important characteristic (coded 2 for analysis)

iii) Characteristic of little importance (coded 3 for analysis)

iv) Characteristic not important (coded 4 for analysis).

The results of responses to this question using the above coding are shown in Table A4.3.12. The results confirm the importance of the six trial design characteristics included in the discrete choice exercise. Four of these characteristics (primary endpoint, statistical significance, effect size and study population) had a mean rating close to 1 (very important) and two (duration of observation and comparators) had a mean rating between 1 (very important) and 2 (quite important). The other characteristics included in this question (lead investigators, countries in which the trial is conducted and organisation sponsoring the trial) had mean ratings tending towards 3 (of little importance). The choice of primary endpoint (mean rating 1.15, SD = 0.49), statistical significance (1.22, SD = 0.46) and effect size (1.26, SD = 0.44) were the three most important design characteristics.

4.5.6 Importance of bisphosphonate trial endpoints

Respondents were asked to rank a predetermined list of bisphosphonate trial primary endpoints in order of importance with 1 being the most important endpoint and 8 being the least important. The results of responses to this question using the above coding are shown in Table A4.3.13. The two endpoints used in the discrete choice exercise, percentage of patients alive without disease recurrence and percentage of patients without metastatic bone disease, were ranked as the most important and

third most important endpoints respectively. The former had a mean ranking of 2.43 (SD = 1.80) and the latter 3.91 (SD = 1.94). The additional cost associated with the use of adjuvant bisphosphonates was ranked as the least important endpoint (mean 6.81, SD = 1.59).

Six respondents (12.97%) indicated that important endpoints were missing from the list provided (Table A4.3.14). These are shown, exactly as articulated by the respondents, in Table A4.3.18. It can be seen that 5 of the 11 omissions cited could be referred to as 'clinical' endpoints (e.g. serum calcium levels) and the remainder as 'economic' endpoints (e.g. cost per QALY).

4.6 RESULTS: DISCRETE CHOICE MODEL ESTIMATION

In this section, the results of the discrete choice model probit regression analysis are presented in terms of the qualitative and quantitative effects.

4.6.1 Qualitative effects

For each of the three models estimated, the signs on the attribute coefficients suggest identical qualitative effects (see Table 4.7). These are summarised below.

1) ***Choice of primary endpoint.*** The coefficient for this attribute (Endpoint) has a positive sign which implies a preference for disease free survival over the incidence of metastatic bone disease as the primary endpoint in adjuvant bisphosphonate trials. Consequently, this suggests that a product is more likely to be chosen if a trial demonstrates an improvement in the proportion of patients alive without disease recurrence compared with one

that shows an improvement in the incidence of metastatic bone disease.

2) ***Effectiveness.*** For the effectiveness attribute (Effectiveness), the coefficient has a positive sign suggesting that the probability of adopting a product is an increasing function of the level of effectiveness demonstrated, regardless of the choice of primary endpoint.

3) ***Degree of uncertainty surrounding the point estimate of effectiveness.*** The sign on the coefficient of the uncertainty variable (Uncertainty) is negative which indicates that the preference for a product decreases as the degree of uncertainty surrounding the point estimate of effectiveness increases.

4) ***Duration of observation.*** The positive sign on the coefficient for this attribute (Duration) suggests a preference for trials of longer durations. In other words, the probability of adopting a product is an increasing function of the duration of evaluation of its benefits.

5) ***Study population.*** A product whose benefits are demonstrated in a trial which enrolls patients with all stages of primary operable breast cancer is more likely to be chosen than one whose enrolment is restricted to subjects with Stage III disease at diagnosis. This is indicated by the positive sign on the study population coefficient (Population).

6) ***Incremental cost of adjuvant bisphosphonate use.*** The negative coefficient for the cost attribute (Cost) suggests that the lower the incremental cost of using a bisphosphonate prevention strategy the more

likely it is to be chosen.

To summarise the above findings, the qualitative effects (signs for the attribute coefficients) are in line with the author's prior expectations which provides evidence of the theoretical validity (internal consistency) of the estimated models.

4.6.2 Quantitative effects

Table 4.7 shows the primary results of this analysis. A Ramsey regression error specification test (RESET) suggests there is no problem with the functional form of any of the three models. The null hypothesis that all of the coefficients are simultaneously zero can be rejected on the basis of the Wald test ($p < 0.01$ in each case). Independently, attributes are statistically significantly different from zero at the 5% level or better with the exception of the duration variable which is borderline significant in both the 'full sample model' and the 'consistent traders sub-group A' model ($p = 0.06$ in both cases). These results indicate that each of the RCT design attributes included in the analysis is important in the decision to adopt adjuvant bisphosphonate treatments and that most respondents were willing to trade off different RCT design characteristics.

Table 4.7
Probit Regression Results[1,2]

RCT Design Attributes	(1) Full Sample	(2) Consistent Traders: Sub-group A[3]	(3) Consistent Traders: Sub-group B[4]
Endpoint	0.2787***	0.2700***	0.3185***
	[0.0522]	[0.0630]	[0.0625]
Effectiveness	0.0457***	0.0390***	0.0464***
	[0.0061]	[0.0068]	[0.0068]
Uncertainty	-0.6210***	-0.6653***	-0.7844***
	[0.0746]	[0.0987]	[0.0725]
Duration	0.0255*	0.0301*	0.0482***
	[0.0134]	[0.0159]	[0.0136]
Population	0.2419***	0.2675***	0.3292***
	[0.0563]	[0.0651]	[0.0704]
Cost	-5.43e-07***	-5.86e-07***	-5.95e-07***
	[5.97e-08]	[7.36e-08]	[6.97e-08]
Observations	864	608	656
Respondents	54	38	41
Log likelihood	-399.35	-285.18	-290.39
Wald chi2 (6)	238.95	163.01	241.27
Prob > chi2	0.0000***	0.0000***	0.0000***
Ramsey chi2(1)	0.30	0.10	0.01
Prob > chi2	0.5861	0.7485	0.9060
Correct predictions	79.28%	79.93%	81.10%

Notes.

1. Models were estimated using the probit (cluster) option available in Stata Version 7.0.[262] This estimator takes into account the potential non-independence of the observations and generates robust standard errors (shown in brackets).
2. Significance levels are denoted as follows: * significant at 10%; ** significant at 5%; *** significant at 1%
3. Inconsistent respondents were defined as those who did not choose the same scenario for choices 10 and 12 (n=10). Non-trading respondents were defined as those who exhibited dominant preferences for any attribute (n=7). Dropping these respondents left a sample of 38 consistent traders (one respondent was both inconsistent and a non-trader).
4. Inconsistent respondents were defined as those who did not choose the options expected for choices 8,9,13 or 14 (n=6). Non-trading respondents were defined as those who exhibited dominant preferences for any attribute (n=7). Dropping these respondents left a sample of 41 consistent traders.

The primary interest in this analysis is with using the regression results to compare the predicted probabilities of product adoption contingent upon alternative RCT designs. These can be computed using the regression results from Table 4.7 and Equations (7) and (8). Therefore sponsors of RCTs could use the regression results from product-specific stated preference surveys in a number of ways, including to:

1) Evaluate the impact of different RCT designs on the probability of product adoption;

2) Determine a technically feasible design which maximises the *expected* predicted probability of product adoption, and

3) Operationalise an investment appraisal approach to RCT design.[203]

Each of these uses is illustrated briefly in section 4.7 below.

4.7 USING DISCRETE CHOICE MODEL RESULTS IN RCT DESIGN

A number of potential uses of discrete choice modelling results in the context of RCT design are considered below. It must be emphasised that although this analysis is based around a case study of adjuvant bisphosphonate trials, the material presented below is purely illustrative. The main body of the text focuses on the results. The formulae, working assumptions and example calculations are presented in Appendix 4.4.

4.7.1 Impact of RCT designs on the probability of product adoption

One potential use of DCM results is to compare and rank alternative RCT designs in terms of the predicted probabilities of product adoption to which they give rise.

Specifically, given a set of candidate designs, sponsors of RCTs could use the results to select the design which gives the highest predicted probability of product adoption, $\Pr(A \mid \mathrm{J})$. This is equivalent to choosing the design with the highest decision-maker preference or utility score, ΔV_{AB}.

Table 4.8 and Figure 4.2 illustrate this application of the results by comparing the predicted probabilities of each of seven hypothetical candidate designs against a hypothetical baseline (existing) treatment. The differences in utility, ΔV_{AB}, are calculated by substituting the regression coefficients from the full sample model (Table 4.7) and the *differences* in the values of the RCT design attributes into Equation (8). The predicted probabilities are calculated by substituting the resulting utility values into Equation (7). An example calculation is provided in Appendix 4.4.

In Table 4.8, designs 1 to 6 differ from the baseline design only in terms of the level (value) of one RCT design attribute. This is done in order to illustrate how the impact on $\Pr(A \mid \mathrm{J})$ of changing the value of only one RCT design characteristic can be evaluated. For example, Design 1 differs from the baseline in terms of the choice of primary endpoint. This gives rise to a predicted probability of adoption of approximately 0.61. With a design identical to the baseline, the predicted probability would be 0.50. In contrast, hypothetical RCT Design 7 is defined as having the "best" attribute values shown in Table 4.2. This means that this design has "better" design characteristics than the baseline for all attributes except study population (which is the same). Consequently, Design 7 has the highest predicted probability amongst the RCT designs compared in Table 4.8. The ranking of the designs in descending order of their predicted probabilities is shown in the last row of Table 4.8.

Table 4.8: Impact of RCT Designs on the Probability of Product Adoption

Attribute	(1) Baseline product evidence	(2) New Product RCT Design 1	(3) New Product RCT Design 2	(4) New Product RCT Design 3	(5) New Product RCT Design 4	(6) New Product RCT Design 5	(7) New Product RCT Design 6	(8) New Product RCT Design 7
Endpoint	Patients without metastatic bone disease	Patients alive without disease recurrence	Patients without metastatic bone disease	Patients without metastatic bone disease	Patients without metastatic bone disease	Patients without metastatic bone disease	Patients without metastatic bone disease	Patients alive without disease recurrence
Effectiveness	2.90%	2.90%	40.00%	2.90%	2.90%	2.90%	2.90%	40%
Uncertainty	0.23% to 5.57% [0.92]	0.23% to 5.57% [0.92]	0.23% to 5.57% [0.92]	2.87% to 2.93% [0.01]	0.23% to 5.57% [0.92]	0.23% to 5.57% [0.92]	0.23% to 5.57% [0.92]	39.60% to 40.40% [0.01]
Duration	2 years	2 years	2 years	2 years	10 years	2 years	2 years	10 years
Population	Patients with Stages I to III disease at diagnosis	Patients with Stages I to III disease at diagnosis	Patients with Stages I to III disease at diagnosis	Patients with Stages I to III disease at diagnosis	Patients with Stages I to III disease at diagnosis	Patients with Stage III disease at diagnosis	Patients with Stages I to III disease at diagnosis	Patients with Stages I to III disease at diagnosis
Cost	£450,000	£450,000	£450,000	£450,000	£450,000	£450,000	£1,800,000	£0
Difference in utility (new minus existing product) ΔV_{AB}		0.2787[1]	1.6955[1]	0.5651[1]	0.2040[1]	-0.2419[1]	-0.7331[1]	2.9876[1]
Predicted probability $\Pr(A \mid J)$		0.6098[2]	0.9550[2]	0.7140[2]	0.5808[2]	0.4044[2]	0.2318[2]	0.9986[2]
Preference ranking of RCT design based on predicted probabilities	**6**	**4**	**2**	**3**	**5**	**7**	**8**	**1**

Footnotes to table can be found on next page.

Table 4.8 (continued)

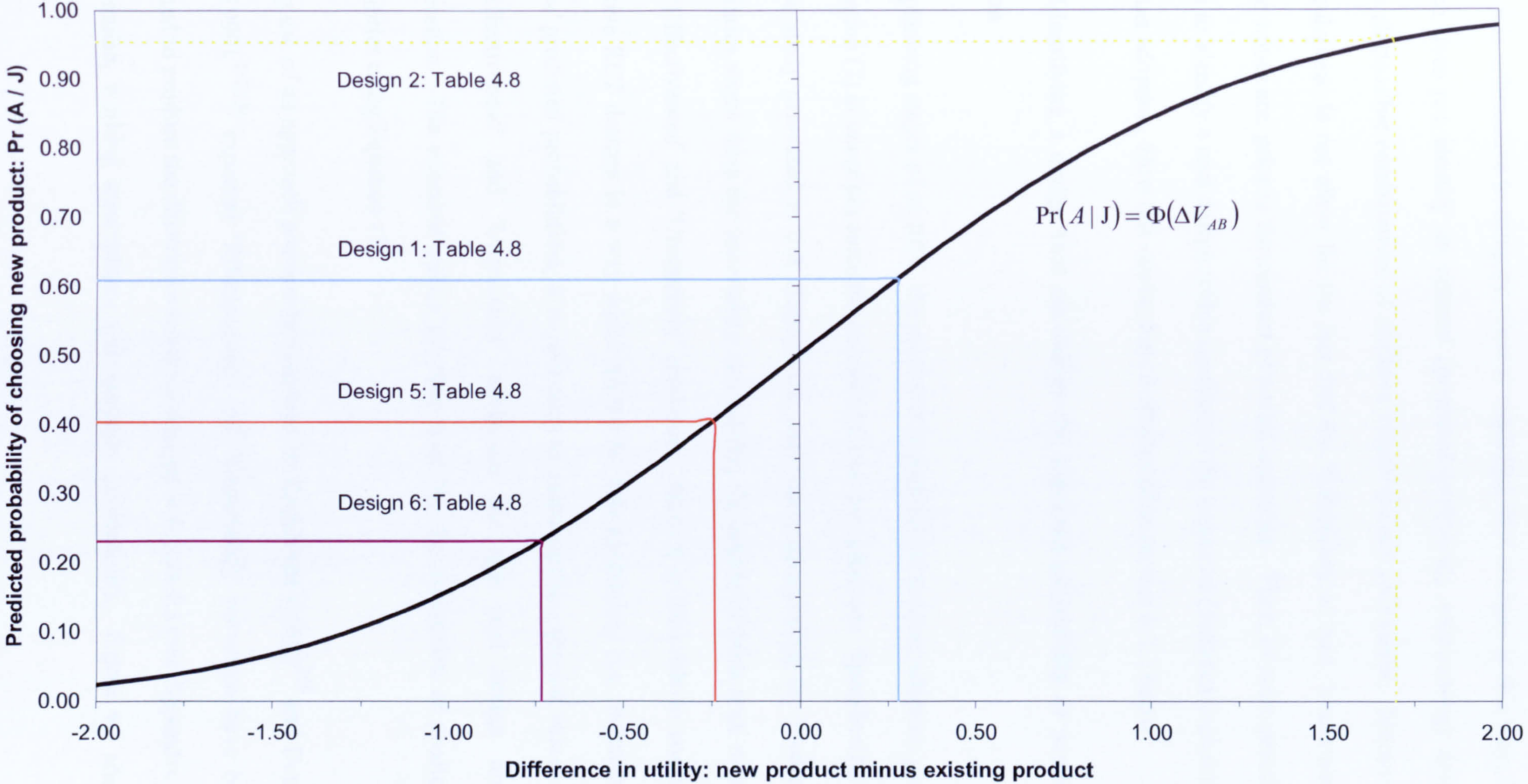
1. ΔV_{AB} is derived using the regression results and Equation (2). A denotes the new product and B denotes the existing product.

2. Derived by substituting the value ΔV_{AB} into Equation (1).

3. Figures in parentheses are the levels of the uncertainty attribute, which expresses the 95% confidence interval as a proportion of the point estimate of effectiveness.

4. See Appendix 4.4 for an example calculation.

Figure 4.2
Predicted probabilities of choosing a new product based on different RCT designs



Whilst it is informative to compare selected candidate RCT designs in this way, this approach does not identify an optimal (predicted probability maximising) design because all possible combinations of attribute values are not considered. Moreover, the calculations do not allow for the fact that the 'Effectiveness' and 'Uncertainty' attribute values are, prior to the conduct of a trial, uncertain. Thus, of more practical value is to identify a trial design which maximises the *expected* predicted probability of product adoption. How this can be done is illustrated in section 4.7.2 below.

4.7.2 Identifying a design that maximises the expected probability of product adoption

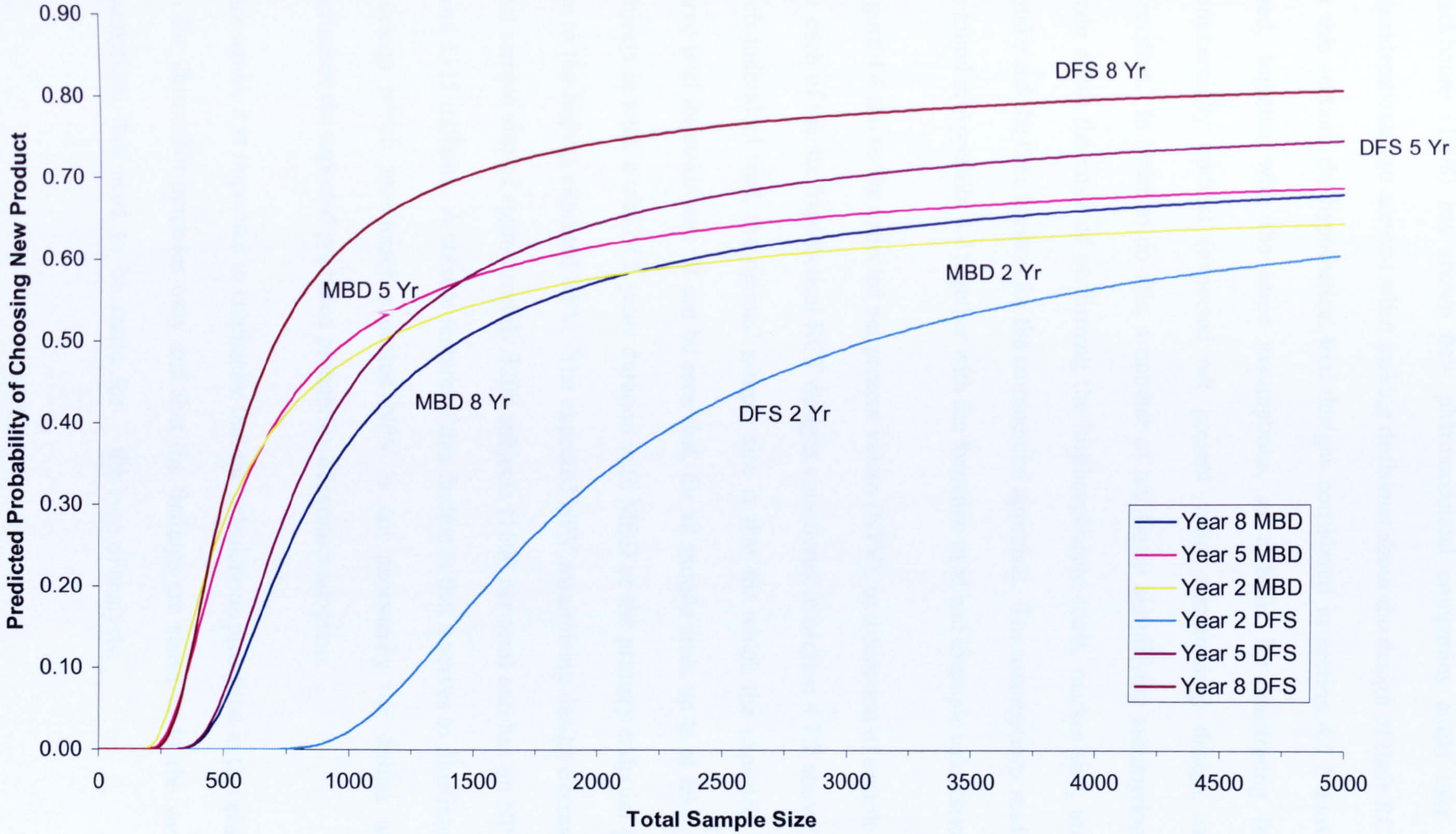
At the planning stages of an RCT, the predicted probability of product adoption given by Equation (7) is uncertain since the results of a trial are unknown. Specifically, for the case study presented in this chapter, the uncertainty surrounding the predicted probabilities stems from the uncertainty surrounding the results of trials with respect to the 'Effectiveness' and 'Uncertainty' attributes. Since it is desirable to consider alternative RCT designs in a way which allows for this uncertainty (i.e. in terms of *expected* predicted probabilities), it is necessary to calculate the *expected* values for the 'Effectiveness' and 'Uncertainty' outcomes for any trial design under consideration. The expected values are then used for the calculation of predicted probabilities using Equation (7).

Making use of an approach previously described by Backhouse (1998)[203] and Detsky (1985;1990),[152;154] expected 'Effectiveness' and 'Uncertainty' outcomes have been calculated to produce the illustrative results presented in Figure 4.3 (see Appendix 4.4 for formulae, working assumptions and example calculations). Figure 4.3 shows

expected predicted probabilities, over a range of trial sample sizes, for six hypothetical RCT designs when compared against the baseline treatment presented in Table 4.8. In order to simplify the exposition, the six designs differ from the baseline only in terms of i) the choice of primary endpoint (patients without metastatic bone disease (MBD)) or patients alive without disease recurrence (DFS)) and / or ii) the duration of the trial (2, 5 or 8 years). It can be seen that, upto a total sample size of approximately 440 subjects (220 per arm), a trial of 2 years duration with MBD as the primary endpoint gives rise to the highest expected predicted probability. Thereafter, a trial with 8 years of follow-up and DFS as the primary endpoint has the highest expected predicted probability of adoption. This is the design which maximises the expected predicted probabilities given the working assumptions. It should be emphasised that these results are purely illustrative and are sensitive to the assumptions made in their derivation, particularly the distributions of the prior expected outcomes (see Appendix 4.4).

A problem with this approach to RCT design is that, whilst an expected predicted probability maximising design can be identified, it may not be optimal from a commercial (profit maximising) perspective. This is because it does not take account of costs and time to market and hence the timing of revenues. It is therefore necessary to extend this analysis to consider the cost and revenue implications of alternative RCT designs.

Figure 4.3
Expected Predicted Probabilities for RCTs of 2, 5 and 8 Years Duration



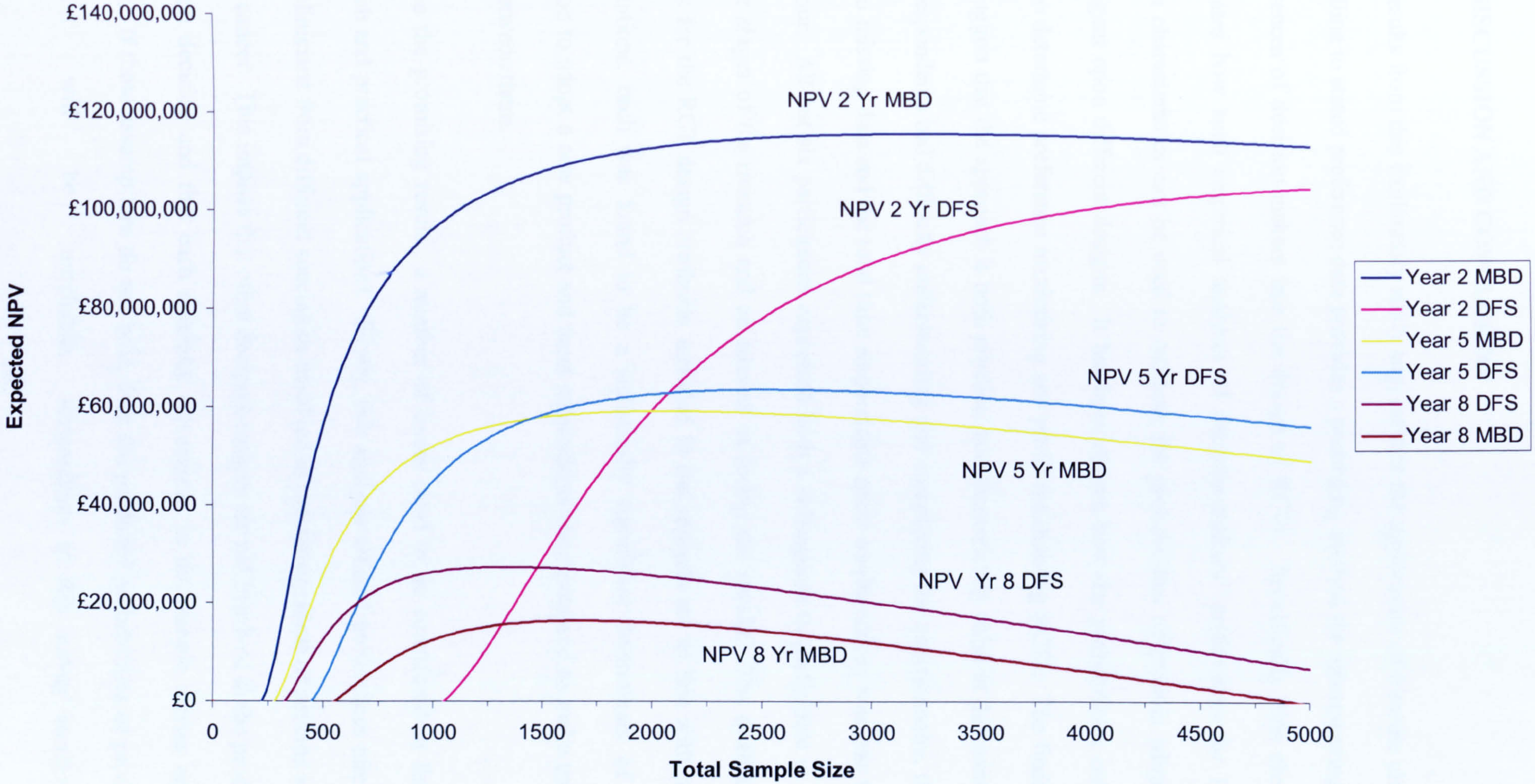
4.7.3 Using DCM results within an investment appraisal framework

Backhouse (1998) has shown how pharmaceutical companies could take profit considerations into account when making decisions about the design of their RCTs.[203] In this section, the hypothetical trial designs considered in section 4.7.2 above are used, together with the same assumptions, as a basis for illustrating how a commercially optimal (expected net present value maximising) design can be identified. In order to do this, a number of additional simplifying assumptions are made about the costs of performing the bisphosphonate trials, market size, product uptake and the time horizon for the commercial appraisal. The assumptions made can be found in Appendix 4.4 together with the formulae used and example calculations.

Figure 4.4 shows the expected net present values (NPV), as a function of sample size, for each of the six hypothetical RCT designs considered in section 4.7.2 above. For each individual trial, the optimal sample size is that for which the expected NPV curve is at its maximum. It can be seen that, for all sample sizes up to at least 5000 subjects in total, a trial of 2 years duration with MBD as the primary endpoint gives rise to the highest expected NPV. The expected NPV maximising design occurs at a total sample size of approximately 3200 subjects (1600 per arm) and has an NPV of about £115 millions. A notable feature of this finding is that it serves to illustrate that a design which maximises expected NPV is not necessarily the design which maximises the expected predicted probability of product adoption.

Once again, it is important to emphasise that the calculations presented in this chapter are for illustrative purposes only and that the findings are sensitive to the various assumptions that need to be made for this type of analysis.

Figure 4.4
Expected Net Present Values for RCTs of 2, 5 and 8 Years Duration



4.8 DISCUSSION AND CONCLUSION

The results from this exploratory study suggest that the application of discrete choice modelling to stated preference data provides a promising method for incorporating the preferences of decision-makers into the design of RCTs. Specifically, this chapter illustrates how such empirical analyses of decision-makers' preferences for RCT design characteristics can be used to estimate the probabilities of product adoption contingent upon different designs. It has been shown how the probabilities can be used to determine preference maximising and profit maximising RCTs. The findings also suggest that the approach is both practical and theoretically valid in this context. Few respondents had difficulty understanding and completing the questionnaire, there was no missing data and the total time respondents spent on the survey was less than one hour. All but six participants expressed both a willingness to participate in any further stages of the research and an interest in seeing the results. The qualitative effects for the RCT design attributes included in the analysis are in line with prior expectations, each was found to be a statistically significant determinant of the decision to adopt a new product and most respondents were prepared to make trade-offs between them.

Despite the promising results, a number of issues need to be considered in future research and practical applications. Firstly, this analysis utilised preferences elicited from clinicians from different specialties involved in the management of patients with breast cancer. This implies that other decision-makers are not involved in the product adoption decision and that each specialty represented in the sample carries equal weight. If these assumptions do not hold, then the predicted probabilities of product adoption will be unreliable. Respondents in this survey confirmed

the importance to the product adoption decision of the specialties that made-up most of the sample. But they also indicated that the influence of other parties should be considered, notably patients and other specialties involved in the management of breast cancer. If the preferences of physicians and patients are aligned, then the results which focus on the former will be robust. There would be practical challenges in applying this type of survey to patients because they may not be familiar with the terminology and practices of RCT design. This would be a valuable area for future research.

It is notable that very few respondents indicated that the influence of decision-making and advisory bodies such as formulary committees and NICE were important. Nevertheless, how such groups formulate their decisions and the preferences underlying them is both a fundamental and topical issue.[272;273] Therefore the potential for applying discrete choice analysis to members of such bodies would be a worthwhile line of future investigation since it offers a feasible means of explicitly quantifying the preferences of key stakeholder groups.

Secondly, the example application chosen for this study lent itself readily to the use of a simple binary choice model of drug prescribing. This is because adjuvant bisphosphonates are not currently established as a therapeutic strategy for the prevention of metastatic bone disease and so the choice problem could be simplified to the decision to use them or not contingent upon the RCT designs and results. Clearly, many treatment choice situations will be less straightforward as physicians are often presented with more than two possible courses of action. In such situations, it may be necessary to construct more complex multinomial models[204] of drug prescribing behaviour which would in turn require more complex stated

preference choice surveys for their application.[206] Moreover, the results from a survey conducted for one product indication will not be generalisable to another which could, for example, lead to a large number of studies required for a sponsor of multiple technologies. So the number, size and complexity of stated preference surveys would necessitate consideration being given to the potential benefits and costs associated with the research effort.

Thirdly, the stated preference survey presented respondents with a series of binary choices for which they were required to indicate a preference for one of the two adjuvant bisphosphonate prevention options. Respondents were not given the opportunity to indicate that they would prefer an alternative other than the two presented in any given choice set. In other words, they were not given the opportunity to 'opt-out'. A review of applications of discrete choice experiments to health care programmes confirms that this approach is consistent with previous practice.[274] However, it has recently been pointed out that the inclusion of an opt-out option may better mimic the circumstances under which actual choices are made and may therefore give rise to more reliable estimates of product or service adoption.[275] But there are also disadvantages of including opt-out alternatives. Subjects may choose the opt-out alternative simply to avoid making difficult trade-offs and it may not be possible to derive the attribute levels (characteristics) of the opt-out option.[275] Both of these factors could significantly reduce the number of observations available for analysis. Furthermore, research conducted outside the health field suggests that estimates of attribute weights and demand can be sensitive to the format of the opt-out alternative presented to respondents.[276] Clearly, further research is required into the issue of obtaining reliable predictions of actual choice behaviour from discrete choice

stated preference surveys. In this respect whether and how to include an opt-out alternative is one of a number of aspects to address.

Fourthly, the model parameters were estimated using discrete choice stated preference survey data which is currently the most common approach used in health economics applications. An alternative would be to use revealed preference data[273] i.e. data pertaining to actual rather than simulated choices. However, it may not be practical to obtain or construct a dataset containing the necessary RCT and product adoption variables and such data will clearly not be available for new products. It should also be noted that it may not be practical to conduct stated preference surveys amongst some decision-makers e.g. NICE appraisal committee members.

Finally, although aspects of this chapter have illustrated the potential use of DCA within a private sector investment appraisal framework, this should not detract from the potential value, in other contexts, of modelling product adoption decisions as a function of RCT design. For example, a useful line of future research would be to explore the conditions under which the private sector perspective on optimal RCT designs would be aligned with the societal perspective adopted by NICE. NICE considers both clinical effectiveness and cost in formulating its advice and its preferred measure for gauging value is the cost per quality adjusted life-year (QALY) (the incremental cost-effectiveness ratio).[272] Approaches to producing optimal trial designs from the societal perspective using cost-effectiveness criteria have been proposed.[155;159] The extent to which the private and societal perspectives will yield equivalent optimal designs will depend upon the importance of the cost-effectiveness ratio in product adoption decisions. Little is known about this relationship and although a recent paper used discrete choice modelling to produce insights

from recommendations made by NICE, the extent of the impact of the recommendations on actual product usage was not explored.[273] In this study, measures of both clinical effectiveness and cost were considered as separate variables but the cost per quality adjusted life year gained was not explicitly evaluated by respondents. However, in considering whether important endpoints were missing from the analysis, only two respondents mentioned the absence of cost per QALY information which raises questions about the alignment of physician and NICE decision-making criteria. Further research into how cost per QALY data could be presented in stated preference surveys would be a beneficial area for further research because it is not a measure that is widely understood amongst many stakeholders.

In conclusion, more sophisticated survey designs and statistical analysis methods may be required in future applications in order to correctly model the treatment decision-making situation of interest. Nevertheless, the results from this analysis suggest that DCA offers a practical and valid method by which sponsors of RCTs could take the preferences of decision-makers into account when planning their studies. Therefore further research into the application of the technique in this context would seem to be worthwhile.

Appendix 4.1
Effectiveness Outcomes Interpolated From Clinical Trials

	Clodronate[1]	Placebo[1]	Difference[2]	Lower 95%CI[3]	Upper 95%CI[3]
% Patients without metastatic bone disease					
From Powles et al (2002)					
Year 1	99.00	97.00	2.00	0.33	3.67
Year 2	96.20	93.30	2.90	0.23	5.57
Year 3	94.00	89.00	5.00	1.67	8.33
Year 4	91.00	87.50	3.50	-0.21	7.21
Year 5	89.00	84.50	4.50	0.45	8.55
Year 6	86.00	83.50	2.50	-1.81	6.81
Year 7	84.50	82.00	2.50	-1.97	6.97
Year 8	83.00	80.00	3.00	-1.65	7.65
From Diel et al (1998)					
Year 1	100.00	92.00	8.00	3.58	12.42
Year 2	98.00	88.00	10.00	4.28	15.72
Year 3	97.00	82.00	15.00	8.20	21.80
Year 4	92.00	78.00	14.00	6.03	21.97
Year 5	88.00	75.00	13.00	4.31	21.69
Year 6	78.00	75.00	3.00	-6.57	12.57
Year 7	78.00	75.00	3.00	-6.57	12.57
% Patients alive					
From Powles et al (2002)					
Year 1	98.00	98.00	0.00	-1.68	1.68
Year 2	92.70	92.40	0.30	-2.85	3.45
Year 3	90.00	87.00	3.00	-0.82	6.82
Year 4	86.50	84.00	2.50	-1.75	6.75
Year 5	82.90	79.30	3.60	-1.09	8.29
Year 6	81.00	76.50	4.50	-0.40	9.40
Year 7	78.50	73.00	5.50	0.37	10.63
Year 8	78.00	72.00	6.00	0.82	11.18
Year 9	74.00	65.50	8.50	3.02	13.98
Year 10	74.00	60.00	14.00	8.43	19.57
From Diel et al (1998)					
Year 1	100.00	82.00	18.00	11.75	24.25
Year 2	95.00	78.00	17.00	9.44	24.56
Year 3	90.00	72.00	18.00	9.31	26.69
Year 4	80.00	65.00	15.00	5.03	24.97
Year 5	80.00	63.00	17.00	6.95	27.05
Year 6	75.00	60.00	15.00	4.54	25.46
Year 7	75.00	60.00	15.00	4.54	25.46

1. Data points were interpolated from the survival curves reported in Powles et al (2002)[269] and Diel et al (1998).[268]
2. Clodronate % minus placebo %.
3. Confidence intervals for the differences in % effectiveness were calculated using the formula provided by Armitage & Berry (1995) pp 128-130.[170]

APPENDIX 4.2

THE STATED PREFERENCE SURVEY QUESTIONNAIRE

Our reference

Your reference

Direct line/e-mail

Nottingham University Business School

Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB
Tel: +44 (0) 115 846 6602
Fax: +44 (0) 115 846 6667
http://www.nottingham.ac.uk/business

8 March 2002

Dear [Doctor]

The Use of Conjoint Analysis in the Design of Clinical Trials

Thank you very much for agreeing to take part in this research.

I am a part-time PhD student at the University of Nottingham. As part of my research I am conducting a survey to assess how a technique known as conjoint analysis might be used to take into account the views of health care professionals when designing clinical trials. The work is not being conducted on behalf of any sponsoring organisation or company.

The research will include interviews with specialists like you and I have asked Accent Marketing and Research to conduct these interviews on my behalf.

I should be grateful if you would assist me with this research by spending about 10 minutes reading the enclosed material. Then on the (insert date), one of Accent's researchers will telephone you to collect your responses to each of the questions. Hence, this material does not need to be returned to me.

The questionnaire does not require you to provide any personal or patient information. Furthermore, Accent will not pass on the names of those who participate in this research to me unless you give your consent for this to happen.

A copy of the results of this survey will be available for all those who have taken part in this research.

If you have any questions relating to the enclosed, please do not hesitate to contact me.

Thank you in advance for your help with this research.

Yours sincerely,

Martin E Backhouse

Enc.

INTRODUCTION

Clinical research has shown that bisphosphonates reduce the incidence of hypercalcaemia and pathological bone fractures in patients with established bone metastases from breast cancer. Moreover, bisphosphonates have been shown to reduce the risk of bone metastases in patients with relapsed breast cancer without obvious bone involvement. However, the effectiveness of adjuvant bisphosphonates as a preventive therapeutic strategy for patients with primary operable breast cancer has yet to be definitively established.

In the choices which follow, you are asked to imagine that **you alone** are deciding which adjuvant bisphosphonate therapy to use based on the trial evidence which is presented. For each choice, you will be asked to compare two alternatives (labelled 'Bisphosphonate Prevention A' and 'Bisphosphonate Prevention B'), which differ only in terms of the following trial design characteristics and results:

- **Primary endpoint:** the main measure chosen to compare the effectiveness of adjuvant bisphosphonate therapy against no such therapy (placebo) in patients with primary operable breast cancer.
- **Difference in % of patients achieving the primary endpoint:** the effectiveness of adjuvant bisphosphonate therapy measured as the <u>difference</u> between the % of patients experiencing the primary endpoint in the 'bisphosphonate' and 'no bisphosphonate' arms of the trial i.e. adjuvant bisphosphonate % ***minus*** no adjuvant bisphosphonate %. The results are also shown in the form of the number of patients that would need to be treated with bisphosphonates in order for one patient to benefit from treatment i.e. number needed to treat (NNT).
- **95% confidence interval on the primary endpoint:** a measure of the uncertainty surrounding the point estimate of the primary endpoint outcome. A range of % difference values is presented within which there is a 95% chance that the true difference will lie. The 95% confidence interval is also shown in the form of the number of patients that would need to be treated with bisphosphonates in order for one patient to benefit from treatment i.e. number needed to treat (NNT).
- **Duration of observation:** the duration of the trial in years (not the duration of adjuvant bisphosphonate therapy). It is assumed that all subjects are followed for this period of time. The primary endpoint results are those observed at the end of this follow-up period.
- **Disease stage at diagnosis:** the eligible study population defined in terms of the stage of primary operable breast cancer at diagnosis (Stages I to III).
- **Additional cost of using adjuvant bisphosphonates:** the ***additional cost*** of using adjuvant bisphosphonates compared with not using them i.e. adjuvant bisphosphonate cost ***minus*** no adjuvant bisphosphonate cost. The cost figure presented is the difference per **<u>100 patients</u>** for the period of the trial. In the choices which follow, no information is provided about the duration of adjuvant bisphosphonate therapy i.e. the cost information can reflect different agents and different durations of bisphosphonate treatment.

In making your choices you should assume that:

1) The efficacy results presented are statistically significant at the 5% level.

2) The evidence comes from well-designed, randomised, double-blind placebo controlled trials in patients with primary operable breast cancer.

3) The evidence is the only evidence that is available to make your decision.

4) The alternatives differ only in terms of the characteristics which are presented.

5) Subjects in both arms of the trial received surgery, chemotherapy, hormonal therapy and radiotherapy as required.

6) In the event of relapse, appropriate local or systemic therapies (including bisphosphonates) were administered as required to subjects in both arms of the trial.

7) The adjuvant bisphosphonates were well tolerated i.e. no significant side effects were observed.

PART A

In this part of the questionnaire you are presented with 16 choices. In each case, you are asked to choose **only one** of the two adjuvant bisphosphonate treatment strategies for patients with primary operable breast cancer. Please indicate your choice by marking a ✔in the appropriate box as shown in the following example:

Example:

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients alive without disease recurrence	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	10% [NNT = 10]	1% [NNT = 100]
95% confidence interval on the primary endpoint	2.50% to 17.50% [NNT = 5.71 to 40.00]	0.75% to 1.25% [NNT = 80.00 to 133.33]
Duration of observation	10 years	8 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 1,800,000	£ 1,800,000

	Prefer Prevention A	Prefer Prevention B
Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	✔	

Now please complete the following choice questions making sure that you choose one option for each of the 16 choices.

CHOICE 1

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients alive without disease recurrence	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	10% [NNT = 10]	1% [NNT = 100]
95% confidence interval on the primary endpoint	2.50% to 17.50% [NNT = 5.71 to 40.00]	0.75% to 1.25% [NNT = 80.00 to 133.33]
Duration of observation	10 years	8 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 1,800,000	£ 1,800,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 2

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients without metastatic bone disease	Patients alive without disease recurrence
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	10% [NNT = 10]	25% [NNT = 4]
95% confidence interval on the primary endpoint	9.90% to 10.10% [NNT = 9.90 to 10.10]	6.25% to 43.75% [NNT = 2.29 to 16.00]
Duration of observation	8 years	8 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 900,000	£ 0

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 3

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients alive without disease recurrence	Patients alive without disease recurrence
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	10% [NNT = 10]	1% [NNT = 100]
95% confidence interval on the primary endpoint	7.50% to 12.50% [NNT = 8.00 to 13.33]	0.01% to 1.99% [NNT = 50.25 to 10000.00]
Duration of observation	4 years	10 years
Disease stage at diagnosis for patients enrolled in the trial	Stages I, II and III	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 450,000	£ 900,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 4

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients without metastatic bone disease	Patients alive without disease recurrence
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	10% [NNT = 10]	1% [NNT = 100]
95% confidence interval on the primary endpoint	0.10% to 19.90% [NNT = 5.03 to 1000.00]	0.99% to 1.01% [NNT = 99.01 to 101.01]
Duration of observation	2 years	4 years
Disease stage at diagnosis for patients enrolled in the trial	Stages I, II and III	Stage III only
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 0	£ 0

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 5

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients alive without disease recurrence	Patients alive without disease recurrence
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	25% [NNT = 4]	25% [NNT = 4]
95% confidence interval on the primary endpoint	6.25% to 43.75% [NNT = 2.29 to 16.00]	18.75% to 31.25% [NNT = 3.20 to 5.33]
Duration of observation	8 years	2 years
Disease stage at diagnosis for patients enrolled in the trial	Stages I, II and III	Stage III only
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 0	£ 900,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 6

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients without metastatic bone disease	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	25% [NNT = 4]	40% [NNT = 2.5]
95% confidence interval on the primary endpoint	24.75% to 25.25% [NNT = 3.96 to 4.04]	10.00% to 70.00% [NNT = 1.43 to 10.00]
Duration of observation	10 years	4 years
Disease stage at diagnosis for patients enrolled in the trial	Stages I, II and III	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 450,000	£ 900,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 7

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients alive without disease recurrence	Patients alive without disease recurrence
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	25% [NNT = 4]	10% [NNT = 10]
95% confidence interval on the primary endpoint	18.75% to 31.25% [NNT = 3.20 to 5.33]	2.50% to 17.50% [NNT = 5.71 to 40.00]
Duration of observation	2 years	10 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stage III only
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 900,000	£ 1,800,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 8

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients without metastatic bone disease	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	25% [NNT = 4]	40% [NNT = 2.5]
95% confidence interval on the primary endpoint	0.25% to 49.75% [NNT = 2.01 to 400.00]	30.00% to 50.00% [NNT = 2.00 to 3.33]
Duration of observation	4 years	10 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stage III only
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 1,800,000	£ 0

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 9

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients without metastatic bone disease	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	40% [NNT = 2.5]	25% [NNT = 4]
95% confidence interval on the primary endpoint	10.00% to 70.00% [NNT = 1.43 to 10.00]	0.25% to 49.75% [NNT = 2.01 to 400.00]
Duration of observation	4 years	4 years
Disease stage at diagnosis for patients enrolled in the trial	Stages I, II and III	Stage III only
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 900,000	£ 1,800,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 10

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
	Patients alive without disease recurrence	Patients alive without disease recurrence
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	40% [NNT = 2.5]	40% [NNT = 2.5]
95% confidence interval on the primary endpoint	39.60% to 40.40% [NNT = 2.48 to 2.53]	0.40% to 79.60% [NNT = 1.26 to 250.00]
Duration of observation	2 years	8 years
Disease stage at diagnosis for patients enrolled in the trial	Stages I, II and III	Stage III only
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 1,800,000	£ 450,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 11

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients without metastatic bone disease	Patients alive without disease recurrence
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	40% [NNT = 2.5]	10% [NNT = 10]
95% confidence interval on the primary endpoint	30.00% to 50.00% [NNT = 2.00 to 3.33]	7.50% to 12.50% [NNT = 8.00 to 13.33]
Duration of observation	10 years	4 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 0	£ 450,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 12

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients alive without disease recurrence	Patients alive without disease recurrence
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	40% [NNT = 2.5]	40% [NNT = 2.5]
95% confidence interval on the primary endpoint	0.40% to 79.60% [NNT = 1.26 to 250.00]	39.60% to 40.40% [NNT = 2.48 to 2.53]
Duration of observation	8 years	2 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 450,000	£ 1,800,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 13

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients without metastatic bone disease	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	1% [NNT = 100]	25% [NNT = 4]
95% confidence interval on the primary endpoint	0.25% to 1.75% [NNT = 57.14 to 400.00]	24.75% to 25.25% [NNT = 3.96 to 4.04]
Duration of observation	2 years	10 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 450,000	£ 450,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 14

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients alive without disease recurrence	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	1% [NNT = 100]	1% [NNT = 100]
95% confidence interval on the primary endpoint	0.99% to 1.01% [NNT = 99.01 to 101.01]	0.25% to 1.75% [NNT = 57.14 to 400.00]
Duration of observation	4 years	2 years
Disease stage at diagnosis for patients enrolled in the trial	Stage III only	Stage III only
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 0	£ 450,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 15

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients without metastatic bone disease	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	1% [NNT = 100]	10% [NNT = 10]
95% confidence interval on the primary endpoint	0.75% to 1.25% [NNT = 80.00 to 133.33]	0.10% to 19.90% [NNT = 5.03 to 1000.00]
Duration of observation	8 years	2 years
Disease stage at diagnosis for patients enrolled in the trial	Stages I, II and III	Stages I, II and III
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 1,800,000	£ 0

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

CHOICE 16

Trial Design Characteristics	Bisphosphonate Prevention A	Bisphosphonate Prevention B
Primary endpoint	Patients alive without disease recurrence	Patients without metastatic bone disease
Difference in % of patients achieving primary endpoint at the end of the trial (bisphosphonate *minus* placebo)	1% [NNT = 100]	10% [NNT = 10]
95% confidence interval on the primary endpoint	0.01% to 1.99% [NNT = 50.25 to 10000.00]	9.90% to 10.10% [NNT = 9.90 to 10.10]
Duration of observation	10 years	8 years
Disease stage at diagnosis for patients enrolled in the trial	Stages I, II and III	Stage III only
Additional cost of using adjuvant bisphosphonates per 100 patients treated (bisphosphonate *minus* placebo)	£ 900,000	£ 900,000

Which adjuvant bisphosphonate prevention option would you prefer for patients presenting with primary operable breast cancer (Stage I to Stage III)?	Prefer Prevention A	Prefer Prevention B

PART B

1. What is your area of specialisation (please tick one of the following):

Medical Oncologist	☐	Radiologist	☐
Surgical Oncologist	☐	Radiotherapist	☐
General Practitioner	☐	Pharmacist	☐
Other	☐	Please specify:.....	

2. Approximately how many **new cases** of breast cancer do you see each year? ☐☐☐

3. Of the **new cases** of breast cancer that you see each year, approximately what percentage have the following stages of disease at diagnosis:

Per cent of new cases with **Stage I** disease at diagnosis	=☐☐☐%
Per cent of new cases with **Stage II** disease at diagnosis	=☐☐☐%
Per cent of new cases with **Stage III** disease at diagnosis	=☐☐☐%
Per cent of new cases with **Stage IV** disease at diagnosis	=☐☐☐%
Please check that the total adds to 100%	=☐☐☐%

4. In deciding whether to start using adjuvant bisphosphonates in patients with primary operable breast cancer, please indicate with a ✓ the degree of influence that you think each of the following specialties would have on the decision:

Specialty	High degree of influence	Some influence	No influence
Radiologist			
Medical Oncologist			
Radiotherapist			
Surgical Oncologist			
Pharmacist			
General Practitioner			

5. Are there any important decision-makers or influences missing from the list provided in Question 4 ? (please tick Yes or No) :

Yes ☐

No ☐

If Yes, please specify: .....
.....
.....
.....
.....

6. When considering the evidence from a clinical trial relating to the use of adjuvant bisphosphonates in patients with primary operable breast cancer, please indicate with a ✓ the importance to you of the following trial design characteristics:

Trial characteristic	**Very important**	**Quite important**	**Of little importance**	**Not important**
Primary endpoint				
Comparator				
Study population				
Duration of follow-up				
Size of effect demonstrated				
Statistical significance of results				
Organisation sponsoring the trial				
Countries in which the trial is conducted				
Lead investigators				

7. If you were designing a clinical trial to inform you whether to use adjuvant bisphosphonates in patients with primary operable breast cancer, please rank the following endpoints in order of importance from 1 (the most important endpoint to you) to 8 (the least important endpoint to you):

Side effects ☐

% patients alive without disease recurrence ☐

Quality of life experienced by patients ☐

% patients alive ☐

Cost of patient management with bisphosphonates ☐

% patients without metastatic bone disease ☐

% patients without non-skeletal metastases ☐

% patients not experiencing skeletal morbidity ☐

8. Are there any important endpoints missing from the list provided in Question 7 ? (please tick Yes or No) :

Yes ☐

No ☐

If Yes, please specify: .....
.....
.....
.....
.....

9. Do you have any responsibility for managing budgets related to the treatment of patients with breast cancer? (please tick Yes or No):

Yes ☐

No ☐

If Yes, please provide a brief description of your responsibilities:
.....
.....
.....
.....
.....

10. Did you find this questionnaire:

Very difficult to complete ☐

Moderately difficult to complete ☐

Slightly difficult to complete ☐

Not difficult to complete ☐

11. Please provide any comments you would like to make about this questionnaire below:

.....

.....

.....

.....

.....

.....

.....

.....

.....

12. How long has it taken you to complete this questionnaire? ☐☐☐ minutes

Please check that you have answered all the questions and then return this questionnaire in the envelope provided.

Thank you for completing this questionnaire.

APPENDIX 4.3
RESULTS OF THE NON-CHOICE COMPONENTS OF THE STATED PREFERENCE SURVEY
(QUESTIONNAIRE PART B)

Table A4.3.1

Question 1: What is your area of specialisation?

Specialty	Title					
	Professor	Consultant	Senior Registrar	Registrar	Other	Total
Medical oncologist	1	5	3	3	2	14
Surgical oncologist	1	15	0	1	0	17
Clinical oncologist	0	11	0	4	1	16
Other	0	4	0	0	3	7
Total	2	35	3	8	6	54

Table A4.3.2

Question 2: Approximately how many new cases of breast cancer do you see each year?

Specialty	Summary of new cases		
	Mean	Standard deviation	Frequency
Medical oncologist	156.86	112.25	14
Surgical oncologist	144.29	65.36	17
Clinical oncologist	224.38	188.82	16
Other	183.57	203.32	7
Total	176.37	142.42	54

Table A4.3.3

Question 3: Approximately what percentage of new cases have the following stages of disease at diagnosis?

Stage	Observations	Mean	Standard deviation	Min	Max
% Stage I	54	38.24	22.63	0	90
% Stage II	54	30.19	15.42	0	60
% Stage III	54	15.83	11.89	0	50
% Stage IV	54	12.04	14.84	0	80
Stage I	54	69.66	65.54	0	360
Stage II	54	55.04	44.59	0	213
Stage III	54	30.24	50.02	0	340
Stage IV	54	21.43	29.12	0	170

Table A4.3.4

Question 9: Do you have any responsibility for managing budgets related to the treatment of patients with breast cancer?

Specialty	Budget responsibility?		
	Yes	No	Total
Medical oncologist	5	9	14
Surgical oncologist	5	12	17
Clinical oncologist	4	12	16
Other	0	7	7
Total	14	40	54

Table A4.3.5

Question 10: How difficult was this questionnaire to complete?

Specialty	Difficulty of questionnaire to complete				
	Very difficult	Moderately difficult	Slightly difficult	Not at all difficult	Total
Medical oncologist	0	4	6	4	14
Surgical oncologist	1	6	9	1	17
Clinical oncologist	0	7	3	6	16
Other	1	5	0	1	7
Total	2	22	18	12	54

Table A4.3.6

Question 12: How long have you spent on this questionnaire?*

Variable	Observations	Mean	Standard deviation	Min	Max
Reviewing time	54	26.76	13.11	0[1]	60
Interview time	54	11.41	4.56	0[2]	26.32
Total time	54	38.17	13.32	19	80

* Times are in minutes.
1. The minimum of zero was caused by one respondent reporting no preparation prior to interview.
2. The minimum of zero was caused by one respondent mailing responses but not participating in the interview.

Table A4.3.7

Are you happy to have your personal details disclosed?

Specialty	Disclosure of personal details		
	Yes	No	Total
Medical oncologist	12	2	14
Surgical oncologist	15	2	17
Clinical oncologist	14	2	16
Other	7	0	7
Total	48	6	54

Table A4.3.8

Would you like to be sent a copy of the results of this survey?

Specialty	Receive copy of the survey?		
	Yes	No	Total
Medical oncologist	9	5	14
Surgical oncologist	17	0	17
Clinical oncologist	15	1	16
Other	7	0	7
Total	48	6	54

Table A4.3.9

Question 11: Do you have any comments on the questionnaire?

Specialty	Comments on questionnaire?		
	Yes	No	Total
Medical oncologist	5	9	14
Surgical oncologist	13	4	17
Clinical oncologist	8	8	16
Other	4	3	7
Total	30	24	54

Table A4.3.10

Question 4: Degree of influence of different specialties on the decision to use adjuvant bisphosphonates?

Specialty	Observations	Mean	Standard deviation	Min	Max
Medical oncologist	54	1.15	0.41	1	3
Radiotherapist	54	1.35	0.55	1	3
Surgical oncologist	54	1.80	0.59	1	3
Pharmacist	54	2.57	0.57	1	3
Radiologist	54	2.63	0.56	1	3
GP	54	2.63	0.52	1	3

Table A4.3.11

Question 5: Are there any important decision makers or influences missing from the list in Question 4?

Specialty	Missing influences?		
	Yes	No	Total
Medical oncologist	7	7	14
Surgical oncologist	11	6	17
Clinical oncologist	11	5	16
Other	7	0	7
Total	36	18	54

Table A4.3.12

Question 6: Importance to you of the following design characteristics of an adjuvant bisphosphonates trial

Variable	Observations	Mean	Standard deviation	Min	Max
Primary endpoint	54	1.15	0.49	1	4
Statistical significance	54	1.22	0.46	1	3
Effect size	54	1.26	0.44	1	2
Study population	54	1.44	0.57	1	3
Duration	54	1.50	0.50	1	2
Comparator	54	1.57	0.57	1	3
Lead investigator	54	2.57	0.69	1	4
Countries	54	2.59	0.69	1	4
Sponsor	54	2.80	0.68	1	4

Table A4.3.13

Question 7: Ranking of importance of adjuvant bisphosphonate trial endpoints

Variable	Observations	Mean	Standard deviation	Min	Max
Disease free survival	54	2.43	1.80	1	8
Alive	54	3.78	2.45	1	8
No metastatic bone disease	54	3.91	1.94	1	7
Quality of life	54	4.11	2.09	1	8
Side effects	54	4.63	1.88	1	8
No skeletal morbidity	54	4.98	2.05	1	8
No other metastases	54	5.35	1.82	2	8
Cost	54	6.81	1.59	3	8

Table A4.3.14

Question 8: Are there any important endpoints missing from the list in Question 7?

Specialty	Missing endpoints?		
	Yes	No	Total
Medical oncologist	1	13	14
Surgical oncologist	1	16	17
Clinical oncologist	4	12	16
Other	1	6	7
Total	7	47	54

Table A4.3.15

Budget responsibility as articulated by respondents

Respondent	Comment
13	For the surgical side of things but not for the drugs.
14	As lead clinician I have some input into where our significant expenditure should be.
15	I start the treatment and see the patient through to when they die.
16	Not directly but we all have some influence. We have the North Trent Breast Care Group which we all have input in and decisions taken through this.
19	I'm the lead clinician for cancer. Separately responsible for prioritising money for the cancer agenda trust.
20	I'm Director of Surgery with a budget of £15millions. Also I'm the lead Cancer Clinician for the hospital. The hospitals overall budget is £115millions.
27	Indirectly in an advisory capacity re drugs and radiation therapy.
31	I sit on the Network Committee and we make decisions about where the money will go.
35	I was Head of Department and made some decisions regarding drugs to be used. Otherwise decisions are joint with other consultants.
36	I'm Clinical Director of Royal Free University College. I Chair the Breast Tumour Board for North London Network.
40	Answered no but made the comment: "Only priority setting at consultant meetings."
42	I sit on the Joint Hospital Board. Decide which drugs we will use.
43	Formulary sub-committee. Chairman of Cancer Network Systems, Network Therapeutics Group: dealing with all new cancer drugs.
44	Answered no but made the comment: "but on consultants' committee."
46	Answered no but made the comment: "but we have to prescribe responsibly within evidence based guidelines."
50	I'm involved in the Hospital Pharmacy Committee and the High Cost Drug Committee.

Table A4.3.16

Comments on the questionnaire as articulated by respondents

Respondent	Comment
1	Interesting.
3	Only that I'm curious to see what endpoint. For me it's been a useful introduction to conjoint analysis from a learning point of view.
4	Certain of them not comparing like with like.
5	Didn't ask how often the patient needed treatment or the type i.e. whether it was tablet or iv. If iv then how often. The interval between treatments by this method is very important to the patient.
9	I found it more difficult to do Showcard 1 because there was no information on the number of patients in the trials.
10	The questionnaire could have been clearer in certain aspects. Each showcard has too much information in order to come up with a choice.
11	I was interested in some of the things being compared in that they don't seem really comparable.
13	It's a concept I hadn't actually appreciated.
16	First lot of questions not that easy: too much information to take in-surgeons are a bit thick! However, it was good – I think it's very important trying to get across what's important in clinical trials.
17	You seem to have covered everything.
19	No – it was quite interesting. I'm interested in bisphosphonates.
21	I found it a bit difficult with some of the choices reconciling them in my mind.
23	No but 10 minutes is unrealistic.
24	There were two problems for me: 1) We were asked to decide prevention strategy for all stages yet a lot of the data was only for Stage 3. 2) I wouldn't make a decision on a single set of data given like this. It was a very false way of looking at scientific data and I was very unhappy with it. I would be very happy for him to contact me to discuss this further.
27	The analysis of the data – it's the first time I've come across this type of vehicle at Showcard 1. I found it quite a useful exercise.
28	Only that on Showcard 2 we don't use this terminology (Stage I etc) – it's American but it didn't bother me unduly.
30	1. There is one important factor in the decision to use bisphosphonates which is not featured anywhere: the need for the staffing and infrastructure to give the treatments, especially if it is being done intravenously. 2. The way Showcard 1 was devised, I feel sure I have contradicted myself at times.
31	Some of the options in Showcards 1-16 don't look very feasible. Some of the scenarios are a bit difficult to understand how a trial can be designed this way. I can't get my head round why they've been written this way. Willing to participate in the next phase.

Table A4.3.16 (continued)

Respondent	Comment
32	I was intrigued by the format. I understand the research was about conjoint analysis but it seemed to be about the use of bisphosphonates which is a pretty controversial subject at the moment.
35	I didn't know what conjoint analysis was-I looked it up on the web. I found assigning values to the different characteristics listed was difficult.
36	Quite an interesting one.
37	I thought it was very well designed.
38	Some of the cost differences are very large-some had zero! It was fully comprehensive but the issue of patients alive without disease recurrence and patients without metastatic bone disease made it difficult to way up when you've got different endpoints.
40	I just found it difficult on the 1-16 choices, to make sure I'd noted the differences on each one. It was the number of choices.
41	It's a lot of fun. Quite challenging. Half the choice questions were very easy and half I didn't feel either option was acceptable but came down on one for the purpose of this exercise. 1) You know what % of stage III patients will be alive at a particular time point and that will naturally affect the way you look at it but you're told not to have any other information. 2) How desirable is the outcome? How likely is the outcome? Do you think your intervention is going to count on the outcome?
42	I've not done anything like this before. I found it very interesting.
44	It is not easy really. I have to think about the formulation of the questions and I can't come up with any bright ideas.
45	It took longer than 20 minutes. I was told it would take 10 minutes.
46	Good questionnaire. Have done this before.
50	Respondent wished to point out that 90% of breast cancer patients are treated by clinical oncologists and if they are called radiotherapists it could upset a lot of people. The situation is politically very sensitive. The economic analysis needs to be good.

Table A4.3.17

Missing decision makers or influences as articulated by respondents

Respondent	Comment
1	Pathologists.
2	Nurse specialist and the patient.
3	Nurse specialists or research nurses.
5	Patient.
6	Medical endocrinologist.
7	1. Clinical nurse specialists. 2. Pathologists.
8	Pathologist.
9	Breast nurse.
10	Breast care nurses.
11	Clinical nurse specialist – has some influence.

Table A4.3.17 (continued)

13	Breast care nurse.
17	Trust manager.
19	The respondent had originally indicated NICE but then deleted and changed the answer to no important decision-makers missing.
20	Palliative care people.
21	Orthopaedic surgeon.
23	Breast reconstruction surgeon.
24	Policy makers = managers.
25	Patient and relatives.
26	Breast care nurse.
27	The patient.
31	Pathologist.
34	1. Palliative medicine consultant. 2. Hospice consultant.
36	1. Patient support groups (some influence). 2. NICE (high degree of influence).
38	1. Plastic surgeon. 2. Palliative care consultants. 3. Breast care nurses/MacMillan nurses.
39	1. Breast care nurse. 2. Palliative care team i.e. MacMillan nurse. 3. Orthopaedic surgeon.
40	Patients.
41	1. The patient. 2. The media. 3. Specialist nurse.
42	Breast care sister.
43	NICE.
44	Breast care nurses.
46	Primary care Trusts as they have to fund increased costs.
47	The patient.
48	1. The patient. 2. The patient's relatives. 3. The breast care nurses.
49	1. Clinical chemist. 2. Rheumatologist
50	Answered no but made the comment: "You need to reframe radiotherapist as a clinical oncologist."

Table A4.3.18

Missing endpoints as articulated by respondents

Respondent	Comment
8	1.Serum calcium levels. 2.Number of pathological bone fractures.
31	1. Quality adjusted life-years 2. Cost per QALY
41	1. % of patients with spinal cord compression – this can go undetected. 2. Nothing about disability or time spent in hospital.
48	1. Health economics assessment. 2. Bone density assessment.
49	Cost of patient management without bisphosphonates.
50	1. Some management can reduce costs. 2. Endpoint 5 (cost): "Is this overall management cost? It's unclear."

APPENDIX 4.4

TECHNICAL APPENDIX

The purpose of this Technical Appendix is to set out the assumptions, equations and sources of data used for the illustrative analyses and results reported in section 4.7, "Using discrete choice model results in RCT design". Example calculations are provided. Although the examples draw on published data pertaining to a recently reported bisphosphonate trial[269], the calculations are illustrative and do not purport to solve an optimisation problem in this context.

1. Interpolation of effectiveness outcomes

In order to perform the illustrative calculations presented in this paper, effectiveness and uncertainty outcomes were interpolated, for three time points, from the survival curves reported in Powles et al (2002).[269] These are presented in Table A4.4.1 below.

Table A4.4.1

	Clodronate[1]	Placebo[1]	Difference[2]	Lower 95% CI[3]	Upper 95% CI[3]	Prior[4]
% Patients without metastatic bone disease						
Year 2	96.20	93.30	2.90	0.23	5.57	1-6
Year 5	89.00	84.50	4.50	0.45	8.55	1-9
Year 8	83.00	80.00	3.00	-1.65	7.65	1-8
% Patients alive[5]						
Year 2	92.70	92.40	0.30	-2.85	3.45	1-4
Year 5	82.90	79.30	3.60	-1.09	8.29	1-8
Year 8	78.00	72.00	6.00	0.82	11.18	1-11

1. Data points were interpolated from the survival curves reported in Powles et al (2002).[269]
2. Clodronate % minus placebo %.
3. Confidence intervals for the differences in % effectiveness were calculated using the formula provided by Armitage & Berry (1995) pp 128-130.[170]
4. The prior expectations for each outcome were assumed to be given by a uniform distribution in the range 1% to the upper limit of the 95% confidence interval (rounded up or down to the nearest whole number). Therefore, in the calculations, each value within the range is assumed to have an equal chance of representing the true difference in effectiveness.
5. Powles et al (2002) did not report disease free survival rates.[269] However, for the illustrative analyses performed in this chapter, the overall survival rates reported were used as if they were the disease free survival rates.

2. Choice of baseline design

It is important to note that the values of the predicted probabilities calculated for different trial designs depend upon the baseline values against which a new design is compared. Thus, such calculations require a 'baseline design' to be chosen. The calculations used to generate the results presented in Table 4.8 and Figures 4.2 to 4.4 are based on the following baseline design:

Table A4.4.2

Attribute	Baseline product evidence	Justification
Endpoint	Patients without metastatic bone disease	The primary endpoint in the Powles et al (2002) trial[269]
Effectiveness	2.90%	The statistically significant difference observed in the Powles et al (2002) study at the end of the treatment period (2 years).[269]
Uncertainty	0.23% to 5.57% [0.92]	The confidence interval was computed using the formula provided by Armitage & Berry (1995) pp 128-130[170]
Duration	2 years	The period (medication period) over which the statistically significant difference in the above measure of effectiveness was observed.
Population	Patients with Stages I to III disease at diagnosis	The Powles et al (2002) trial enrolled patients with primary operable breast cancer regardless of the stage of disease at diagnosis.[269]
Cost	£450,000	This level reflects the approximate cost of treating patients with oral clodronate for a 2 year period as allowed for in the Powles et al (2002) trial dosing regimen.[267]

This is the baseline design presented in column (1) Table 4.8. It is important to note that, despite being based on the Powles et al (2002) trial, the choice of this baseline is purely illustrative.

3. Example calculation of predicted probabilities of product adoption presented in Section 4.7.1, Table 4.8 and Figure 4.2

Consider the comparison between the baseline design and "New product RCT Design 7" shown respectively in columns (1) and (8) of Table 4.8. By substituting the regression coefficients from the full sample model (Table 4.7) and the *differences* in the values of the two RCT design attributes into Equation (8) we derive:

$$\Delta V_{AB} = \begin{bmatrix} (0.2787*1)+(0.0457*37.10)+(-0.6210*-0.91)+ \\ (0.0255*8)+(0.2419*0)+(-5.43e-07*-450000) \end{bmatrix} = 2.9876 \ .$$

Substituting the above utility value into Equation (7) gives the predicted probability of preferring RCT Design 7 (A) to the baseline design (B):

$$\Pr(A \mid \mathrm{J}) = \Phi(\Delta V_{AB}) = \Phi(2.9876) = 0.9986 \text{, and}$$

$$\Pr(B \mid \mathrm{J}) = 1 - \Phi(\Delta V_{AB}) = 1 - 0.9986 = 0.0014 .$$

All the results presented in Table 4.8 and Figure 4.2 were calculated in this way.

4. Example calculation of expected predicted probabilities of product adoption presented in Section 4.7.2, Figure 4.3

For the illustrative calculations used to produce Figure 4.3, it is assumed that trial sponsors would only consider studies of 2, 5 or 8 years duration and that they would only accept a study population which included subjects with all stages of primary operable breast cancer at diagnosis. It is further assumed that the new bisphosphonate treatment would be priced at parity with the existing treatment (equivalent to £450,000 per 100 patients treated) and that a total sample size above 5000 subjects would not be contemplated. These assumptions are made in order to limit the computational effort involved.

As an example calculation, consider the comparison between a design which differs from the above baseline only in terms of the period of follow-up (5 years instead of 2 years). The calculations are illustrated for a hypothetical trial which enrols a total of 1500 subjects (750 per arm). The calculation involves the following steps:

4.1. Calculation of expected effectiveness

For each sample size, n, we first need to calculate the *expected* values for the effectiveness attribute. Utilising an approach previously described by Backhouse (1998) [203] and Detsky (1985,1990) [152;154], the expected effectiveness, $E_n(\Delta X)$ likely to be demonstrated by a trial of sample size n, is given by the following formula:

$$E_n(\Delta X) = \sum_{\Delta X=-\infty}^{\infty} \Pr\left(D_n^{\phi} = \Delta X | \Delta X\right).\Pr(\Delta X).\Delta X$$

where $\Pr\left(D_n^{\phi} = \Delta X | \Delta X\right)$ is the conditional probability that a difference of ΔX will be established in a trial with significance level ϕ if that difference is in fact there (the power of a trial), and where $\Pr(\Delta X)$ is the prior probability of a true difference of ΔX.

For this illustrative calculation, $E_{n=1500}(\Delta X)$ =

(0.0781 * 0.1111 * (85.5%-84.5%) +
(0.1948 * 0.1111 * (86.5%-84.5%) +
(0.3874 * 0.1111 * (87.5%-84.5%) +
(0.6207 * 0.1111 * (88.5%-84.5%) +
(0.8216 * 0.1111 * (89.5%-84.5%) +
(0.9406 * 0.1111 * (90.5%-84.5%) +
(0.9869 * 0.1111 * (91.5%-84.5%) +
(0.9982 * 0.1111 * (92.5%-84.5%) +
(0.9999 * 0.1111 * (93.5%-84.5%) = **4.1953%.**

In the above calculation, the first number in each row is the power of the trial which is calculated using the formula described in Machin et al (1997) p19. [195] A two tailed test with significance level ϕ = 5% was assumed for all calculations. The second

number in each row is the prior probability of the difference in effectiveness shown as the last number in each row. Thus, from Table A4.4.1, the assumed prior expectation of difference in effectiveness is 1% to 9% with each value in that range assumed to have an equal prior probability: 1/9 = 0.1111.

4.2. Calculation of expected uncertainty

To calculate the expected uncertainty, 95% confidence intervals for the expected effectiveness outcome were computed using the formula provided by Armitage & Berry (1995) pp 128-130 referred to in the footnote to Table A4.4.2 above.[170] Based on this formula, the expected 95% confidence interval is: 0.7537% to 7.6369% (4.1953% ± 3.4416%). The expected uncertainty attribute value for this particular trial is therefore given by:

$$E_{1500} = [(7.6369 - 0.7537)/4.1953]/2 = 0.8203.$$

This is the uncertainty value which enters the calculation below.

4.3. Calculation of expected predicted probability

By substituting the regression coefficients from the full sample model and the differences in the values of the two RCT design attributes into Equation (8) we derive:

$$E_{1500}(\Delta V_{AB}) = \begin{bmatrix} (0.2787*0)+(0.0457*1.2953)+(-0.6210*-0.0997)+ \\ (0.0255*3)+(0.2419*0)+(-5.43e-07*-0) \end{bmatrix} = 0.1976$$

Substituting the above utility value into Equation (7) gives the *expected* predicted probability of preferring the RCT design as described above (A) to the baseline design (B):

$$E_{1500}\ \Pr(A \mid \mathrm{J}) = \Phi(E_{1500}(\Delta V_{AB})) = \Phi(0.1976) = 0.5783.$$

This value was used to plot, in Figure 4.3, the expected predicted probability for a trial of 5 years duration with total sample size of 1500 and the percentage of patients without metastatic bone disease as the primary endpoint. Similar calculations were repeated across a large range of sample sizes and designs to produce Figure 4.3.

5. Example calculation of expected net present values presented in Section 4.7.3, Figure 4.4

Backhouse (1998) has shown how the net present value (NPV) for a trial can be calculated and used to determine optimal (NPV maximising) designs. [203] This is the approach adopted here. The example below builds on the example presented above.

5.1 Calculation of cost

In order to estimate the discounted cost of performing the trials PTC_n, the following simple cost function was assumed:

$$PTC_n = \sum_{t=0}^{H} \frac{F_t + V_t.n_t + FU_t.n_t}{(1+r)^t}$$

where F_t denotes the fixed cost of performing the trial, V_t denotes the variable cost per subject enrolled, FU_t denotes the follow-up cost per patient per year, n_t denotes the total trial sample size and t denotes the year in which the costs are incurred. In all calculations, the following assumptions are made:

$F_t = £1{,}000{,}000$, incurred in the first year ($t = 0$)

$V_t = £3{,}000$, incurred in the first year ($t = 0$)

$FU_t = £1{,}000$ incurred for each year of follow-up ($t = 1$ to 5)

$H = 15$ years, the time horizon for the NPV calculations

$r = 0.15$, the discount rate.

Based on the above assumptions, the cost calculation for the 5 – year trial illustrated in section 4 above is shown below in Table A4.4.3.

Table A4.4.3

Year	t	n	F_t	$V_t.n_t$	$FU_t.n_t$	TC_t	PTC_t
1	0	1500	£1,000,000	£4,500,000		£5,500,000	£5,500,000
2	1	1500			£1,500,000	£1,500,000	£1,304,348
3	2	1500			£1,500,000	£1,500,000	£1,134,216
4	3	1500			£1,500,000	£1,500,000	£ 986,274
5	4	1500			£1,500,000	£1,500,000	£ 857,630
6	5	1500			£1,500,000	£1,500,000	£ 745,765
						Total $PTC_{n=1500}$	**£10,528,233**

It should be emphasised that the assumed values used here are purely illustrative. They do not necessarily reflect the actual costs of performing such a trial.

5.2 Calculation of revenue

In order to estimate the discounted revenue associated with performing the trials, the following simple demand function was assumed:

$$E_n \Pr_t(A \mid \mathrm{J})\, M_t$$

where M_t denotes the annual number of newly diagnosed cases of breast cancer assumed to be currently treated with the baseline product. The discounted revenue associated with a trial of given design and sample size, PTR_n, is given by:

$$PTR_n = \sum_{t=0}^{H} \frac{E_n \Pr_t(A \mid \mathrm{J})\, M_t .P_t}{(1+r)^t}$$

where P_t denotes the cost per year per patient treated and all other variables are as described above.

In all calculations, the following variable values were assumed:

$M_t = 15{,}000$, assumed to be constant for the time horizon of this illustrative calculation and represents the number of patients assumed to be currently treated with the baseline product.

P_t = £2,250 per patient per year for a two year course of treatment (treatment cost = £4,500 per patient).

H = 15 years, the time horizon for the NPV calculations

r = 0.15, the discount rate.

Based on the above assumptions, the expected revenue calculations for the 5 – year trial illustrated in section 4 above is shown in Table A4.4.4 below.

5.3 Calculation of expected NPV

Finally, the expected net present value of a trial of given design and sample size, NPV_n, is given by:

$$NPV_n = PTR_n - PTC_n$$

which in this case equals £68,392,368 - £10,528,233 = £57,864,135. This value was used to plot, in Figure 4.4, the expected NPV for a trial of 5 years duration with total sample size of 1500 and the percentage of patients without metastatic bone disease as the primary endpoint. Similar calculations were repeated across a large range of sample sizes and designs to produce Figure 4.4.

Table A4.4.4

Year	t^2	M_t	$E_n \Pr_t(A \mid \mathrm{J})$	$E_n \Pr_t(A \mid \mathrm{J}) M_t$	Number Treated[1]	TR_t	PTR_t
8	7	15000	0.5783	8675	8675	£19,518,098	£ 7,337,576
9	8	15000	0.5783	8675	17350	£39,036,196	£12,761,002
10	9	15000	0.5783	8675	17350	£39,036,196	£11,096,523
11	10	15000	0.5783	8675	17350	£39,036,196	£ 9,649,151
12	11	15000	0.5783	8675	17350	£39,036,196	£ 8,390,566
13	12	15000	0.5783	8675	17350	£39,036,196	£ 7,296,144
14	13	15000	0.5783	8675	17350	£39,036,196	£ 6,344,473
15	14	15000	0.5783	8675	17350	£39,036,196	£ 5,516,933
						Total $PTR_{n=1500}$	**£68,392,368**

1. Each patient is assumed to receive treatment for 2 years at an annual cost of £2,250.
2. Year 8 ($t = 7$) is assumed to be spent acquiring marketing authorisation hence no revenue is received.

CHAPTER 5 : DISCUSSION AND CONCLUSION

Randomised controlled trials (RCTs) play a fundamental role in the development and marketing activities of pharmaceutical companies. They are costly to perform and their design and results are a major determinant of the sales of a product. The contribution of RCTs to the performance of pharmaceutical companies emphasizes the importance of considering the application of techniques by which the value of trials may be assessed. Yet a review of the literature pertaining to economic analysis and clinical trials performed in the course of this research revealed that a private sector investment appraisal approach to RCT design has not hitherto been proposed nor has its potential application previously been explored. Therefore the purpose of this thesis has been to begin to fill that gap by setting out how methods of investment appraisal might be applied to RCT design decision-making and by exploring aspects surrounding the practicalities of application. Each individual chapter presents conclusions and directions for future research so they will not be repeated in detail here. Rather the purpose of this chapter is to briefly summarise the *key* conclusions and contributions of this thesis and to present an agenda of research topics that could be pursued to further develop the application of investment appraisal approaches to trial design.

5.1 SUMMARY OF CONCLUSIONS AND CONTRIBUTIONS

In the field of economic analysis and RCTs, this thesis makes original contributions of both a conceptual and applied nature. The main conclusions and contributions of this thesis are summarised below.

The main contributions based on Chapter 2 are two fold. Firstly, a potential role for private sector investment appraisal techniques in the design of RCTs was identified and proposed for the first time. Secondly, a general model was described setting out the key components of the approach. An illustration of the application of the NPV method of investment appraisal was presented using a hypothetical trial design problem facing a pharmaceutical company. It was shown how profit criteria could be applied to decide whether a particular RCT is worth conducting, to determine an optimal (NPV maximising) design, and to rank RCTs in terms of their expected NPVs. The latter would enable a company to select a portfolio of studies which maximises the return on a given development or trial budget.

One of the strengths of this technique lies in the fact that it requires a health care decision maker focused approach to the planning and design of RCTs. It explicitly recognizes that the nature and extent of product adoption (and ultimately profit) is linked to the strength and relevance of the evidence that trials produce for decision makers. This has not been built in to the methods previously proposed for use by companies to assess the value of research and development projects. Therefore the importance of being able to estimate the demand for a product contingent upon RCT design and expected trial outcomes was highlighted and the potential use of discrete choice analysis for that purpose was proposed for the first time.

The work in Chapter 3 was conducted at a time when the pharmaceutical industry started to express its concerns about the potential impact, on development times and costs, of decision makers' requiring cost-effectiveness evidence generated by RCTs. The work set out to explore how such a requirement might impact the major determinants of RCT costs, namely sample size and study duration. The

applied investigation utilized data from a previously conducted clinical evaluation and published formulae to calculate sample sizes for testing cost-effectiveness hypotheses for a number of hypothetical study designs. The designs were chosen to portray different preferences for cost-effectiveness evidence characterized by various design components of an RCT.

The main contributions based on Chapter 3 are two fold. Firstly, the analyses conducted demonstrate, for the first time, that the impact on RCT costs depends upon the *specific nature* of the cost-effectiveness evidence requirements. In so doing, the work draws attention to one of the practical implications of the findings: such studies cannot be adequately designed without detailed prior information about decision-makers' preferences for evidence defined in terms of RCT design attributes. Secondly, the analyses conducted also demonstrate that circumstances can be such that a requirement to produce evidence of cost-effectiveness based wholly on RCTs can significantly increase product development costs and times above that which would be required to test hypotheses only on clinical endpoints. The direction and magnitude of any impact will always depend upon the specific requirements and circumstances surrounding a product.

Empirical demand analysis is a key component of an investment appraisal approach to RCT design. Specifically, the technique cannot be operationalised without estimating the demand for a product contingent on the design and results of RCTs. Therefore Chapter 4 focused on this critical component of the investment appraisal approach. The purpose of the work presented in Chapter 4 was to explore how DCA could be used as a means of incorporating decision makers' preferences into the design of RCTs. Specifically, a primary goal was to explore the use of DCA as a method

for estimating the probabilities of product adoption contingent upon different RCT designs and expected results in order to operationalise the investment appraisal approach to trial design.

In Chapter 4, a discrete choice model of drug demand was presented in general form. This was followed by an overview of the key elements to be considered in the design of a stated preference survey to estimate the parameters of a discrete choice model. The approach was then illustrated through the design and application of a stated preference survey to elicit clinician preferences for RCT evidence pertaining to the use of bisphosphonates in the management of patients with primary operable breast cancer. The data collected in the survey were used to estimate the parameters of a discrete choice model. It was shown how those results could be used to derive predicted probabilities of product adoption which, in turn, can be used to determine optimal RCT designs within an investment appraisal framework.

The main contributions of the work presented in Chapter 4 are as follows. Firstly, the use of DCA as a method for incorporating decision-maker preferences for evidence into the design of RCTs has not hitherto been proposed. In Chapter 4, the potential use of DCA for that purpose has been identified and described. Secondly, the first application of DCA to RCT design was presented. The empirical case-study provides evidence that the method is practical and theoretically valid in that context. Finally, and in contrast to other health economics applications of DCA, the research has focused on the use of the predicted choice probabilities that can be derived from discrete choice models. Specifically, the research has shown for the first time how these can be used to determine preference maximizing or profit maximizing trial

designs.

5.2 AN AGENDA FOR FURTHER RESEARCH

This thesis has begun to explore how pharmaceutical companies might apply investment appraisal techniques to assist with RCT design decision-making. In this section, some major directions for further research into the application of investment appraisal techniques to RCT design are briefly highlighted.

The application of investment appraisal to RCT design was illustrated using the NPV method of investment appraisal. This was chosen because of its known superiority over rival orthodox methods and because of its widespread use by pharmaceutical companies to appraise project investment decisions. However, future research should consider how the conclusions reached through the NPV method compare with those reached by alternatives such as decision-analytic methods and value of information analysis[277;278] and the real options approach to investment decisions.[279;280]

Further research should also be directed at gaining experience with the application of investment appraisal methods to various investment decisions involving the design of RCTs. Specifically, it would be worthwhile focusing on applying investment appraisal techniques to determine an optimal programme of trials to meet the evidence needs of both the regulatory agencies and pricing and reimbursement authorities. Any application will require a company to be able to estimate both the costs and the revenues contingent upon different RCT designs and results. It is the estimation of the product adoption and revenues that will pose the greatest challenge to successful application of the approach.

Empirical work presented in this thesis has shown that discrete choice analysis is a potentially promising technique for incorporating decision-makers preferences for evidence into the design of trials generally and for operationalising an investment appraisal approach to RCT design in particular. Further research into the application of DCA will need to address how the technique can be successfully applied to decision-makers other than clinicians and to more complex treatment choice situations than those illustrated here. Furthermore, future research will need to address how product diffusion can reliably be modeled based on the findings of stated preference surveys[204] or other techniques.[281]

REFERENCES

1. Friedman LM, Furberg CD, DeMets DL. Fundamentals of clinical trials, Second ed. Littleton, MA: PSG Publishing Company; 1985.

2. Bulpitt CJ. Randomised controlled clinical trials, First ed. The Hague: Martinus Nijhoff; 1983.

3. Meinert CL. Clinical trials: design, conduct and analysis, First ed. Oxford University Press; 1986.

4. Spilker B. Guide to clinical trials. New York: Raven Press; 1991.

5. Pocock SJ. Clinical trials: a practical approach. John Wiley & Sons; 1991.

6. US Food and Drug Administration (FDA): http://www.fda.gov/.

7. European Medicines Evaluation Agency (EMEA): http://www.emea.eu.int/.

8. Clark T. A guide to the international pharmaceutical industry, 2005 ed. New Orleans: VOI Consulting Inc; 2005.

9. Caspari M, DeFazio L, Geanotes A, Gerlach E, Hampton R, Harper C *et al.* Pharmaceutical pricing and reimbursement, 2003 ed. Cambridge, UK.: PPR Communications Ltd.; 2003.

10. Drummond MF, Dubois D, Garattini L, Horisberger B, Jonsson B, Kristiansen IS *et al.* Current trends in the use of pharmacoeconomics and outcomes research in Europe. Value in Health 1999;2:323-32.

11. Drummond, M. F. The use of health economic information by reimbursement authorities. Rheumatology 42(Suppl 3), 60-63. 2003.

12. Taylor, R. S., Drummond, M. F., Salkeld, G., and Sullivan, S. D. Inclusion of cost effectiveness in licensing requirements of new drugs: the fourth hurdle. British Medical Journal 23(329), 972-975. 2004.

13. Tarn, T. Y. H. and Dix Smith, M. Pharmacoeconomic guidelines around the world. ISPOR Connections 10(4), 5-15. 2004.

14. National Institute for Clinical Excellence. Guide to the methods of technology appraisal, April 2004 ed. London: NICE; 2004.

15. National Institute for Health and Clinical Excellence (NICE). NICE . 2005.

16. Williams H. New treatments for atopic dermatitis: good news, but when and how to use them is a muddle. British Medical Journal 2002;324:1533-34.

17. Williams H. Another vehicle-controlled study of 1% pimecrolimus in atopic dermatitis: how does it help clinicians and patients and is it ethically sound? Archives of Dermatology 2002;138:1602-03.

18. Grabowski H, Vernon J. Returns to R&D on new drug introductions in the 1980s. Journal of Health Economics 1994;13:383-406.

19. Grabowski H. The effect of pharmacoeconomics on company research and development decisions. PharmacoEconomics 1997;11:389-97.

20. El Feki, S. Prescription for change: a survey of pharmaceuticals. The Economist (June 18-24), 3-17. 18-6-2005.

21. Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes, 2nd ed. Oxford: Oxford University Press; 1997.

22. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. Oxford: Oxford University Press; 1996.

23. Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomized control trials. Medical Care 1992;30:231-43.

24. Culyer AJ, Wiseman J, Walker A. An annotated bibliography of health economics. London: Martin Robinson; 1977.

25. Drummond MF, Ludbrook A, Lowson KV, Steele A. Studies in economic appraisal in health care, vol. 2. Oxford: Oxford University Press; 1986.

26. Elixhauser A, Luce BR, Taylor WR, Reblando J. Health care cost-benefit and cost-effectiveness analysis from 1979 to 1990:a bibliography. Medical Care 1993;31:JS1-JS150.

27. Elixhauser A. Health care cost-benefit and cost-effectiveness analysis: an update. Medical Care 1998;36:MS1-MS147.

28. Ganiats TG, Wong AF. Evaluation of cost - effectiveness research: a survey of recent publications. Family Medicine 1991;23:457-62.

29. Udvarhelyi IS, Colditz GA, Rai A, Epstein AM. Cost-effectiveness and

cost-benefit analyses in the medical literature: are the methods being used correctly? Annals of Internal Medicine 1992;116:238-44.

30. Backhouse ME, Backhouse RJ, Edey SA. Economic evaluation bibliography. Health Economics 1992;1:1-236.

31. O'Brien BJ, Drummond MF, Labelle RJ, Willan A. 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.

32. Detsky AS, Naglie G, Krahn MD, Naimark D, Redelmeier DA. Primer on medical decision analysis: Part 1 - getting started. Medical Decision Making 1997;17:123-25.

33. Detsky AS, Naglie G, Krahn MD, Redelmeier DA, Naimark D. Primer on medical decision analysis: Part 2 - building a tree. Medical Decision Making 1997;17:126-35.

34. Naglie G, Krahn MD, Naimark D, Redelmeier DA, Detsky AS. Primer on medical decision analysis: Part 3 - estimating probabilities and utilities. Medical Decision Making 1997;17:136-41.

35. Krahn MD, Naglie G, Naimark D, Redelmeier DA, Detsky AS. Primer on medical decision analysis: Part 4 - analyzing the model and interpreting the results. Medical Decision Making 1997;17:142-51.

36. Drummond MF. Economic analysis alongside controlled trials. London: Department of Health; 1994.

37. Drummond MF, Stoddart GL. Economic analysis and clinical trials. Controlled Clinical Trials 1984;5:115-28.

38. Drummond MF. Experimental versus observational data in the economic evaluation of pharmaceuticals. Medical Decision Making 1998;18:S12-S18.

39. Buxton M, Drummond MF, VanHout BA, Prince RL, Sheldon TA, Szucs T *et al.* Modelling in economic evaluation: an unavoidable fact of life. Health Economics 1997;6:217-27.

40. Holly N. A call for randomized controlled cost-effectiveness analysis of percutaneous transluminal coronary angioplasty. International Journal of Technology Assessment in Health Care 1988;4:497-510.

41. Schoenbaum SC. Cost/effectiveness considerations in clinical trials. Triangle 1980;19:103-06.

42. Copley-Merriman C, Egbuonu-Davis L, Kotsanos JG, Conforti P, Franson T, Gordon G. Clinical economics: a method for prospective health resource data collection. PharmacoEconomics 1992;1:370-76.

43. Mauskopf J, Schulman KA, Bell L, Glick H. A strategy for collecting pharmacoeconomic data during phase II/III clinical trials. PharmacoEconomics 1996;9:264-77.

44. Backhouse, M. E., Gnanasakthy, A., Schulman, K. A., Akehurst, R. A., and Glick, H. The development of standard economic datasets for use in the economic evaluation of medicines. Drug Information Journal 34, 1273-1291. 2000.

45. Briggs A. A Bayesian approach to stochastic cost-effectiveness analysis. Health Economics 1999;8:257-61.

46. Briggs A, Mooney CZ, Wonderling D. Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using Monte Carlo simulation. Statistics in Medicine 1999;18:3245-62.

47. Briggs A, Gray A. Handling uncertainty in economic evaluations of healthcare interventions. British Medical Journal 1999;319:635-38.

48. Gardiner J, Holmes-Rovner M, Goddeeris J, Rovner D, Kupersmith J. Covariate-adjusted cost-effectiveness ratios. Journal of Statistical Planning and Inference 1999;75:291-304.

49. Laska EM, Meisner M, Siegel CS. Power and sample size in cost-effectiveness analysis. Medical Decision Making 1999;19:339-43.

50. Heitjan DF, Moskowitz AJ, Whang W. Bayesian estimation of cost-effectiveness ratios from clinical trials. Health Economics 1999;8:191-201.

51. Obenchain, R. L. Resampling and multiplicity in cost-effectiveness inference.

Journal of Biopharmaceutical Statistics 9, 563-582. 1999.

52. Rittenhouse BE, Dulisse B, Stinnett AA. At what price significance? The effect of price estimates on statistical inference in economic evaluation. Health Economics 1999;8:213-19.

53. Torgerson D, Campbell MK. Cost effectiveness calculations and sample size. British Medical Journal 2000;321:697.

54. Gardiner JC, Huebner M, Jetton J, Bradley CJ. Power and sample size assessments for tests of hypotheses on cost-effectiveness ratios. Health Economics 2000;9:227-34.

55. Raikou M, Briggs A, Gray A, McGuire A. Centre-specific or average unit costs in multi-centre studies? Some theory and simulation. Health Economics 2000;9:191-98.

56. O'Hagan A, Stevens JW, Montmartin J. Inference for the cost-effectiveness acceptability curve and cost-effectiveness ratio. PharmacoEconomics 2000;17:339-49.

57. Lothgren M, Zethraeus N. Definition, interpretation and calculation of cost-effectiveness acceptability curves. Health Economics 2000;9:623-30.

58. Al, M. J. and VanHout, B. A. A Bayesian approach to economic analyses of clinical trials: the case of stenting versus balloon angioplasty. Health Economics 9, 599-609. 2000.

59. Cook, J. R. and Heyse, J. F. Use of an angular transformation for ratio estimation in cost-effectiveness analysis. Statistics in Medicine 19, 2989-3003. 2000.

60. Heitjan, D. F. Fieller's method and net health benefits. Health Economics 9, 327-335. 2000.

61. Briggs A, O'Brien BJ. The death of cost-minimisation analysis? Health Economics 2000;4.

62. McIntosh MW, Ramsey SD, Berry K, Urban N. Parameter solicitation for planning cost effectiveness studies with dichotomous outcomes. Health Economics 2001;10:53-66.

63. Tambour M, Zethraeus N, Johannesson M. A note on confidence intervals in cost-effectiveness analysis. International Journal of Technology Assessment in Health Care 1998;14:467-71.

64. Tambour M, Zethraeus N. Bootstrap confidence intervals for cost-effectiveness ratios: some simulation results. Health Economics 1998;7:143-47.

65. Al MJ, VanHout BA, Michel BC, Rutten FFH. Sample size calculation in economic evaluations. Health Economics 1998;7:327-35.

66. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Medical Decision Making 1998;18:S68-S80.

67. Willke RJ, Glick HA, Polsky D, Schulman KA. Estimating country-specific cost-effectiveness from multinational clinical trials. Health Economics 1998;7:481-93.

68. Briggs A, Gray A. Power and sample size calculations for stochastic cost - effectiveness analysis. Medical Decision Making 1998;18:S81-S92.

69. Briggs A, Gray A. The distribution of health care costs and their statistical analysis for economic evaluation. Journal of Health Services Research and Policy 1998;3:233-45.

70. Briggs A, Tambour M. The design and analysis of stochastic cost-effectiveness studies for the evaluation of health care interventions. Stockholm School of Economics Working Paper Series in Economics and Finance 1998;234:1-22.

71. Heitjan DF, Moskowitz AJ, Whang W. Problems with interval estimates of the incremental cost-effectiveness ratio. Medical Decision Making 1999;19:9-15.

72. Willan AR, O'Brien BJ. Sample size and power issues in estimating incremental cost-effectiveness ratios from clinical trials data. Health Economics 1999;8:203-11.

73. Lord J, Asante MA. Estimating uncertainty ranges for costs by the bootstrap procedure combined with probabilistic sensitivity analysis. Health Economics 1999;8:323-33.

74. Severens JL, De Boo TM, Konst EM. Uncertainty of incremental cost-effectiveness ratios. International Journal of Technology Assessment in

Health Care 1999;15:608-14.

75. Whang W, Sisk JE, Heitjan DF, Moskowitz AJ. Probabilistic sensitivity analysis. International Journal of Technology Assessment in Health Care 1999;15:563-72.

76. Etzioni R, Feuer E, Sullivan S, Lin D, Chengcheng H, Randall A. On the use of survival analysis techniques to estimate medical care costs. Journal of Health Economics 1999;18:365-80.

77. Liu G, Zhao Z. Stochastic cost-effectiveness analysis. Value in Health 1999;2:420-28.

78. Bala M, Mauskopf J. The estimation and use of confidence intervals in economic analysis. Drug Information Journal 1999;33:841-48.

79. Cook J, Drummond MF, Glick H, Heyse JF. Analyzing data from multinational clinical trials. HESG 1997.

80. Laska EM, Meisner M, Siegel C. Statistical inference for cost - effectiveness ratios. Health Economics 1997;6:229-42.

81. Lin D, Feuer E, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data:when the cost histories are not recorded. Biometrics 1997;53:419-34.

82. Polsky D, Glick H, Willke R, Schulman KA. Confidence intervals for cost-effectiveness ratios: a comparison of four methods. Health Economics 1997;6:243-52.

83. Stinnett AA, Paltiel AD. Estimating CE ratios under second-order uncertainty: the mean ratio versus the ratio of means. Medical Decision Making 1997;17:483-89.

84. Briggs A, Fenn P. Trying to do better than average: a commentary on 'statistical inference for cost-effectiveness ratios'. Health Economics 1997;6:491-95.

85. Laska EM, Meisner M, Siegel C. The usefulness of average cost-effective ratios. Health Economics 1997;6:497-504.

86. Fenn P, McGuire A, Phillips V, Backhouse ME, Jones DA. The analysis of censored treatment cost data in economic evaluation. Medical Care 1995;33:851-63.

87. Gardiner J, Hogan A, Holmes-Rovner M, Rovner D, Griffith L, Kupersmith J. Confidence intervals for cost-effectiveness ratios. Medical Decision Making 1995;15:254-63.

88. Sacristan JA, Day SJ, Navarro O, Ramos J, Hernandez JM. Use of confidence intervals and sample size calculations in health economic studies. The Annals of Pharmacotherapy 1995;29:719-25.

89. Wakker P, Klaasen MP. Confidence intervals for cost/effectiveness ratios. Health Economics 1995;4:373-81.

90. Torgerson D, Ryan M, Ratcliffe J. Economics in sample size determination for clinical trials. Quarterly Journal of Medicine 1995;88:517-21.

91. Briggs A, Wonderling D. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Economics 1997;6:327-40.

92. Chaudhary MA, Stearns SC. Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. Statistics in Medicine 1996;15:1447-58.

93. Coyle D. Statistical analysis in pharmacoeconomic studies:a review of current issues and standards. PharmacoEconomics 1996;9:506-16.

94. Fenn P, McGuire A, Backhouse ME, Jones DA. Modelling programme costs in economic evaluation. Journal of Health Economics 1996;15:115-25.

95. Lin D, Etzioni R. Estimating medical costs from incomplete follow-up data:when the cost histories are recorded. Department of Biostatistics Technical Report, University of Washington 1996;138.

96. Sacristan JA. Use of confidence intervals in health economic studies. PharmacoEconomics 1996;10:432-33.

97. Siegel C, Laska EM, Meisner M. Statistical methods for cost - effectiveness analysis. Controlled Clinical trials 1996;17:387-406.

98. Willan A, O'Brien BJ. Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. Health Economics 1996;5:297-305.

99. Gray A, Marshall M, Lockwood A, Morris J. Problems in conducting economic evaluations alongside clinical trials: lessons from a study of case

management for people with mental disorders. British Journal of Psychiatry 1997;170:47-52.

100. Drummond MF, O'Brien BJ. Clinical importance, statistical significance and the assessment of economic and quality-of-life outcomes. Health Economics 1993;2:205-12.

101. O'Brien BJ, Drummond MF. Statistical versus quantitative significance in the socioeconomic evaluation of medicines. PharmacoEconomics 1994;5:389-98.

102. VanHout BA, Al MJ, Gordon JS, Rutten FFH. Costs, effects and C/E - ratios alongside a clinical trial. Health Economics 1994;3:309-19.

103. Rutten-van Molken MPMH, VanDoorslaer EKA, VanVliet RCJA. Statistical analysis of cost outcomes in a randomized controlled clinical trial. Health Economics 1994;3:333-45.

104. Roland M, Torgerson D. Understanding controlled trials: what outcomes should be measured? British Medical Journal 1998;317:1975.

105. Schulman KA, Burke J, Drummond MF, Davies L, Carlsson P, Gruger J *et al.* Resource costing for multinational neurologic clinical trials: methods and results. Health Economics 1998;7:629-38.

106. Baker, A. M., Arnold, R. J., and Kaniecki, D. J. Economic analysis in clinical trials: practical considerations. Drug Information Journal 33, 1053-1060. 2000.

107. Karnon J, Qizilbash N. Economic evaluation alongside n - of - 1 trials: getting closer to the margin. Health Economics 2001;10:79-82.

108. Dowie J. Clinical trials and economic evaluations? No, there are only evaluations. Health Economics 1997;6:87-89.

109. O'Brien BJ. Economic evaluation of pharmaceuticals: Frankenstein's monster or vampire of trials? Medical Care 1996;34:DS99-DS108.

110. Freemantle N, Drummond MF. Should clinical trials with concurrent economic analyses be blinded. JAMA 1997;277:63-64.

111. Rittenhouse BE. Exorcising protocol - induced spirits: making the clinical trial relevant for economics. Medical Decision Making 1997;17:331-39.

112. Haycox A, Drummond MF, Walley T. Pharmacoeconomics: integrating economic evaluation into clinical trials. British Journal of Clinical Pharmacology 1997;43:559-62.

113. Coyle D, Davies L, Drummond MF. Trials and tribulations: emerging issues in designing economic evaluations alongside clinical trials. International Journal of Technology Assessment in Health Care 1998;14:135-44.

114. Hillman BJ, Kahan JP, Neu CR, Hammons GT. Clinical trials to evaluate cost - effectiveness. Investigative Radiology 1989;24:167-71.

115. Drummond MF, Davies L. Economic analysis alongside clinical trials: revisiting the methodological issues. International Journal of Technology Assessment in Health Care 1991;7:561-73.

116. Bonsel GJ, Rutten FFH, Uyl-de Groot CA. Economic evaluation alongside cancer trials: methodological and practical aspects. European Journal of

Cancer 1993;29A:S10-S14.

117. Bennett CL, Armitage JL, Buchner D, Gulati S. Economic analysis in phase III cancer trials. Cancer Investigation 1994;12:336-42.

118. Coyle, D., Davies, L., and Drummond, M. F. Trials and tribulations: conducting economic evaluations alongside clinical trials. 1994.

119. Hillis A. Cost containment: death knell or new opportunity for randomized trials? Archives of Ophthalmology 1994;112:174-75.

120. Baker AM, Goldberg ARJ, Kaniecki DJ. Considerations in measuring resource use in clinical trials. Drug Information Journal 1995;29:1421-28.

121. Berger ML. Design of prospective cost-effectiveness clinical trials: the critical role of the comparator group. Drug Information Journal 1995;29:1415-20.

122. Drummond MF. Economic analysis alongside clinical trials: problems and potential. Journal of Rheumatology 1995;22:1403-07.

123. Drummond MF, O'Brien BJ. Economic analysis alongside clinical trials: practical considerations. The Journal of Rheumatology 1995;22:1418-19.

124. Glick H. Strategies for economic assessment during the development of new drugs. Drug Information Journal 1995;29:1391-403.

125. Powe NR, Griffiths RI. The clinical-economic trial: promise, problems and challenges. Controlled Clinical trials 1995;16:377-94.

126. Rittenhouse BE. The relevance of searching for effects under a clinical trial

lamppost: a key issue. Medical Decision Making 1995;15:348-57.

127. Schulman KA, Yabroff KR, Glick H. A health services approach for the evaluation of innovative pharmaceutical and biothechnology products. Drug Information Journal 1995;29:1405-14.

128. Donaldson C, Hundley V, McIntosh E. Using economics alongside clinical trials: why we cannot choose the evaluation technique in advance. Health Economics 1996;5:267-69.

129. Ellwein LB, Drummond MF. Economic analysis alongside clinical trials: bias in the assessment of economic outcomes. International Journal of Technology Assessment in Health Care 1996;12:691-97.

130. Guess HA, Rudnick SA. Use of cost-effectiveness analysis in planning cancer chemoprophylaxis trials. Controlled Clinical trials 1983;4:89-100.

131. Levy RI, Sondik EJ. Large scale clinical trials: are they worth the cost? Annals of the New York Academy of Sciences 1982;411-22.

132. Levy RI. Clinical trials: are the benefits worth the costs? Clinical Research 1982;116-22.

133. Levy RI, Sondik EJ. Decision-making in planning large-scale comparative studies. Annals of the New York Academy of Sciences 1978;304:441-57.

134. Schechtman KB, Gordon ME. A comprehensive algorithm for determining whether a run-in strategy will be a cost-effective design modification in a randomized clinical trial. Statistics in Medicine 1993;12:111-28.

135. McKinlay SM. Cost-efficient designs of cluster unit trials. Prev Med 1994;23:606-11.

136. Allison DB. When is it worth measuring a covariate in a randomized clinical trial? Journal of Consulting and Clinical Psychology 1995;63:339-43.

137. Silverblatt FJ. Clinical trials and statistical rigor:are the benefits necessarily worth the cost?[editorial]. Journal of Infectious Diseases 1987;155:168-69.

138. Eliasziw M, Donner A. A cost - function approach to the design of reliability studies. Statistics in Medicine 1987;6:647-55.

139. Bloch DA. Sample size requirements and the cost of a randomized clinical trial with repeated measurements. Stat Med 1986;5:663-67.

140. Thornquist MD, Urban N, Tseng A, Edelstein C, Lund B, Omenn GS. Research cost analyses to aid in decision making in the conduct of a large prevention trial, CARET. Carotene and Retinol Efficacy Trial. Control Clin Trials 1993;14:325-39.

141. Ciampi A, Till JE. Null results in clinical trials: the need for a decision-theory approach. Br J Cancer 1980;41:618-29.

142. Baker SG, Heidenberger K. Choosing sample sizes to maximize expected health benefits subject to a constraint on total trial costs. Medical Decision Making 1989;9:14-25.

143. Mosteller F, Weinstein MC. Toward evaluating the cost-effectiveness of medical and social experiments, In: Hausman JA, Wise SA, editors. Social

experimentation. Chicago: University of Chicago Press; 1985. p. 221-49.

144. Weinstein MC. Cost-effective priorities for cancer prevention. Science 1983;221:17-23.

145. Mendoza G, Iglewicz B. An extension of Colton's model for comparing two medical treatments. Journal of the American Statistical Association 1978;73:646-49.

146. Begg CB, Mehta CR. Sequential analysis of comparative clinical trials. Biometrika 1979;66:97-103.

147. Mehta CR. Sequential comparison of two exponential distributions with censored survival data. Biometrika 1981;68:669-75.

148. Hamburg M. Statistical analysis for decision making. New York: Harcourt - Brace; 1977.

149. Thompson MS. Decision-analytic determination of study size: the case of electronic fetal monitoring. Medical Decision Making 1981;1:165-79.

150. Thompson MS, Cohen AB. Decision analysis of electronic fetal monitoring, In: Wortman P, editor. Methods for evaluating health services. Beverly Hills: Sage; 1981. p. 109-40.

151. Thompson MS. Evaluation for decision in social programmes. Westmead: Saxon House; 1975.

152. Detsky AS. Using economic analysis to determine the resource consequences

of choices made in planning clinical trials. Journal of Chronic Diseases 1985;38:753-65.

153. Detsky AS. Are clinical trials a cost-effective investment? JAMA 1989;262:1795-800.

154. Detsky AS. Using cost-effectiveness analysis to improve the efficiency of allocating funds to clinical trials. Statistics in Medicine 1990;9:173-84.

155. Claxton K, Posnett J. An economic approach to clinical trial design and research priority setting. Health Economics 1996;5:513-24.

156. Claxton K. Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. Health Economics 1999;8:269-74.

157. Claxton K. The irrelevance of inference: a decision making approach to the stochastic evaluation of health care technologies. Journal of Health Economics 1999;18:341-64.

158. Claxton, K., Neumann, P. J., Araki, S., and Weinstein, M. C. Bayesian value-of-information analysis. An application to a policy model of Alzheimer's disease. International Journal of Technology Assessment in Health Care 17(1), 38-55. 2001.

159. Claxton K, Thompson KM. A dynamic programming approach to the efficient design of clinical trials. Journal of Health Economics 2001;20:797-822.

160. Claxton, K., Ginnelly, L., Sculpher, M., Phillips, Z., and Palmer, S. A pilot study on the use of decision theory and value of information analysis as part of

the NHS Health Technology Assessment programme. Health Technology Assessment 8(31), 1-103. 2004.

161. Ades, A. E., Lu, G., and Claxton, K. Expected value of sample information calculations in medical decision modelling. Medical Decision Making 24(2), 207-227. 2004.

162. Claxton, K., Eggington, S., Ginnelly, L., Griffin, S., McCabe, C., Philips, Z., Tappenden, P., and Wailoo, A. A pilot study of value of information analysis to support research recommendations. CHE Research Paper 4. 2005. The University of York, Centre for Health Economics.

163. Bergman S, Gittins J. Statistical method for pharmaceutical research planning. New York: Dekker; 1985. p. 167-246.

164. Senn S. Some statistical issues in project prioritization in the pharmaceutical industry. Statistics in Medicine 1996;15:2669-702.

165. Senn S. Further statistical issues in project prioritisation in the pharmaceutical industry. Drug Information Journal 1998;32:253-59.

166. Hawkins CJ, Pearce DW. Capital investment appraisal. London: Macmillan; 1979.

167. Brealey RA, Myers SC. Principles of corporate finance, Fifth ed. New York: McGraw-Hill; 1996.

168. Bierman H, Smidt S. The capital budgeting decision: economic analysis of investment projects, Eighth ed. New York: Macmillan; 1993.

169. Machin DM, Campbell MJ. Statistical tables for the design of clinical trials. Oxford: Blackwell Scientific Publications; 1987.

170. Armitage P, Berry G. Statistical methods in medical research, Third ed. Oxford: Blackwell Science; 1995.

171. Myers SC, Shyam-Sunder L. Measuring pharmaceutical industry risk and the cost of capital, In: Helms RB, editor. Competitive strategies in the pharmaceutical industry. Washington DC: American Enterprise Institute; 1996. p. 208-37.

172. DataEdge LLC. *PICAS database*. DataEdge Information (UK), 1998.

173. Koutsoyiannis A. Modern microeconomics, 1st ed. London: Macmillan; 1978.

174. Dent THS, Hawke S. Too soon to market. British Medical Journal 1997;315:1248-49.

175. Greenhalgh T. More convincing evidence of efficacy needs to be cited. British Medical Journal 1997;315:1623.

176. Stein K, Milne R, Best L. BMJ should require advertisements to detail actual state of evidence. British Medical Journal 1997;315:1623-24.

177. Hooper P, Whitehead AM. Drug company's reply. British Medical Journal 1997;315:1624.

178. Wilkinson D, et al. Review of drug in Drug and Therapeutics Bulletin is uninformed. British Medical Journal 1997;315:1625.

179. Ryan M. Using consumer preferences in health care decision making: the application of conjoint analysis. London: Office of Health Economics; 1996.

180. Cattin P, Wittink DR. Commercial use of conjoint analysis: a survey. Journal of Marketing 1982;46:44-53.

181. Mahajan V, Muller E, Bass FM. New product diffusion models in marketing: a review and directions for research. Journal of Marketing 1990;54:1-26.

182. Mahajan V, Peterson RA. Models for innovation diffusion, vol. 48. Sage; 1985.

183. Sultan F, Farley JU, Lehmann DR. A meta-analysis of applications of diffusion models. Journal of Marketing Research 1990;27:70-77.

184. Battista RN. Innovation and diffusion of health-related technologies. International Journal of Technology Assessment in Health Care 1989;5:227-48.

185. Greer AL. The state of the art versus the state of the science: the diffusion of new medical technologies into practice. International Journal of Technology Assessment in Health Care 1988;4:5-26.

186. Sloan FA, Valvona J, Perrin JM, Adamache KW. Diffusion of surgical technology: an exploratory study. Journal of Health Economics 1986;5:31-61.

187. Lilien GL, Rao AG, Kalish S. Bayesian estimation and control of detailing effort in a repeat purchase diffusion environment. Management Science 1981;27:493-506.

188. Becker DM. Decision to adopt new medical technology: a case study of thrombolytic therapy. Social Science and Medicine 1985;21:291-98.

189. Spiegelhalter DJ, Freedman LS, Parmar MKB. Bayesian approaches to randomised trials. Journal of the Royal Statistical Society 1994;157:357-416.

190. ICH. Statistical principles for clinical trials. Harmonised Tripartite Guideline 1997;E9.

191. National Institute for Clinical Excellence. Technical guidance for manufacturers and sponsors on making a submission to a technology appraisal. March 2001. National Institute for Clinical Excellence.

192. O'Sullivan, A. K., Thompson, D., and Drummond, M. F. Collection of health-economic data alongside clinical trials: is there a future for piggyback evaluations? Value in Health 8(1), 67-79. 2005.

193. Cook, J., Drummond, M. F., and Heyse, J. F. Economic endpoints in clinical trials. Statistical Methods in Medical Research 13(2), 157-176. 2004.

194. Ramsey, S., Willke, R., Briggs, A., Brown, R., Buxton, M., Chawla, A., Cook, J., Glick, H., Liljas, B., Petitti, D., and Reed, S. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR task force report. Value in Health 8(5), 521-533. 2005.

195. Machin DM, Campbell MJ, Fayers PM, Pinol APY. Sample size tables for clinical studies, Second ed. Oxford: Blackwell Science Ltd; 1997.

196. Rosner B. Fundamentals of biostatistics, 5th ed. Pacific Grove, CA: Duxbury

Press; 2000.

197. Briggs, A. and Tambour, M. The design and analysis of stochastic cost-effectiveness studies for the evaluation of health care interventions. Drug Information Journal 35, 1455-1468. 2001.

198. Briggs, A., O'Brien, B. J., and Blackhouse, G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. Annual Review of Public Health 23, 377-401. 2002.

199. Zethraeus, N., Johannesson, M., Jonsson, B., and Lothgren, M. Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies. PharmacoEconomics 21(1), 39-48. 2003.

200. NHS Management Executive. The New NHS Reference Costs. Leeds: Department of Health; 1999.

201. Smith DH, Gravelle H. The practice of discounting in economic evaluations of healthcare interventions. International Journal of Technology Assessment in Health Care 2001;17:236-43.

202. Backhouse ME, Mauskopf J. Formal requirements for economic analysis of medicines: the potential implications of the UK NICE requirements for global product development, vol. April 3, Spectrum ed. Boston, MA: Decision Resources; 2000. p. 6.1-6.14.

203. Backhouse ME. An investment appraisal approach to clinical trial design. Health Economics 1998;7:605-19.

204. Ben-Akiva M, Lerman SR. Discrete choice analysis. London: The MIT Press; 1986.

205. Hair JF, Anderson RE, Tatham RL, Black WC. Conjoint analysis, Multivariate data analysis, Fifth ed. New Jersey: Prentice Hall; 1998. p. 387-441.

206. Louviere JJ, Hensher DA, Swait JD. Stated choice methods: analysis and application. Cambridge: Cambridge University Press; 2000.

207. Wittink DR, Cattin P. Commercial use of conjoint analysis: an update. Journal of Marketing 1989;53:91-96.

208. Wittink DR, Vriens M, Burhenne W. Commercial use of conjoint analysis in Europe: results and critical reflections. International Journal of Research in Marketing 1994;11:41-52.

209. Bingham MF, Johnson FR, Miller D. Modeling choice behaviour for new pharmaceutical products. Value in Health 2001;4:32-44.

210. Bryan S, Buxton M, Sheldon R, Grant A. Magnetic resonance imaging for the investigation of knee injuries: an investigation of preferences. Health Economics 1998;7:595-603.

211. Bryan, S., Gold, L, Sheldon, R., and Buxton, M. Preference measurement using conjoint methods: an empirical investigation of reliability. Health Economics 9, 385-395. 2000.

212. Sculpher M, Bryan S, Fry P, de Winter P, Payne H, Emberton M. Patients' preferences for the management of non-metastatic prostate cancer: discrete

choice experiment. British Medical Journal 2004;328:382.

213. Bryan S, Parry D. Structural reliability of conjoint measurement in health care: an empirical investigation. Applied Economics 2001;1:561-67.

214. Bunch WH, Chapman RG. Patient preferences in surgery for scoliosis. J Bone Joint Surg [Am] 1985;67:794-99.

215. Carroll NV, Gagnon JP. Consumer demand for patient-oriented pharmacy services. Am J Public Health 1984;74:609-11.

216. Chakraborty G, Gaeth G, Cunningham M. Understanding consumers preferences for dental service. Journal of Health Care Marketing 1993;21:48-58.

217. Chakraborty G, Ettenson R, Gaeth G. How consumers choose health insurance. Journal of Health Care Marketing 1994;14:21-33.

218. Farrar S, Ryan M, Ross D, Ludbrook A. Using discrete choice modelling in priority setting: an application to clinical service developments. Social Science and Medicine 2000;50:63-75.

219. Garb HN. A conjoint measurement analysis of clinical predictions. J Clin Psychol 1983;39:295-301.

220. Graf MA, Tanner DD, Swinyard WR. Optimizing the delivery of patient and physician satisfaction: a conjoint analysis approach. Health Care Manage Rev 1993;18:34-43.

221. Hakim Z, Pathak D. Modelling the EuroQol data: a comparison of

discrete choice conjoint and conditional preference modelling. Health Economics 1999;8:103-16.

222. Harrison DD, Cooke CW. An elucidation of factors influencing physicians' willingness to perform elective female sterilization. Obstet Gynecol 1988;72:565-70.

223. Harwood RH, Rogers A, Dickinson E, Ebrahim S. Measuring handicap: the London handicap scale, a new outcome measure for chronic disease. Quality in Health Care 1994;3:11-16.

224. Hershey JC, Kunreuther H, Schwartz JS, Williams SV. Health insurance under competition: would people choose what is expected? Inquiry 1984;21:349-60.

225. Holtgrave DR, Weber EU. Dimensions of risk perception for financial and health risks. Risk Anal 1993;13:553-58.

226. Lee A, Lum ME, Beehan SJ, Hillman KM. Interhospital transfers: decision-making in critical care areas. Crit Care Med 1996;24:618-22.

227. Maas A, Stalpers L. Assessing utilities by means of conjoint measurement: an application in medical decision analysis. Medical Decision Making 1992;12:288-97.

228. Verhoef CG, Maas A, Stalpers LJA, Verbeek ALM, Wobbes T, van Daal WAJ. The feasibility of additive conjoint measurement in measuring utilities in breast cancer patients. Health Policy 1991;17:39-50.

229. Magat WA, Viscusi WK, Huber J. Paired comparison and contingent

valuation approaches to morbidity risk valuation. Journal of Environmental Economics and Management 1988;15:395-411.

230. Malhotra NK, Jain AK. A conjoint analysis approach to health care marketing and planning. Journal of Health Care Marketing 1982;2:35-44.

231. McClain JO, Rao VR. Trade-offs and conflicts in evaluation of health system alternatives: methodology for analysis. Health Services Research 1974;9:35-52.

232. Nickerson CA, McClelland GH, Petersen DM. Measuring contraceptive values: an alternative approach. J Behav Med 1991;14:241-66.

233. Parker BR, Srinivasan V. A consumer preference approach to the planning of rural primary health-care facilities. Operations Research 1976;24:991-1025.

234. Propper C. Contingent valuation of time spent on NHS waiting lists. The Economic Journal 1991;100:193-99.

235. Propper C. The disutility of time spent on the United Kingdom's National Health Service waiting lists. The Journal of Human Resources 1995;30:677-700.

236. Ratcliffe J, Buxton M. Patient's preferences regarding the process and outcomes of life-saving technology: an application of conjoint analysis to liver transplantation. International Journal of Technology Assessment in Health Care 1999;15:340-51.

237. Ratcliffe J. The use of conjoint analysis to elicit willingness-to-pay values.

Proceed with caution? International Journal of Technology Assessment in Health Care 2000;16:270-90.

238. Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. Health Economics 2000;9:137-48.

239. Ratcliffe J, Van Haselen R, Buxton M, Hardy K, Colehan J, Partridge M. Assessing patients' preferences for characteristics associated with homeopathic and conventional treatment of asthma: a conjoint analysis study. Thorax 2001;Submitted.

240. Reardon G, Pathak DS. Segmenting the antihistamine market: an investigation of consumer preferences. Journal of Health Care Marketing 1990;10:23-33.

241. Reed Johnson F, Desvousges WH, Ruby MC, Stieb D, De Civita P. Eliciting stated health preferences: an application to willingness to pay for longevity. Medical Decision Making 1998;18:S57-S67.

242. Johnson FR, Banzhaf MR, Desvousges WH. Willingness to pay for improved respiratory and cardiovascular health: a multiple-format, stated-preference approach. Health Economics 2000;9:295-317.

243. Rosko MD, Walker LR, McKenna W, DeVita M. Measuring consumer preferences for ambulatory medical care arrangements. J Med Syst 1983;7:545-54.

244. Rosko MD, DeVita M, McKenna WF, Walker LR. Strategic marketing applications of conjoint analysis: an HMO perspective. Journal of Health Care

Marketing 1985;5:27-38.

245. Ryan M, Hughes J. Using conjoint analysis to assess women's preferences for miscarriage management. Health Economics 1997;6:261-73.

246. Ryan M, McIntosh E, Shackley P. Using conjoint analysis to assess consumer preferences in primary care: an application to the primary health card. Health Expectations 1998;1:117-29.

247. Ryan M, McIntosh E, Shackley P. Methodological issues in the application of conjoint analysis in health care. Health Economics 1998;7:373-78.

248. Ryan M, et al. Framing effects: a methodological issue in conjoint analysis. Health Economics 1999.

249. Ryan M. Using conjoint analysis to go beyond health outcomes: an application to in vitro fertilisation. Social Science and Medicine 1999;8:535-46.

250. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. British Medical Journal 2000;320:1530-33.

251. Ryan M, McIntosh E, Dean T, Old P. Trade-offs between location and waiting times in the provision of health care: the case of elective surgery on the Isle of Wight. Journal of Public Health Medicine 2000;22:202-10.

252. San Miguel F, Ryan M, McIntosh E. Demonstrating the use of conjoint analysis in health economics: an application to menorrhagia. Applied Economics 2000;32:823-33.

253. Vick S, Scott A. Agency in health care. Examining patients'

preferences for attributes of the doctor-patient relationship. Journal of Health Economics 1998.

254. Scott A. Eliciting GPs' preferences for pecuniary and non-pecuniary job characteristics. Journal of Health Economics 2001;20:329-47.

255. Singh J, Cuttler L, Shin M, Silvers JB, Neuhauser D. Medical decision-making and the patient: understanding preference patterns for growth hormone therapy using conjoint analysis. Medical Care 1998;36:AS31-AS45.

256. Spoth R. Simulating smokers' acceptance of modifications in a cessation program. Public Health Rep 1992;107:81-92.

257. Spoth R, Redmond C. Identifying program preferences through conjoint analysis:illustrative results from a parent sample. American Journal of Health Promotion 1993;8:124-33.

258. Szeinback SL, Mason HL. Variables affecting pharmacists' willingness to accept third-party prescription contracts: a conjoint analysis. Journal of Health Care Marketing 1990;10:45-50.

259. van der Pol M, Cairns J. Establishing patient preferences for blood transfusion support: an application of conjoint analysis. Journal of Health Services Research and Policy 1998;3:70-76.

260. Wigton RS, Hoellerich VL, Patil KD. How physicians use clinical information in diagnosing pulmonary embolism: an application of conjoint analysis. Med Decis Making 1986;6:2-11.

261. Bradley M. Stated preference experiment editor and designer (SPEED) user manual. Hague Consulting Group; 1991.

262. Stata Corporation. Stata reference manual release 7., vol. Volume 2 H-P. College Station, Texas: Stata Press; 2001. p. 580-94.

263. Ryan M. Establishing arguments in the infertile person's utility function using the economic instrument of conjoint analysis. HERU Discussion Paper 1995;8.

264. Ryan M. A role for conjoint analysis in technology assessment in health care? International Journal of Technology Assessment in Health Care 1999;15:443-57.

265. Jones A. Applied econometrics. London: Office of Health Economics; 2001.

266. National Institute for Clinical Excellence. Guidance on cancer services: improving outcomes in breast cancer. London: National Institute for Clinical Excellence; 2002.

267. Monthly Index of Medical Specialities, vol. January. London: Haymarket Medical Limited; 2002.

268. Diel IJ, Solomayer E, Costa SD, Gollan C, Goerner R, et al. Reduction in new metastases in breast cancer with adjuvant clodronate treatment. New England Journal of Medicine 1998;339:357-63.

269. Powles T, Paterson S, Kanis JA, McCloskey E, Ashley S, et al. Randomized, placebo-controlled trial of clodronate in patients with primary operable breast

cancer. Journal of Clinical Oncolgy 2002;20:3219-24.

270. NSABP protocol B-34. http://www.nsabp.pitt.edu/B34_Information.htm. Last accessed 9-10-2002.

271. The Medical Directory. FT Business Ltd; 1999.

272. Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgements. British Medical Journal 2004;329:224-27.

273. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Economics 2004;13:437-52.

274. Ryan, M. and Gerard, K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. Applied Health Economics and Health Policy 2(1), 1-10. 2003.

275. Ryan, M. and Skatun, D. Modelling non-demanders in choice experiments. Health Economics 13, 397-402. 2004.

276. Kontoleon, A. and Yabe, M. Assessing the impacts of 'opt-out' formats in choice experiment studies: consumer preferences for genetically modified content and production information in food. Journal of Agricultural Policy Research (5), 1-43. 2003.

277. Raiffa H, Schlaifer R. Probability and statistics for business decisions. New York: McGraw-Hill; 1959.

278. Raiffa H. Decision analysis: introductory lectures on choices under

uncertainty. Reading: Addison-Wesley; 1968.

279. Dixit AK, Pindyck RS. Investment under uncertainty. Princeton: Princeton University Press; 1994.

280. Dixit AK, Pindyck RS. The options approach to capital investment. Harvard Business Review 1995;73:105-15.

281. Mahajan V, Muller E, Bass FM. New product diffusion models, In: Eliashberg J, Lilien GL, editors. Handbooks in OR & MS, vol. 5.Elsevier Science Publishers; 1998. p. 349-408.