

**AN APPRAISAL OF METHODS TO INCREASE
GENERALISABILITY OF ECONOMIC
EVALUATION OF PHARMACEUTICALS**

Hege Fjell Urdahl
PhD in Health Sciences

University of York
2001

ABSTRACT

Health-care decision-makers are increasingly in need of information on the relative cost-effectiveness of healthcare interventions, evaluating meaningful outcomes in relevant populations. A gold standard for clinical and economic evaluation is the randomised controlled trial, in which context-specific factors are minimised to enable the assigning of treatment effect to the intervention under investigation. The limited generalisability of such evaluations may limit the relevance of study results to decision-makers.

This thesis presents three technical critiques of methods by which researchers may increase the generalisability of economic evaluation whilst focusing on transferability of results across patient populations and clinical settings. Firstly, meta-regression techniques were applied to trials with pragmatic and explanatory features to investigate whether increased external validity from pragmatic trials are achieved at the cost of reduced internal validity. Secondly, a checklist to explore the scope for economic modelling techniques to increase generalisability of results was applied to economic models identified in a literature review. Finally, a study design was developed for an observational patient record database (UK Mediplus ®) to determine the feasibility for its use in real-life cost-effectiveness analyses.

The findings indicate that pragmatic design features may be introduced to trials without jeopardising internal validity. Decision-analytic models have scope to synthesise data on cost and effect from several data sources to estimate cost-effectiveness for different clinical scenarios, however few published models make attempts to increase generalisability. The observational database has a potential to provide data on real-life drug use, resource use and clinical outcomes, however shortcomings in data quality and data management prevented the conduct of a full economic evaluation.

Pragmatic trials and observational studies may provide valuable data on cost and effectiveness reflective of specific clinical practice settings. Economic models have scope for presenting cost-effectiveness estimates representative across a range of settings. These methods should be adopted alone or in combination in order to assess the generalisability of the results of economic evaluations.

ACKNOWLEDGEMENTS

I am indebted to Nick Freemantle, Karen Bloor and Alan Maynard for their encouraging supervision throughout the period of study, and to Ian Russell for his constructive comments on the draft thesis.

I am also indebted to the Norwegian association for proprietary pharmacists (Norges Apotekerforening) for the funding of my independent studentship, and to the European Health Outcomes team at Eli Lilly and Company and Eli Lilly and Company UK, who jointly funded the observational study reported in this thesis.

At a more general level I am particularly grateful to Anne Burton for editing the manuscript; to Alison Dawson who provided support and practical advice; to Philip Young who advised on the statistical analysis of the observational study; to David Torgerson who provided constructive support at key research stages; and finally to Mike Drummond and the Centre for Health Economics who generously provided the facilities for practical study.

AUTHORS DECLARATION

All the research presented in this thesis was initiated and conducted by the author. In chapter 3 the author conceived the idea, developed the checklist, extracted all papers and conducted all analysis. In chapter 4 the author conceived the idea, developed the checklist, extracted all papers and provided the synthesis. The author conceived the idea for the study presented in chapter 5, ensured funding, developed the study design, and conducted all analysis. The author is completely responsible for the research presented in this thesis.

CONTENTS

ABSTRACT

ACKNOWLEDGEMENTS & DECLARATION

CHAPTER 1	Introduction	1
	Background	1
	Aims, objectives & research questions	3
	Source projects	5
	Structure & content of thesis	7
CHAPTER 2	Policy & Research Context: Generalisability of Economic Evaluation of Pharmaceuticals	9
	Economic evaluation in healthcare decision-making	9
	Generalisability in economic & clinical evaluation	15
	Economic evaluation guidelines & generalisability	24
	Themes of the thesis	29
CHAPTER 3	Effect size in pragmatic trials: striking the balance between internal and external validity of trials in schizophrenia and post MI	30
	Introduction	30
	Background	31
	Objectives	33
	Methods	33
	Results	41
	Discussion	45
	Conclusions & recommendations	50
CHAPTER 4	The merits of economic modelling as a means of generalising cost-effectiveness estimates: assessment of economic models in osteoporosis.	52
	Introduction	52
	Background	52
	Objectives	55
	Methods	55
	Results	59
	Discussion	66
	Conclusions & recommendations	71
CHAPTER 5	Developing and demonstrating methods for economic analysis in an observational patient record database	92
	Introduction	92
	Background & study material	93
	Aim & objectives	113
	Methods	115
	Results	130
	Discussion	154
	Conclusions & recommendations	167
CHAPTER 6	Discussion, conclusions and recommendations	171
	Source projects' contribution to methodological development	172
	Relative merits of alternative approaches; conclusions	178
	Recommendations	182
REFERENCES		184
APPENDICES		

CHAPTER 1

INTRODUCTION

1.1 Background

There is a growing interest world-wide in economic evaluation as a tool for the systematic assessment of costs and consequences of interventions in increasingly resource-constrained healthcare environments. Opportunity costs are associated both with the use of inefficient therapies in healthcare and with delayed introduction of new, efficient therapies and the results of these may be jeopardising nations' health and welfare. Pharmaceutical evaluation is time- and resource consuming, so there is increasing pressure to transfer findings across settings and countries.

Economic evaluations of pharmaceuticals combine an estimate of effect and an estimate of costs associated with the use of a medication, relative to an alternative treatment strategy. Results may be presented in a cost-effectiveness ratio expressing incremental costs and effects of one therapy compared with another or no therapy. Frequently, randomised controlled trials (RCT) provide the basis for the effect estimate, whether the economic evaluation is based on a decision-analytic model or individual patient-level data. Randomised controlled trials may also provide the basis for cost estimates, predominantly in trial-based economic evaluations.

Pivotal RCTs are designed to meet licensing requirements. A high degree of internal validity may be achieved by conducting evaluations in specialist settings, through the selective inclusion of patients, by allocating treatment to eligible patients through the process of randomisation, by blinding physicians and patients to allocated treatment and through the process of monitoring patient care.

Concerns have been raised that laboratory care conditions of such a clinical trial may limit the external validity of the findings, and whether findings derived within the setting of a randomised controlled trial would automatically translate into clinical

effectiveness as experienced in actual clinical practice. (Drummond and Davies 1991, O'Brien 1996, Fayers and Hand 1997, Coyle et al 1998) For economic evaluations, this may also result in the recording of protocol-driven costs. Increasingly, attention has been focused on the external validity of economic evaluations and the interest to developing methods for evaluating healthcare interventions across settings has been growing. The issue of generalisability of economic evaluation studies is currently on the research agenda of the Health Technology Assessment Programme for the UK National Health Service (HTA project 98/22/05) and for the WHO guidelines for cost-effectiveness analyses. (Murray et al 2000)

The users of economic evaluations are likely to ask themselves whether the results of a given evaluation are useful and applicable in their health care setting. First and foremost, the methods employed to produce the evidence should normally be scrutinised when assessing the validity of results. The question of usefulness would however also encompass a judgement about whether the results may be transferred in time, across populations or between clinical contexts. These three are the main axes across which the generalisability of an economic evaluation may be assessed.

Key clinical components include the age and gender of the patient group, co-morbidities, severity of disease and patient acceptability to the intervention. Clinical trial setting, the practice within which patients are treated, the health care system within which a study is framed, or the country in which the study is being conducted may determine context-specific components. Finally, the introduction of new treatments, change in the perceptions of a clinical condition or an intervention or simply the more effective use of an intervention as experience accumulates are components that may determine generalisability of results in time.

When assessing generalisability of economic evaluation evidence, whether pertaining to time, setting or population, decision-makers should examine closely the data used to produce the evidence. This thesis primarily focuses on generalisability across alternative patient populations and contexts, and therefore views generalisability in terms of the degree to which results from one evaluation

hold true for different patients and clinical settings. Key elements of economic evaluations that may be modified in order to reflect alternative settings and patient populations, or study designs that may provide real-life cost effectiveness estimates are outlined. Such an evaluation of alternative methods that may be used in isolation or in conjunction, may assist users of economic evaluations in assessing generalisability of results they encounter in published studies and to identify ways in which further data may be generated in order to make appropriate extrapolations.

One of the problems with economic evaluation methodology is the paucity of empirical data to illustrate the extent to which environmental factors limit the generalisability of the evaluations and the degree to which effect estimate, resource use, and unit cost may differ across settings. There are also shortcomings in methods illustrating how these problems could be overcome. Neither has there been, previously, a systematic appraisal of the methods available to health economists to take these issues into consideration in economic evaluations. Methods need to be developed to evaluate outcomes of use of pharmaceuticals in clinical practice.

The thesis provides a critical evaluation of methods that are available to [researchers that evaluate the generalisability of primarily clinical parameters in economic evaluations of pharmaceuticals] those wishing to increase generalisability of economic evaluations of pharmaceuticals. The research draws on insights of several fields of research in medicine, including health policy and economics, epidemiology and medical statistics, research disciplines that have added to an accumulating knowledge base surrounding clinical and economic evaluation methods.

1.2 *Aim, objectives and research questions*

The aim of this thesis is to assess the relative merits of methods that can be applied to increase generalisability of economic evaluations of pharmaceuticals across

clinical settings and to further develop methods currently available methods. Specific objectives and key research questions are outlined below.

1.2.1 Research objectives

- To identify methods to increase external validity of estimates of costs and effectiveness in economic evaluation;
- To develop a framework for assessment of whether economic models accommodate variation between settings in analysis and presentation of results;
- To examine the degree to which economic models attempt to generalise cost-effectiveness estimates;
- To evaluate whether pragmatic trial design features provide effect estimates systematically different from explanatory features;
- To explore the feasibility of an observational patient record database to accommodate full economic evaluation;
- To apply observational study methodology to generate cost-effectiveness estimates based on clinical practice;
- To identify and consider key methodological issues for the development and future conduct of economic evaluations with increased external validity.

1.2.2 Research questions

- How can learning from clinical evaluation methodology be incorporated into economic evaluation methodology to meet the need for data transferable between settings?
- Which design features characterise a pragmatic trial and do these provide a biased result?
- How can economic models incorporate setting-specific aspects in the evaluation?

- Which purpose can observational data serve in the provision of economic evaluation data relevant to clinical practice?

1.3 Source projects

The thesis seeks to meet the research objectives through design, conduct and evaluation of three source projects, each exploring a single method that may be used in the process of evaluating generalisability of results of economic evaluations between clinical settings and populations. Pragmatic trial design, modelling approach and observational study design will be proposed as methods for increasing external validity of economic evaluation studies. Through the reporting of these projects, this thesis aims to contribute to an increasing knowledge-base on generalisation methodology by evaluating aspects of these approaches. Although these methods are explored individually in this thesis they are not viewed as alternative approaches but rather as complementary processes by which generalisability may be explored and achieved. The source projects do not focus on one particular therapeutic area, though pharmaceutical interventions in the clinical areas of osteoporosis, schizophrenia and treatment after myocardial infarction were used in the source projects for the purpose of demonstrating methods.

1.3.1 Pragmatic trials

Relevant clinical and economic data can be derived from real-life settings, but only at the cost of an increased risk of bias and confounding. Pragmatic trials adopt features to ensure a higher external validity to the setting within which the trial is conducted. (Schwartz and Lellouch 1967) The most important modifications of clinical trial design in pragmatic trials include research question and objectives; patient selection; omitting blinding of trial participants; relaxing of monitoring protocol; evaluating a final clinical endpoint and adopting intention-to-treat statistical

analysis. The adoption of pragmatic design features in economic evaluation trials has been extensively advocated. (Revicki & Frank 1999, Simon 1995, Coyle et al 1998, Guscott and Taylor 1994, Roland and Torgerson 1998, Brixton et al 1998) Previous assessments of trial estimates as a function of quality aspects of randomised controlled trials through the use of meta-regression provided methods on which this source project could build and expand. (Schulz et al 1996, Moher et al 1998)

The first source project of this thesis examines whether pragmatic design features in clinical trials systematically moderate the effect size compared to trials adopting explanatory features, and therefore achieve higher external validity at the expense of internal validity. The evaluation makes use of two systematic reviews of pharmaceutical interventions in two therapeutic areas to examine the predictive effect of pragmatic design features on effect size.

1.3.2 Economic models

Economic models have been promoted as the method with which economic evaluation researchers may vary the model input in order to make the results applicable to different healthcare settings. (Buxton et al 1997, Briggs and Gray 1999, Commonwealth of Australia 1995, CCOHTA 1996)

Researchers and authors of guidelines have pointed to economic models as a relevant methodology to accommodate differences in regional variability. The second source project aims to assess the extent to which models in one therapeutic area have accounted for such differences to date. A checklist was developed which described aspects of clinical and economic features of the model, as well as checks for external validity. This was then applied to a selection of models identified in a systematic review.

1.3.3 Observational studies

The scarcity of methods for generalising trial results to clinical practice settings led to the search for alternative approaches capable of providing real-life cost-effectiveness estimates. Observational databases in which patient information is collected longitudinally may be valuable sources of drug use data, reflective of clinical practice. The potential value of observational data in economic evaluation is outlined and a review carried out of the methodological and statistical tools adopted by researchers in epidemiology. The third source project presents the design and results of two studies that were developed as parts of a pilot project to evaluating cost-effectiveness for pharmaceuticals in an observational patient-record database.

The studies examined whether an observational database in the UK (the Mediplus® database) was a feasible tool for the conduct of a full economic evaluation of cost-effectiveness in clinical practice. The Mediplus® database contains primary care patient records collected in clinical practice and may be a potential vehicle for economic evaluation in the primary care setting, provided that the database meets the requirements to clinical and economic evaluation.

1.4 *Structure and content of thesis*

A synthesis of different strands of literature outlining the need for economic evaluation and in particular, the generalisability of such evaluation, is provided in Chapter 2. This contains a brief review of the methodological foundations of economic evaluation analysis, as well as a review of the issue of generalisability in the context of clinical and economic evaluation. Finally, the decision-maker's expressed need for generalisability of economic evaluation is evaluated from the viewpoint of those regulating formal submission of economic evidence required within some jurisdictions. Three methods through which analysts may explore external validity of evaluations emerge from this review, methods that researchers may use either individually or in combination. Chapters 3 through 5 are devoted to evaluating the relative merits of each of these methods separately.

Chapter 3 evaluates the pragmatic trial approach to generalising economic evaluations and Chapter 4 offers a critique of the degree to which economic models adopt values for adjustment of model parameters or for sensitivity analysis that would facilitate judgement of external validity. The aim of Chapter 5 is to provide an assessment of the feasibility of an observational database to provide data for a full economic evaluation, reflective of clinical practice. All chapters have a similar structure: a background that puts the source project into the context of the overall theme of generalisability; an outline of research aims and objectives; a detailed description of methods; presentation of results; a discussion of key findings and the merits of the examined methods to increase generalisability of economic evaluations.

The focus of the three source projects differs. The emphasis of the first was to apply meta-regression techniques to *analyse* the balance between internal and external validity in pragmatic trials. Central to the second project was the systematic search for economic evaluation models and structured *assessment* of the degree to which models may accommodate context-specific variation in key model parameters. The last project focused on further methodological development by exploring the *feasibility* of observational patient record data for use in economic evaluations striving for high external validity. These differences in emphasis are reflected in the reporting of the source projects. For example, particular attention is given to the statistical methods adopted in Chapter 3, the development of an assessment framework is central to Chapter 4, and the application of economic evaluation to observational study designs has been given detailed consideration in Chapter 5.

In the concluding chapter of the thesis, attention is turned to the impact of the individual source projects on the further development of methods to increase generalisability of economic evaluations. A discussion of the relative merits of pragmatic trials, models and observational data to synthesise relevant (generalisable) data for economic evaluation, provides an assessment of the contribution of this research in the field of generalisability. Methodological recommendations conclude the final chapter of this thesis.

CHAPTER 2

POLICY AND RESEARCH CONTEXT: GENERALISABILITY OF ECONOMIC EVALUATION OF PHARMACEUTICALS

The development of methods to increase the generalisability of economic evaluations of pharmaceuticals across settings is a focus throughout this thesis. In healthcare, the methods for clinical evaluation and economic evaluation are increasingly interrelated, so both play significant roles in this research. Strands of literature covering research areas relevant to generalisability of clinical and economic evaluation are reviewed in this chapter. Attention is also given to the requirements for the generalisability of economic evaluations by those jurisdictions that systematically use such information in policy-making. Finally, emerging themes underlying the research undertaken for the thesis are outlined. Before turning to the review of empirical research in the area of generalisability of clinical and economic evaluation, it is useful to give the theoretical foundations of economic evaluation some consideration.

2.1 Economic evaluation in healthcare decision-making

Medical care has changed over the last decades, with new therapies being discovered that may prevent or treat previously untreated conditions, and new technologies replacing old procedures. The availability of new and costly health-related interventions has gradually exceeded society's ability to afford them. (Mosteller 1985) The reasons for the growth in healthcare spending include demographic changes in populations as the proportion of elderly increases, the development of new and more costly therapies and the increasing availability of lifestyle treatments. Also, previously untreated conditions may now be treated with new discoveries.

A peculiarity of the market for medical care adds to this increase in aggregate healthcare costs. According to conventional economic theory, independent

producers and consumers trade goods freely according to price and quantity. Price in such a market will reflect the margin of consumer valuation of the goods. (McGuire et al 1988) The characteristics of the healthcare market deviate from assumptions underlying this model. In healthcare, the physician acts as an informed decision-making agent on behalf of the patient, and the costs are borne by a third-party, for example an insurance organisation or a national health service. Economic theories of supply and demand and their interaction with price and consumption of goods fail to achieve equilibrium in the healthcare environment. Supply and demand sides are no longer separated resulting in the absence of a working price mechanism and the phenomenon of supplier-induced demand. (McGuire et al 1988)

The continuing rise in healthcare costs has forced a change in health-policy paradigm, where healthcare systems have moved from what is possible given technical constraints to what is possible given economic constraints. (Maynard and Bloor 1998) Writing on the UK National Health Service (NHS), Cochrane (1971) made the link at an early stage between the issue of clinical-effectiveness and efficiency in the health service, emphasising the role of clinical decision-makers to take this into account:

“The main job of medical administrators is to make choices between alternatives. To enable them to make the correct choices they must have accurate comparable data about the benefits and costs of alternatives.”

Williams (1992) took this new emphasis on cost-consciousness one step further by arguing that:

“A caring, responsible and ethical doctor has to take costs into account. Indeed it is unethical not to do so!”

Williams refers to the concept of ‘opportunity cost’ of resources, or the value of the resources in their best alternative use, alluding to the fact that resources spent in one area of the healthcare sector or on one patient group may not be used in other areas or on other patients. This focus on efficient use of resources in healthcare

has provided strong incentives for clinical decision-makers to take make use of health economic data when informing decisions. Clinical decision-makers include physicians, healthcare providing organisations and governments. The individual roles of these stakeholders differ, but their broad remit is to ensure safe and effective provision of care to the population.

Economic evaluation techniques have been adapted to healthcare provision and developed for the analysis of individual services, interventions and programmes within healthcare from their origin in disciplines such as engineering. (Williams 1993) In healthcare, economic evaluations may provide an alternative to the failing price mechanism in the healthcare market and a guide to resource allocation. (Russell et al 1996) A central measure in economic evaluation is the cost-effectiveness ratio. Implicit in the cost-effectiveness analysis is comparison between a treatment and an alternative. The cost-effectiveness ratio for the comparison is the difference in costs divided by the difference in effectiveness, and therefore expresses the incremental cost of obtaining an additional unit of health effect. (Torrance et al 1996) The health effect component of the cost-effectiveness ratio can be expressed differently. Economic evaluation analyses that express the health effect in terms of a physical unit such as life years gained are called cost-effectiveness analysis. Those analyses that value the health outcome by assigning a utility such as the quality adjusted life year gained (QALY) are denoted cost-utility analyses. Finally, cost-benefit analyses value the health outcome in financial terms. (Drummond et al 1997)

Theoretical foundations for research methods may guide the conduct of research, understand its limitations and guide the decision-making process following introduction of new evidence. To date, health economists debate the theoretical foundations of the methods used in economic evaluation. The roots have predominantly been traced to welfare economics, which is concerned with the societal allocation of resources in order to maximise the welfare of an affected population. (Garber et al 1996, Hurley 2000) If the fundamental purpose of health economic analysis is to improve general welfare through improving health then it may be placed within this context. The strand of welfare economic theory advocated by some health economists is the von Neuman-Morgenstern

utilitarianism, a theory that assumes that individuals have a well-defined utility-function and that the overall welfare of society is a function of such individual preferences. (McGuire et al 1988)

Welfare economists disagree amongst themselves as to how healthcare decision-makers should prioritise societal resources, and economists broadly advocate two normative frameworks, the neo-classical framework and the extra-welfarist framework. (Hurley 2000) Neo-classical welfare theorists aim to maximise societal utility, with little emphasis on the distribution of the gain. Key assumptions underlying the utilitarian view is that social welfare is made up of from the welfare of each individual member of society, and that individuals are best judges of their own welfare. The aggregate social utility function resulting from a situation where an allocation makes at least one person better off and no one worse off it is said to present an actual Pareto improvement. A potential Pareto improvement may occur when those that are better off as a consequence of a resource allocation compensate for those that are worse off after the allocation. A resource allocation is Pareto optimal (i.e. allocatively efficient) if it is impossible to increase one persons utility without simultaneously decreasing another's. The benefit of healthcare interventions is of primary interest to those subscribing to the neo-classical view, and benefits in empirical welfare analysis are predominantly valued in monetary terms. Neo-classical Paretian welfare economic theory therefore provides the conceptual framework of cost-benefit analysis, consistent with the aim of maximising societal utility. (Hurley 2000)

Williams (1993) made a case for the extra-welfarist perspective, which provides an alternative to defining social utility as an aggregate of individual utilities by giving special weight to health in social accounting and the distribution of health. From this perspective, health is viewed as the output of the healthcare sector and the social objective is to maximise health subject to resource constraints. (Garber et al 1996) Benefits of healthcare interventions are predominantly expressed as a subjective health measure, for example quality adjusted life years (QALYs) gained.

Cost-effectiveness analysis and cost-utility analysis compare health care programmes that produce similar units of outcome. Cost-effectiveness analysis has

been considered a framework for informing questions of technical efficiency, where two ways of achieving the same goal are compared, for example, the comparison of competing therapies in same clinical condition. In contrast, cost-benefit analysis is broader in scope because it assigns relative values to health-related goals to determine which goals are worth achieving. The cost-benefit analysis framework has therefore been viewed as appropriate for policy questions of allocative efficiency, where decision-makers aim to assess whether an intervention is worthwhile in the first place or whether the resources should be invested elsewhere. Cost-utility analysis may also be considered appropriate for decisions of allocative efficiency since the expression of outcome in one generic clinical term, the QALY, informs policy-decisions on the efficient allocation of resources between programmes that produce different specific health outcomes.

The interpretation of results from cost-benefit analyses is, in principle, straightforward: if benefits exceed costs then the programme should be implemented. (Briggs and Gray 2000) Two broad rules can be adopted by decision-makers who use evidence produced by cost-effectiveness and cost-utility analyses. Rank-ordering all possible uses of resources in terms of their incremental cost-effectiveness ratio and working down the list implementing the most cost-effective interventions until the healthcare budget is exhausted is called the “league table approach”. Inherently, this approach aims to maximise health within a given budget, but has been criticised for inappropriately comparing cost-effectiveness ratios derived from inconsistent use of methods and ‘unthinking’ decisions. (Maynard 1991, Drummond et al 1993) Alternatively, the shadow price decision rule incorporates a cost per QALY that decision-makers are willing to pay for additional health, i.e. an acceptable cost-effectiveness ratio reflecting decision-makers’ willingness to pay. Interventions with an incremental cost-effectiveness below this threshold would then be adopted. Besides the fact that it has proven difficult to quantify the decision-makers’ willingness to pay for additional health benefits, the shadow price rule has been criticised for being a “prescription for growth in healthcare expenditures”. (Briggs and Gray 1999)

The debate over theoretical foundation has guided research into issues such as costing methods and valuation of health, (Hurley 2000) but despite this, economic

evaluation is commonly looked upon as a pragmatic solution to the need for information to prioritise healthcare resources and maximise health. (Garber et al 1996) Economic evaluation has a multi-disciplinary research approach to evaluation, encompassing disciplines such as economics, decision analysis, medical sociology, epidemiology and trial methodology. The most pragmatic view of economic evaluation may be that of a complementary process to clinical evaluation, but with the aim of informing healthcare utilisation decisions on a broader public health scale.

Within the increased interest in health economics' contribution to healthcare provision, the focus on pharmaceutical expenditure has been particularly strong. (Maynard and Bloor 1998) Many measures have been undertaken in attempts to reduce and control healthcare costs, such as reference price systems, negative lists, and price and profit controls. (Freemantle and Bloor 1996, Bloor and Freemantle 1996, Bloor et al 1996) Government cost containment policies and price regulation may have driven the demand for economic evaluation analysis of pharmaceuticals. However, the regulatory demand for data on effectiveness of pharmaceuticals for registration purposes, and the resulting availability of such data, also provides a good basis for economic evaluation.

Despite increasing interest in economic evaluation analysis of pharmaceuticals, there has been little evidence that such data contribute systematically to resource allocation decisions. (Drummond and Cooke et al 1997) There may be barriers to the use of economic evaluation data at both political and practical level. Firstly, there are a number of challenges to the discipline and concern that the methodology is being used adequately. (Udvarhelyi et al 1992, Drummond 1992, Byford and Palmer 1998) The thorny methodological issues include the valuation of health outcomes, discount rate, inclusion of future costs, statistical analysis and assessment of uncertainty. Secondly, economic models have been criticised for having a 'black box' feel to them. (Sheldon 1996) The lack of insight and ability to scrutinise methodology has led some researchers to propose that computer models might be submitted for peer review. (Hill et al 2000) Finally, economic evaluations of pharmaceuticals are frequently based on randomised controlled trial data, but critics have argued that pragmatic prescribing decisions may not best be informed

by estimates of 'cost-efficacy' in an ideal trial setting, rather by an estimate of 'cost-effectiveness' in settings reflecting clinical practice. (Coyle et al 1998)

Drug evaluation, including economic evaluation, is a costly undertaking and there is a pressure to use data generated in one setting and apply the findings to different settings. Adapting economic evaluations to local circumstances by adequately accounting for different patients, settings, regions and countries may be essential to the value of the evidence to local decision-making. Little is known about how well cost and clinical data translate between clinical settings and across countries, and therefore whether cost-effectiveness ratios vary as a function of the setting in which it is conducted. Potential biases can enter the process of transferring data between settings, resulting in premature introduction of inefficient therapies or similarly, delaying the use of efficient therapies. Hence the importance of developing methods that can be adapted to make a systematic evaluation of generalising across settings.

2.2 Generalisability in economic and clinical evaluation

A convenient introduction to issues of generalisability in clinical and economic evaluation research is provided by the following quotation:

“Internal validity implies that the differences observed between groups of patients allocated to different interventions may, apart from random error, be attributed to the treatment under investigation. In contrast, external validity, or generalisability, is the extent to which the results of a study provide a correct basis for generalisations to other circumstances. In itself, there is no external validity. The term is only meaningful with regard to specified ‘external’ conditions, such as other patient populations or treatment regimens. Internal validity is a prerequisite for external validity: the results of a flawed trial are invalid, and the question of external validity becomes redundant.” (Juni et al 2001)

A high degree of internal validity can be achieved in the randomised controlled trial (RCT), where the randomisation procedure, subject selection, monitoring procedures and scientific process, aim to address bias. Sackett (1985) argued that the double-blind placebo controlled trials that follow a defined protocol represent the 'gold standard' for drug-trials, primarily because of the ability of randomisation to deal adequately with bias. The focus of the conventional RCTs on internal validity may have resulted in limiting their usefulness and raised concern that the uncertainty in the assumption that overall results from clinical trials may be extrapolated to patients outside the trial jeopardises the external validity of trial results. (Lancet 1994, Rothwell 1995, Black 1996) The setting within which a study is conducted may encompass specific geographical and national location, type of healthcare facility, patient population and period of study. As the authors of the quote above point out, generalisability may be irrelevant without reference to one or more specific settings outside that of the evaluation.

The randomised controlled trial has been adopted as the 'gold standard' also for clinical estimate of effectiveness in the denominator of the cost-effectiveness ratio. (Drummond 1997) Researchers undertaking economic evaluations of pharmaceuticals over the last decades have rightly focused on internal validity of the study, ensuring validity of effect estimates and acceptance by the medical community. However, concerns have been raised that prescribing decisions may not be best informed by the results of randomised trials frequently conducted for licensing purposes, but rather by data reflecting clinical practice patterns of drug use and resource consumption. (Coyle et al. 1998)

The limitations of trial evidence for use in economic evaluation need to be addressed by the development of methods to maximise external validity of the findings to other settings. In the process of developing such methods, lessons may be learned from the clinical literature regarding methods that could be used to optimise generalisability of clinical evaluations. Pragmatic trial methodology, meta analyses and observational data may serve this purpose in clinical research. Literature covering generalisability in clinical evaluation will be reviewed before returning to the special case of generalisability within economic evaluation.

2.2.1 Issues of generalisability in clinical evaluation

Pharmaceutical development has a long tradition of systematic evaluation of clinical evidence. The practice of clinical trials has developed rapidly during the last 50 years. This development was parallel to an increase in the regulatory requirements to the licensing of pharmaceuticals, some of which were introduced after the Thalidomide-scandal in the 1960s. (Pocock 1983) The evaluation of quality, safety and efficacy is central to these requirements; they are the 'three hurdles' that a manufacturer needs to jump in order to get a drug licensed by influential regulatory bodies such as the Food and Drug Administration in the US (FDA), and the European Medicines Evaluation Agency (EMA).

Because treatment effects are often relatively modest, biases may overwhelm the effect of treatment unless these are adequately controlled for. The process of randomisation is considered crucial to the minimisation of bias in clinical evaluations. (Schulz et al.1996) Random allocation of eligible patients to the intervention and control groups of a trial ensures that the features of the patients and their therapists are distributed across the treatment arms of the experiment by the play of chance. Randomisation does not make groups equal, but forms a good basis for comparison. (Freemantle et al. in press) Statistical methods examine the extent to which observed differences between groups may be attributable strictly to chance, indirectly providing information on the likelihood that a difference in outcomes observed in a trial may be attributable to the different treatments allocated to those groups.

The clinical stage of drug development is broadly divided into four phases. (Piantadosi 1997) Phase I clinical trials are the first experiments using a pharmaceutical compound in a population, when investigators seek to assess the safety of the drug in the human body and identify an appropriate dose-range required to intervene with the clinical condition without causing serious side-effects. During phase II of clinical drug development, investigators assess the feasibility of the treatment and estimate treatment effects. Phase III trials are experiments of the therapy escalated to a larger sample of the patient population, where the impact of the therapy on the clinical condition of interest is compared to placebo or standard

therapy. In a regulatory context such trials are called *pivotal*, and these are frequently designed in order to meet, for example, the US Food and Drug Administration (FDA) requirement to licensing, of two trials in favour of new treatment showing significant inference at single sided 0.025 level for the primary clinical outcome measure. (Freemantle 2001) Finally, phase IV trials are often large-scale experiments looking for uncommon treatment effects after licence has been granted. Key to trial conduct is a protocol uniform to all trial participants that determines the clinical management of patients in order to minimise the impact of the strategy of care on the treatment outcome and ensure internal validity of the findings from the trial.

Inherent in the trial protocol of an RCT is a series of adjustments to clinical practice. Subjects included in clinical trials may be selected by narrow inclusion criteria. For example, the exclusion of the elderly and of women from cardiovascular trials restricts the scope for generalising results to these patients. (Wenger 1992) Run-in periods in trials, during which patients are selected for inclusion, may also contribute to an atypical trial patient population, for example, more compliant patients may be chosen. (Pablos-Mendez et al. 1998) Finally, blinding of patients and physicians to treatment allocation, the manipulation of pattern and quality of care by the trial protocol and the monitoring of trial participants may restrict the scope for generalisation of trial results. (Rothwell 1995)

Trials that are designed to measure treatment outcome under ideal conditions have been labelled 'explanatory'. (Schwarz and Lellouch 1967) They can be viewed as models of clinical practice in which reality has been simplified and standardised to provide 'laboratory' conditions for the experiment. In clinical research, the term 'efficacy' is frequently used to describe the performance of a treatment under such conditions. The adjustments in care demanded by the trial protocol are in place to ensure internal validity, but may limit the external validity of the trial. Extrapolating the findings from explanatory trials to clinical practice, where patients and physicians seek positive outcome in a real life environment, may not be appropriate. Pragmatic trials, meta-analysis and evaluation of observational data are methodological approaches that may be used in an attempt to increase the generalisability of evaluations.

Firstly, 'pragmatic' trials are designed to measure clinical-effectiveness under conditions reflective of clinical practice. (Swartz and Lellouch 1967, Roland and Torgerson 1998) Frequently, such trials relax the design features of explanatory trials, for example by inclusion of a more heterogeneous patient population, avoidance of blinding of participants to assigned treatment, comparison with active substance and use of a final clinical endpoint, such as mortality, as the primary clinical outcome. Trials are complex structures, comprising a number of features ranging from subject inclusion and care provision to statistical analysis. The suggestion that trials can be dichotomised as either explanatory or pragmatic is somewhat simplistic. Rather, trials can be viewed as part of a spectrum which stretches from those in which most aspects of the design reflect the clinical setting in which they are conducted to those that modify most aspects of subject constituency and care provision, according to the features they adopt. The more it reflects a real life practice situation, the closer the trial is to the 'pragmatic' end of the continuum. Achieving a study design that adequately ensures internal validity of the findings while increasing the external validity of the results may be a matter of striking the right balance in trial design features. The degree of external validity may, for example, increase throughout the four clinical development phases.

Secondly, in clinical research, statistical pooling techniques have been developed as a means of summarising evidence from several trials investigating similar outcomes. (Egger and Smith 1997) Meta analysis has several merits. By combining a number of trials, it increases the power of the evaluation of a clinical question by reducing the probability of a false negative result. But more importantly, in this context, meta-analysis may also aid the generalising of trial results by combining, for example, trials with different patient case-mix or conducted in different healthcare settings, making results applicable to a more diverse patient population and to a variety of clinical settings.

Finally, observational study designs represent a third alternative, achieving high external validity by basing the estimation of treatment outcomes on care received in clinical practice. It has been proposed that observational data are essential in clinical research and may expand the evidence base for healthcare therapies.

(Black 1996, Lewsey et al 2000, Radford and Foody 2001) Nevertheless, the major drawback is the lack of randomisation, which leaves the results of the studies susceptible to biasing factors that jeopardise study validity. Statistical modelling is routinely used to minimise confounding from observed variables (Hennekens and Buring 1987), with propensity analysis and instrumental variables methods providing further risk-adjustment techniques to control for residual confounding from unobserved differences in observational studies. (Radford 2001, McClelland et al 1994)

To understand the consequences of potential bias in estimating effect size, researchers have evaluated the way in which these estimates travel between randomised studies and observational (i.e. non-randomised) studies. Authors of empirical studies debate whether observational and randomised studies that evaluate the same treatment in similar populations systematically provide different magnitude and direction of the effect size estimate. Several authors have reported systematic comparisons of the two methods, but findings from these are inconclusive. (Colditz et al 1989, Kunz and Oxman 1998, Reeves B et al 1998, Britton et al 1998, MacLehose et al 2000) For example, Concato et al (2000) evaluated 99 trials and observational studies in five therapy areas and found that the results of well-designed observational studies do not systematically overestimate the magnitude of effects of treatment as compared with randomised controlled trials on the same topic. Similarly, Benson and Harz (2000) found little evidence that estimates of treatment effects between observational and randomised evidence were consistently different, but both sets of authors were criticised for having used selected and unrepresentative samples of trials and studies. (Pocock et al 2000) In contrast, the findings of Ioannidis et al (2001), who reviewed 240 randomised trials and 168 observational studies in 45 topics, found that discrepancies beyond chance between randomised and observational studies do occur and that differences in estimated magnitude of treatment effect are in fact very common.

Clinical researchers have a variety of designs in their toolbox to answer different research questions about the clinical impact of pharmaceutical treatments. We have seen that pragmatic design features may be adopted in trial design to increase

external validity, that meta-analytic techniques may be used in generalisation and that observational studies may provide data with high external validity, albeit at the risk of jeopardising internal validity. These clinical evaluation tools are also available to those undertaking economic evaluations, providing the opportunity to present economic evaluation estimates with varying degree of external validity relevant to the clinical situation of interest to a decision-maker.

2.2.2 Issues of generalisability in economic evaluation

In the economic evaluation of pharmaceuticals, the cost-effectiveness ratio may be estimated based on patient-level clinical and economic data or estimated in a decision-analytic modelling technique in which the analyst synthesises cost and outcome data from multiple sources. (Weinstein and Fineberg 1980, Drummond and Stoddart 1984, Johnston et al 1999) The cost-effectiveness ratio is a collapsed measure of both cost and effectiveness, and is sensitive to variation in estimates of both resource use and clinical-effectiveness. There is limited empirical data on how this statistic responds to setting-specific variation, but an analysis by Willke et al. (1998) provided convincing evidence that differences in healthcare setting are important to the cost-effectiveness ratio.

Because of its status as the gold standard for clinical evaluation, the randomised controlled trial is commonly considered the preferred methodology on which to base the effect estimate in economic evaluations, whether it is a patient-level data analysis or a model-based evaluation. (Drummond et al 1997, Gold et al 1996) Problems encountered when running an economic evaluation alongside a clinical trial are extensively documented. (Drummond and Stoddart 1984, Drummond and Davies 1991, Donaldson et al 1996, O'Brien 1996, Gray et al 1997, Fayers and Hand 1997, Coyle et al 1998) Discrepancy between the setting of randomised clinical trials and that of clinical practice is a common theme. Of particular importance to economic evaluation is the fact that care modified by the trial protocol may result in the recording of protocol-driven costs and subsequent reduced ability to present a realistic view of resource consequences or distinguish between treatment arms of a trial. (Mauskopf et al 1996) Inadequacies in the design of trials

by selective choice of comparator, exclusion of patients with certain characteristics and limited duration of follow-up may not only limit the generalisability of trials but indeed bias the results. (Freemantle and Maynard 1994) Furthermore, technical challenges for researchers undertaking trial-based economic evaluations include sample size and statistical evaluation of trial-based evaluations originating from statistical properties of the cost-effectiveness ratio. (Briggs 2000, O'Hagan and Stevens 2001) To overcome the limitations of the conventional randomised clinical trial for economic evaluation, economic evaluations may be based on pragmatic trials, modelling exercises or observational studies.

Firstly, clinical management patterns in trials adopting pragmatic features may, to a lesser degree, be manipulated by the protocol so it is increasingly common to incorporate economic parameters alongside them. (Oster et al 1995, Simon et al 1995, Simon et al 1996, Revicki and Frank 1999) The increased conduct of economic evaluations alongside randomised clinical trials of pharmaceuticals in several research centres and across more than one country has also offered progress to the generalisation of cost-effectiveness estimates. Differences in demography and epidemiology of disease, clinical practice and conventions, incentives and regulations for healthcare providers, relative price levels and consumer preferences are key components of the threat to generalisability of economic evaluations across countries. (O'Brien 1997) A pharmaceutical intervention can be cost effective in one country but not in another. In an evaluation of tirilazad mesylate for subarachnoid haemorrhage, Willke et al (1998) used a regression-based approach to the trial data and found that there were significant country-specific differences in total patient costs. The authors concluded that generalisation of trial-wide cost results to specific countries would be inappropriate.

Unfortunately, one limitation of multi-centre studies is that they are rarely powered to detect centre-specific difference in treatment effect and costs, so only relatively large cost differences can be detected. The most common method of estimating treatment costs per patient is to take average or standard unit cost for each resource item of interest and apply this to all costs recorded. (Schulman et al. 1996) A less common approach is to use unit costs specific to each centre. (Glick et al. 1998) Raikou et al (2000) showed in a simulation experiment that there is a

difference in the estimates gained from these two methods, and that a calculation based on average or standard unit costs would overestimate treatment costs. This poses a dilemma for the analysis of multi-centre cost-effectiveness studies.

Generalisation according to Juni et al (2000) is reliant upon relevance to a specific setting, so results would need to be presented with unit costs relevant to each participating centre. The remaining question is whether individual centre-specific cost-effectiveness can be presented by assigning unit costs to the aggregate trial results or only to the patients in the relevant centre.

Secondly, the model-based approach to economic evaluation can, potentially, evaluate the relative cost-effectiveness of therapies under a variety of circumstances and settings. A model may be based on inclusion of data from more than one clinical trial, incorporate the opinion of clinical experts, and make use of large longitudinal databases with patient-level information. The construct of a model enables testing of generalisability by incorporating a range of model inputs reflective of different scenarios. Briggs and Gray (1999) proposed the sensitivity analysis as an appropriate stage in a modelling exercise for exploring uncertainty relating to location and population. External validity of an economic evaluation may, for example, be increased by the use of a pooled effect estimate from many clinical trials. (Pang et al 1999)

Finally, there have been suggestions that observational data could be applied when generating information on resource consumption and treatment patterns in clinical practice. (Gold 1996, Drummond 1998, Sheldon 1996) One advantage of using clinical trials as a framework for economic evaluation is that they provide the opportunity to collect and analyse patient-specific resource use data, a feature also shared by most observational data. (Johnston et al 1999) More frequently, large observational databases are being used to assess clinical and cost-effectiveness of pharmaceuticals, particularly in the US, where these databases exist for the purposes of reimbursement. (Hornberger and Wrone 1997, Sacristan and Soto 1994) Economic evaluation undertaken in an observational context dispenses with the need for a protocol for care, enabling recording of resource quantities as they would normally occur. Also, being less resource-consuming to run than clinical trials, observational studies may enable researchers to study larger patient

population over longer periods of time, measuring clinically relevant final endpoints rather than intermediate endpoints. (Drummond 1998)

Effect estimates with observational data are not protected from bias, and the degree to which an estimate is biased is not knowable. (Davies and Crombie 1999) For data to be useful for clinical evaluation, the researcher must be able to make an attribution of causality between the pharmaceutical intervention and the observed outcomes. This hinges on appropriate study design, valid data recording and management and appropriate analysis to control for observed and unobserved confounding factors. The feasibility of any potential source of data for economic evaluation should be assessed prior to its undertaking.

The methods outlined above can be applied to meet the current limitations to generalisability of clinical and economic evaluations, the underlying assumption being that decision-makers really do need this information. An overview of current guidelines for researchers undertaking economic evaluation may give an indication of the extent to which those regulating the conduct and submission of economic evidence pay attention to this issue.

2.3 Economic evaluation guidelines and generalisability

The first section of this chapter stated that clinical decision-makers include physicians, health care providing organisations and governments. Regulatory authorities, whose role frequently is to police the conduct of evaluations and the uptake of new medical technologies, are also an audience for economic evaluation evidence on health care interventions and their guidance may be influential on those conducting such analyses

The primary focus of clinical trials conducted prior to marketing is meeting the requirements of license-granting bodies such as the FDA, EMEA or national regulatory agencies. There is no value-statement inherent in the licensing process concerning the use of a compound in clinical practice. Due to the breakdown of the price mechanism in the health care market, price or profit regulations have been

introduced in most countries to regulate the price paid by the government for pharmaceuticals. (Bloor et al 1996) Also, healthcare providers consider pharmaceutical reimbursement systems to be a crucial part of the healthcare provision. (Freemantle and Bloor 1996, Maynard and Bloor 1997) Decisions on price and reimbursement status of a pharmaceutical are separate from licensing. A challenge to the manufacturer is therefore to persuade a second set of regulators that their compound is worth reimbursement and a given unit price, and the submission of economic data to support these processes is a 'fourth hurdle' to market access. (Freemantle 1999) The uses of economic evaluation information of pharmaceutical therapies include price negotiations, reimbursement negotiations and prescribing decisions. (Johannesson 1995)

The perspective of an economic evaluation, as well as the choice of patient- and context-specific data, may have relevance to the generalisability of an economic evaluation. The perspective of an economic evaluation determines which health outcomes and costs are relevant to the study. For example, a cost-effectiveness analysis (CEA) done from the societal perspective is comprehensive, counting the health effects and costs experienced by all those who are significantly affected by the intervention including the health service, employers, patients and carers. In comparison, a CEA done from the health services perspective would primarily count costs that are directly relevant to the care organisation and may omit costs and outcomes that are not of direct interest to the decision-maker. The appropriate perspective for economic evaluation is generally considered that of the society. The US panel on Cost-Effectiveness in Health and Medicine analysis advised that the societal perspective should be used in all such analyses. (Gold et al 1996) There is however a risk that a study that incorporates a range of resource use applicable to patients and society in one setting may have less scope for transferability to a setting in which patterns and costs of care vary profoundly.

Recently, jurisdictions have introduced the right to formally require economic evaluation data for reimbursement decisions, most notably Australia, the province of Ontario in Canada and in the UK. (Commonwealth Department of Human Services and Health 2000, Canadian Co-ordinating Office for Health Technology Assessment 1997, National Institute for Clinical Excellence 2001) The need for

good quality data and comparable studies has resulted in the publication of guidance on the conduct of economic evaluations, but do these explicitly recognise the need for generalisability of study results?

The Australian guidelines were introduced to aid submissions for new pharmaceuticals to be listed on the government's reimbursement list, the Pharmaceutical Benefits Scheme (PBS). The submissions are reviewed by the Pharmaceutical Benefits Advisory Committee (PBAC). (Hill et al 2000) The guidelines emphasise that "an explicit preference is given for randomised trials over non-randomised studies" for preliminary submissions. However, the guideline acknowledges the limitations of such studies by issuing a section on how to address these limitations through modelling. (Appendix J of the Guideline):

The use of models include "To examine the impact of differences between subjects enrolled in the trials and patients who would be likely to obtain the drug on the PBS and between the settings of the trials and the community setting of the PBS in Australia. Both affect the generalisability of the trials to the PBS context. Important patient factors which may affect outcomes are identified in (c) of Appendix C. There may also be important differences in the mix of patients who will receive the drug on the PBS. Two concerns of the PBAC here are that there may be patients in the community who have disease which is less severe than that of subjects who participated in the randomised trials. There also may be patients in the community for whom the main comparator can be expected to perform better than in the trials. Both could diminish the difference in effectiveness between the proposed drug and main comparator and, therefore, increase the incremental cost-effectiveness ratio. Factors relating to the setting include extrapolating results of trials conducted in hospitals to use outside the hospital and the effect of more rigorous follow-up, which may swamp important differences in the convenience and acceptability of the drug compared with alternative treatments, with resulting effects on patient compliance and thence response to treatment.

The guideline also states that models may also be important to “modify resource use patterns measured in the trials to reflect more closely those in Australia (and/or to add likely changes in resource use patterns not measured in the trials).”

The Canadian guidelines provide assistance for the conduct of studies accompanying applications reimbursement status on the province of Ontario’s reimbursement formulary. (CCOHTA 1995) The Canadian Health Technology Assessment programme has assumed responsibility for the academic contents of the guideline. Guideline 11 of the document concerns “efficacy versus effectiveness” and states:

“Ideally, pharmacoeconomic studies should report on drug effectiveness rather than efficacy. Because effectiveness data are generally not available, appropriate modelling techniques based on sound pharmacoepidemiology (e.g. using epidemiologic studies to estimate patient compliance with therapy in the real world) are permissible. All assumptions used in such extrapolation techniques must be stated explicitly and thoroughly tested in sensitivity analysis.”

In this context, retrospective data represents “viable but not ideal alternative information sources”, and the use of meta-analysis is advocated. For resource use, guideline 19 states:

“In considering international trials it should be noted that resource quantities cannot be directly imported to the Canadian system, because of the major differences in the way that healthcare is delivered in many countries. As a minimum, resource quantities must be re-validated for Canadian practice. Some may, in fact, be transportable into Canada, but an explanation and justification is required. The default assumption is that resource quantities are not directly transportable.”

In the UK, the National Institute for Clinical Excellence was set up in 1999 to appraise clinical and economic evidence for pharmaceuticals assessed for wider

use on the National Health Service (NHS). (Hutton and Maynard 2000) Relating to generalisability of study results, the guidance document states that:

“The settings, populations and methods by which outcomes and costs are measured in the original studies from which the data are drawn should be described and the implications of generalising the data to the NHS in England and Wales explained.”

This brief review of the guidelines for Australia, Canada and the UK reveals that advice to those who conduct and review economic evaluations on generalisability is limited. This is summarised well in the Canadian guideline after a section outlining “portability of economic evaluations” (guideline 24):

“There is no precise process which is recommended at the present time to adapt studies from one jurisdiction to another. It is the responsibility of the investigator to think carefully about the issues discussed above in the planning, interpretation and communication of study results; and it is the responsibility of the “user” to think diligently when using study results in the context of decision-making.”

Philosophically, the guidelines support the use of ‘effectiveness data’ but in practice randomised clinical trial data are preferred. All the guidelines call for setting-specific information and at present modelling is predominantly the tool recommended to provide this. The dilemma faced by decision-makers is that many submitted economic evaluations will have been conducted at phase III in the stage of clinical development of the pharmaceutical compound, where clinical trials tend to be towards the explanatory end of the spectrum of external validity (see section 2.2.1). In future, regulatory bodies such as NICE and the PBAC may require follow-up information on cost-effectiveness to retain a certain reimbursement status or guideline recommendation. To inform such decisions, researchers will need to be equipped with the appropriate tools for increasing the transferability of economic evaluation between clinical settings. Economic evaluation data relevant to the current setting needs to be at hand when therapy recommendations are made, so

an evaluation of the current methodologies available to increase external validity of studies is a research priority.

2.4 *Themes of the thesis*

Economic evaluation is a rigorous, comparative approach with strong links in welfare economic theory. A series of pragmatic compromises are made to balance quality and feasibility when economic evaluation is conducted to inform clinical decision-makers. Economic modelling and the adoption of pragmatic design features in clinical trials are both relevant current methodologies with potential for increasing generalisability, and observational studies may help meet the need for locally applicable data. Against the backdrop of this literature, the need to assess the relative merits of these available methodologies that can be adopted by investigators to increase generalisability of economic evaluations of pharmaceuticals emerged.

Out of this multi-disciplinary review, a number of key themes emerge of particular relevance:

- The use of pragmatic trials in economic evaluation and the impact of study design on effect estimate. (Revicki and Frank 1999, Simon et al 1995, Schulz et al 1995)
- The increasing of generalisability through the use of economic modelling techniques. (Buxton et al 1997, Briggs and Gray 1999, Sculpher et al 2000)
- Applying lessons learned from observational study designs and medical statistics to expand economic evaluation methodologies to generalise findings. (Hennekens and Buring 1987, Hornberger and Wrona 1997, Drummond 1998)

Each of these themes recur through the three source projects of the thesis which are reported in chapter 4 through 6, and in the concluding chapter of the thesis.

CHAPTER 3

EFFECT SIZE IN PRAGMATIC TRIALS: STRIKING THE BALANCE BETWEEN INTERNAL AND EXTERNAL VALIDITY OF TRIALS IN SCHIZOPHRENIA AND POST-MI

3.1 Introduction

Clinical trials adopting a pragmatic design compare treatments under conditions in which they would be applied in clinical practice. (Schwarz and Lellouch 1967) Limited protocol modification of the treatment patterns enables the collection of resource use data reflective of regular practice. Pragmatic trials therefore provide a basis upon which economic evaluations may be conducted, and such trials are increasingly conducted to provide clinical decision-makers with data relevant to clinical practice setting. (Revicki and Frank 1999)

Concerns have been raised that poorly designed and conducted trials, as well as studies with an observational design, may bias effect estimate. (Schulz et al 1995, Ioannidis et al 2001) So it is of interest to examine the design features of pragmatic trials that distinguish them from conventional randomised controlled trials and evaluate whether increased external validity is achieved at the cost of internal validity.

This chapter examines aspects of pragmatic trial methodology in more depth. Firstly, a framework that aims to assess features of pragmatism in randomised controlled trials is developed. Secondly, methods used to investigate the impact of trial quality on effect size are applied to a sample of trials with both pragmatic and explanatory features to investigate the impact of trial design on effect size. Finally, the results are discussed to determine the merits of pragmatic trials for provision of externally valid results without loss of internal validity, a finding that would support adoption of such trials for economic evaluation.

3.2 Background

The basic function of clinical research is to compare the outcome of different treatments. Researchers may exercise a varying degree of control over the circumstances in which the comparison is undertaken, ranging from 'laboratory' to 'clinical practice' conditions. Schwartz and Lellouch (1967) proposed the terminology 'explanatory' and 'pragmatic' attitudes to randomised controlled trials (RCTs), respectively, to characterise the two extremes. For example, an explanatory attitude would be delivery of two treatment strategies under equalised conditions, while a pragmatic attitude would be delivery of two treatment strategies under conditions where the physician rather than the protocol determines the conditions. Furthermore, trials designed with explanatory attitudes may assess results by means of intermediate endpoints, such as regression of cancer tumour, or a functional disability such as pain levels. A pragmatic design would assess results by a measure of practical importance, such as mortality. Other design attitudes that Schwarz and Lellouch judged pragmatic include comparison with active substance rather than placebo, selection of a broad range of suitable patients rather than restrictive inclusion criteria, and analysis of patients according to the intervention to which they were assigned, regardless of whether they dropped out or swapped treatment during the course of the trial. The authors conclude that an explanatory comparison of treatments seeks to verify a biological hypothesis and understand the differences between two treatments, whereas a pragmatic comparison seeks to inform a treatment decision on the basis of "which of the two treatments should we prefer?"

Explanatory RCTs run the risk of providing a low external validity to general clinical practice because of subject and centre exclusions, non-participation and blinding of the study participants to the treatment allocation inherent in the trial protocol. Clinical decision-makers face heterogeneous populations and conditions that may differ from those in a RCT setting, so adopting pragmatic study design features, such as expanding the inclusion criteria, applying a multicentre design and encouraging practitioners to invite all eligible patients, provide solutions to this problem. Also, avoiding the blinding of the participants and leaving some control of the therapeutic regimen to the clinician and the patient would reflect the circumstances under which a drug will be used in

clinical practice. (Freemantle and Drummond 1997)

Evaluation of an intervention in a naturalistic setting enables collection of resource data that reflects a clinical practice pattern. Pragmatic trials provide particularly good basis for economic evaluation of resource consumption as well as clinical outcomes. (Revicki and Franck 1999) For example, Simon et al (1996) evaluated effectiveness and cost of two treatment strategies in depression, fluoxetine versus tricyclic antidepressants, using an RCT in which inclusion criteria were broad, management of care was determined by the physicians, neither patients nor physicians were blinded, and the data were analysed according to original treatment allocation. Similarly, Oster et al (1995) analysed costs and effectiveness of cholesterol-lowering drugs in the CRIS trial, which applied wide inclusion criteria, minimised intervention by leaving management decisions to the physician and measured number of 'successfully treated patients', i.e. patients that achieved blood cholesterol below a certain guideline threshold, by intention to treat analysis.

Pragmatic trials relax many design features to increase external validity, but maintain the element of randomisation. The process of randomly allocating treatment to patient groups is a key trial design feature aiming to maximise trial validity, and randomisation is incorporated in checklists evaluating clinical trial quality. (Chalmers et al 1981, Colditz et al 1989, Moher et al 1996, Moher et al 2001) Several research teams have sought empirical evidence of the impact of randomisation on effect size in clinical research, but findings are inconclusive (see section 2.2.1). (Concato et al 2000, Ioannidis et al 2001) Similarly, researchers have suggested that randomised clinical trials where quality is 'poor' according to key criteria may bias effect estimates. Schulz et al (1995) analysed the methodological quality of 250 controlled trials from 33 meta-analyses and found a 30% larger effect size where treatment allocation was not appropriately concealed. Moher et al (1998) analysed 127 trials from 11 meta-analyses and found a 37% larger estimate of benefit in trials that used inadequate allocation concealment.

As we have seen, pragmatic trials randomise patients, but adopt design features that make them more reflective of clinical practice. Is there a risk that trials that are 'contaminated by reality' may bias effect estimates? If this is the

case, then increased external validity would be achieved at the cost of low internal validity of such studies. There is little empirical evidence that pragmatic design features influence the estimate of effect. In the analysis reported in this chapter the design features of pragmatic trials were examined, which distinguish them from explanatory randomised controlled trials. First, a checklist of design features characteristic of a pragmatic trial was developed and applied to 318 trials included in 19 meta-analyses in two therapeutic areas. Generalised linear models were fitted to systematically assess whether design features that characterise pragmatic trials impact their outcome. The methods used were derived from those employed by Schulz et al (1995) and subsequently by Moher et al (1998), amongst others.

3.3 Objectives

The objectives of this project were:

- To identify and describe pragmatic design features that trial investigators may adopt to increase external validity of a randomised clinical trial.
- To assess whether pragmatic design features alone, or in any combination, provide effect estimates systematically different from those provided by explanatory design features.

3.4 Methods

3.4.1 Development of a checklist of pragmatic design features

Authors have frequently used a dichotomised trial definition of ‘explanatory’ versus ‘pragmatic trials’ (Roland and Torgerson 1998, Revicki and Frank 1999) in spite of the fact that Schwartz and Lellouch (1967) concluded their original work with “most real problems contain both explanatory and pragmatic elements”. The original authors therefore view pragmatism on a continuum,

where, to a varying extent, trials adopt design features that make them more or less reflective of clinical practice.

A checklist was derived for the purpose of this review, to identify pragmatic design features in clinical trials, the criteria for which were developed based on literature describing attributes of pragmatic trials in different therapy areas. The checklist aimed to be generic and covered different dimensions of trial design, such as description, inclusion criteria, treatment, assessment and analysis. Design features that were considered, *a priori*, the most important components of a pragmatic trial are listed in *Table 3.1*, which comprises the checklist.

Trial description

Pragmatic trials are often designed as long term trials to enable measurement of final clinical endpoints. In this analysis, trials of 12 or more months duration were characterised as long-term. Large trials may be considered pragmatic as these have a higher potential to include a range of patients. The size of 300 was chosen as a cut-off for the purpose of the analysis. The sensitivity of the analysis to a cut-off of the median, 200, 400, 500 and 1000 patients was also assessed. Resource consumption data are sometimes collected alongside pragmatic trials, so the recording and subsequent costing of outcomes such as hospitalisations and concomitant drug use were classified as a pragmatic feature.

Since human behaviour is influenced by what is known or believed, knowledge of the treatment regimen may affect patient or physician behaviour and assessment of treatment outcome. In order to reduce such bias, patients and physicians are frequently blinded to treatment allocation in clinical trials, (Day and Altman 2000) but in clinical practice, neither patients nor physicians are ignorant of the administered treatment, so pragmatic trials evaluating treatment in real-world settings are sometimes conducted unblinded (open). (Freemantle and Drummond 1997, Buxton et al. 1997)

Inclusion criteria

Explanatory trials may include selected patients with a greater capacity to benefit from treatment, and elderly patients may frequently be excluded from clinical trials. (Wenger 1992) It is difficult to make an age cut-off a generic measure, as patient groups targeted by different treatments vary. This analysis used patients over 70 years old as a cut-off point. Explanatory trials are often conducted in specialist settings under expert care, for example in hospital. Trials that are undertaken in regular clinical practice would include outpatients, thereby increasing the external validity of the sample, so the inclusion of outpatients was considered a pragmatic trial feature. Also, multi-centre trials are more likely to reflect a diversity of treatment provision, resulting in higher external validity. Multi-centre design was therefore classified as a pragmatic feature.

Treatment

Regulatory requirements for trials designed to demonstrate the efficacy of an intervention include comparison to inert substance, or placebo. In clinical practice, however, the treating physician is faced with a choice of no therapy or an alternative treatment strategy. Trials aiming to increase external validity of an evaluation may compare the intervention under study with an active substance, such as that considered the most frequently used in the relevant clinical setting. (Buxton et al 1997) Monitoring of patient compliance is often incorporated in the trial protocol of randomised clinical trials and may include interview during frequent physician visits, serum level measurements or pill-counts. To increase external validity, a reduction in such intervention may be made in order to reflect a clinical practice setting, (Buxton et al. 1997) and minimal monitoring of compliance was considered a pragmatic design feature in this analysis.

Analysis

'Intention to treat' is an analytical strategy used in RCTs comparing patients in the groups to which they were originally assigned, regardless of whether they did not receive treatment, stopped taking medication or changed treatment

during the course of the trial. A trial adopting ‘intention to treat’ analysis evaluates the strategy of one treatment compared with another, rather than the outcome of an isolated intervention and is considered a pragmatic trial feature. (Schwarz and Lellouch, 1967) This design principle was first deemed ‘pragmatic’ by Schwarz and Lellouch and since then has become a common feature of trial methodology seeking an unbiased solution to treatment switching or attrition (patients in a control group ultimately receiving treatment intended for the intervention group) or where patients do not actually receive the treatment intended. Hollis and Campbell (1999) found that there is no single definition of an intention to treat analysis and that the approach is often both inadequately described and applied. An in-depth investigation of the quality of the intention to treat analysis was beyond the scope of this study. Therefore, trials that explicitly reported the use of intention to treat analysis were recorded as having adopted this pragmatic design feature.

Table 3.1 A checklist of characteristics of a pragmatic design.

Trial design aspect	Pragmatic design feature	Score
Trial description	The trial is a long term trial (> 12 months)	(Y/N)
	The trial is large (≥ 300)	(Y/N)
	Trial presents resource consumption data	(Y/N)
	Trial is single blind or open	(Y/N)
Inclusion criteria	Patients over 70 are included	(Y/N)
	Trial included outpatients	(Y/N)
	The trial is multi-centre	(Y/N)
Treatment	Comparator is active substance	(Y/N)
	Patient adherence is not monitored	(Y/N)
Analysis	Intention to treat analysis was used	(Y/N)

3.4.2 Material

The checklist in *table 3.1* was applied to 318 randomised controlled trials identified in two systematic reviews in two therapy areas (*Table 3.2*). These were published as Clinical Practice Guidelines in the therapy areas of schizophrenia and secondary prevention of myocardial infarction. (Geddes et al.

2000, Eccles et al 1998, Data on File) Identification of trials for inclusion in these reviews followed systematic reviewing methodology, searching the electronic databases MEDLINE, EMBASE, SIGLE and the Cochrane Controlled Trial Register as well as the grey literature for unpublished trials. Trials were included in the reviews if either published or unpublished by December 1st 1998. These systematic reviews conducted meta-analyses at the level of therapy, for each therapy considered. *Table 3.3* provides an overview of the therapies included, all of which were reviewed by the North of England Clinical Practice Guideline Projects. The majority of the studies evaluated pharmaceutical interventions, although diet and cardiac rehabilitation were also included. A full reference list to all trials included in the review is provided in appendix 1.

Table 3.2 Contents of the systematic reviews

Therapeutic area	Meta-analyses	Trials	Patients
Schizophrenia	7	52	12 649
Post - MI	12	266	248 717
Total	19	318	261366

Table 3.3 Meta-analyses included in the therapy areas.

Therapy area	Meta analyses
Scizophrenia	Amisulpride, Clozapine, Olanzapine, Quetiapine, Risperidone, Sertindole, and Alternative atypical antipsychotic drugs
Post-MI	Cholesterol lowering statins, Beta-blockers (short term), Beta-blockers (long term), Beta-blockers (+conventional), ACE inhibitors (Unselected patients), ACE inhibitors (Selected patients), Calcium channel blocker (Myocardial infarction), Calcium channel blocker (Heart failure), Nitrates, Dietary measures, and Antiplatelet agents

Some information that was of interest to this analysis had already been extracted for the original guideline reports (*Table 3.4*).

A key aspect of trial quality is concealment of treatment allocation. (Schulz et al 1995, Moher et al 1998) The strength of the randomisation relies on the degree to which the procedure conceals the allocation sequence and therefore prevents participants in the trial from including patients on the basis of the treatment next in the sequence. (Altman and Schulz 2001) This quality feature was included in

the analysis in order to analyse the impact of allocation concealment on a different subset of trials from those evaluated by the previous authors. Data extraction for the two original reviews involved double checking and discussion between two independent assessors to resolve disagreements. A separate extraction of these data for the current analysis would have duplicated research effort without any significant contribution to precision, so was not undertaken.

Primary endpoint differed between the trials in the review. Primary outcomes from the post-MI trials included neuroendocrine activity, blood pressure, heart rate and mortality. The trials in schizophrenia recorded a range of outcomes, including the disease-specific scores to measure psychopathology and schizophrenic symptoms (Brief Psychiatric Rating Scale [BPRS] and the Positive and Negative Syndrome Scale [PANSS]). Binary events of mortality in trials in MI and drop-out in trials in schizophrenia were chosen for the purpose of this analysis, and since both outcomes are undesirable in their respective therapy area, effect of treatment would aim to reduce the incidence of each (table 3.4).

Table 3.4 Data extracted in the existing systematic reviews of trials in post-MI and Schizophrenia.

Extracted information	Variable
Authorship details	Author / Research group and year
Comparisons	Intervention drug and control drug
Double-blind	Y/N
Allocation was reported to be concealed	Y/N
Number of patients in trial	n
Loss to follow-up	%
Follow up (duration)	Year
Events (mortality or drop out)	n

3.4.3 Statistical analysis

Microsoft Excel was used for data management. The trials were retrieved and data extracted according to the checklist in *Table 3.1* and the information in *Table 3.4*. The Excel file was then transformed into a SAS statistical system file and SAS v 8 was used for modelling. Generalised linear modelling is adopted in this chapter. (McCullagh and Nelder 1989)

Definition of outcome

All cause mortality was available from all of the trials in post-MI, as were dropout rates from all trials in schizophrenia. The Odds Ratio (OR) of effect was based on these two binary trial outcomes and used as outcome in the statistical analysis. Effect estimates are calculated on the basis of prevention of an adverse outcome (a reduction of events). The greater the reduction in events as a result of treatment in a trial the larger the treatment effect estimated by that trial. The summary OR in all trials was calculated and an OR of less than 1.0 indicated that treatment intervention was more effective than the controls overall.

Interaction terms between each pragmatic trial feature and treatment were then fitted. Results of the analyses are reported in terms of a ratio of ORs (ROR). By the modelling convention adopted here, explanatory trial features are the reference case. A ROR of less than 1.0 for an interaction term indicates that trials adopting pragmatic features yielded larger estimates of treatment effect, on average, than the reference group of trials adopting the corresponding explanatory feature. Conversely, a ROR greater than 1.0 for an interaction term indicates that pragmatic trial features yielded a smaller estimate of treatment effect than explanatory trials. By way of illustration, trials that adopted one particular pragmatic feature on average estimated an effect size of 0.9, and trials that adopted the corresponding explanatory design feature on average estimated an effect size of 0.8 (i.e. reduction of odds of events by 10% and 20% respectively). The RORs would be $0.9/0.8 = 1.125$, indicating that the ORs in the pragmatic trials are on average 12.5% higher than that in the explanatory trials. Since treatment effect is defined as reduction in the odds of an event,

this estimates a 12.5% smaller effect size in the trials adopting the pragmatic feature.

Fitting models

The specific methods used for this project were derived from those first developed by Schulz et al (1995) and subsequently used by Moher et al (1998). The general approach to model development was described by McCullagh and Nelder (1989). Generalised linear models with logit link and binomial error were fitted to assess whether the trials applying one or more of the design features in *Table 3.1* differ systematically from those that do not in predicting power of the ratio of OR (ROR) as estimated in the meta-analyses. (See section 5.1.1 of this thesis for an outline of generalised linear modelling.) The models were fitted using PROC GENMOD. They included specified hierarchical factors reflecting the structure of the data, where a trial feature is shared by trials included in different meta-analyses in different therapy areas, so take into account the measurement variability within each level (analogous to multilevel models). These models accounted for the effect of treatment, trial, meta-analysis and clinical area (post MI or schizophrenia) by the fitting of appropriate classification variables.

The interaction effects between treatment and each factor indicating pragmatic design feature were of primary interest. The main effects of all the indicators of pragmatism were explored in conjunction with the interaction with treatment and factors significant at conventional statistical level $p < 0.05$ were then combined in a model. Interaction effects between significant variables were also explored. The combination of factors that provided the lowest residual deviation was selected for the final model.

Overdispersion

Odds ratios (OR) are themselves estimates and subject to measurement error. The estimates generated from trials conducted in selected cohorts will reflect the distribution of the underlying treatment effect in the whole population with the condition. There is reason to believe that such data will be overdispersed (i.e. have extra-binomial variability) and the final model will be a poor fit. Model

fit can be assessed by investigating the value of the scale factor (standard deviance divided by residual degrees of freedom on the appropriate strata). If this is close to 1 the model fits well, however if the value is high then overdispersion should be addressed. PROC GENMOD does not enable fitting of random effects models and therefore this analysis accounted for extra binomial variability (overdispersion) by inflating the scale parameter by the mean residual deviance on the appropriate stratum. (McCullagh & Nelder 1989)

3.5 Results

Eighty-two (25.9%) trials were long term trials of more than 12 months duration. Resource consumption was reported in 23 (7.1%) trials (table 3.5). Eighty-nine (28%) trials were single blind or open trials. Patient adherence was not reported monitored in as much as 66.3% of the trials. About two thirds of the trials included outpatients in primary care.

Table 3.5 Characteristics of trials (total number of trials 318)

Design feature	Trials (%) with pragmatic feature
The trial is a long term trial (> 12 months) (Y/N)	82 (25.9)
The trial is large (≥300) (Y/N)	100 (31.4)
Trial presents resource consumption data (Y/N)	23 (7.1)
Trial is single blind or open (Y/N)	88 (27.8)
Patients over 70 are included (Y/N)	99 (31.1)
Trial included outpatients (Y/N)	219 (64.1)
The trial is multi-centre (Y/N)	97 (30.4)
Comparator is active substance (Y/N)	147 (46.3)
Patient adherence is not monitored (Y/N)	211 (66.3)
Intention to treat analysis reported (Y/N)	90 (28.2)
Allocation was NOT concealed or unclear (Y/N)	267 (84.1)

Overall, treatment was associated with a 13% reduction in the odds of an event versus controls (summary Odds Ratio = 0.87, 95% CI = 0.82 to 0.93, $p < 0.0001$). In other words, mortality / drop-out was on average 13% lower in

the treatment arms of the trials, estimating an overall treatment effect of 13%. The relationship between treatment and events may be different for trials in schizophrenia and post-MI, and this was explored by examining the value of the interaction between treatment and therapy area. The ROR for trials in the area of post-MI compared to those in schizophrenia was 0.55 (0.41 to 0.73, $p < 0.0001$), that is, trials in post-MI predicted on average 45% larger effect size than trials in schizophrenia.

The associations between pragmatic design features and treatment effect were then explored individually for each pragmatic feature (*Table 3.6*). All models accounted for the effect of treatment, trials, therapeutic area, meta-analysis and the interaction effect between treatment and therapy area. Extra-binomial variation was accounted for by inflating the standard error by the root of the scale factor.

The effects estimate in those trials that applied pragmatic design features including multi-centre design, collection of resource data and inclusion of patients aged over 70 did not vary more than would be expected by chance. Trials with a sample size of over 300 patients, were single blind or open, or included outpatients, appeared to bias the effect estimates (*Table 3.6*). Large trials (>300) provided, on average, an effect estimate 80% smaller than small trials (ROR = 1.80). In order to assess the sensitivity of the analysis to the choice of trial size as cut-off point, the sizes 200, 400 and 1000 were explored. For cut-off size of 200, the ROR was 1.54 (1.55 to 1.22, $p = 0.003$), for cut-off size of 400 it was 1.44 (1.18 to 1.46, $p = 0.0003$), and for cut-off size of 1000 it was 1.58 (1.33 to 1.88, $p < 0.0001$). The analysis of large trials was not sensitive to change in trial size outside of the range 200 – 1000 patients.

Trials that included outpatients provided an effect estimate 24% larger than trials that only included inpatients. Trials that were single-blind or open provided an effect estimate that was 23% smaller than trials that were double-blind. Trials that reported to have used intention-to-treat analysis provided more conservative effect estimates, with a 19% lower effect estimate than those applying alternative statistical analyses. Those trials that did not report an adequate concealment of allocation reported 15% higher effect estimates.

Table 3.6. Association between indicators of individual pragmatic trial features and estimates of treatment effect. Ratio of odds ratio for the interaction effects of the trial design features with treatment.

Pragmatic design feature	ROR (95% CI)	P - value
The trial is a long term trial (> 12 months)	0.65 (0.38 to 1.12)	0.12
The trial is large (≥ 300)*	1.80 (1.47 to 2.21)	<0.0001
Trial presents resource consumption data	1.02 (0.89 to 1.17)	0.78
Trial is single blind or open*	1.23 (1.08 to 1.40)	0.002
Patients over 70 are included	1.08 (0.93 to 1.24)	0.31
Trial included outpatients*	0.76 (0.66 to 0.88)	0.0002
The trial is multi-centre	1.09 (0.90 to 1.31)	0.39
Comparator is active substance	0.97 (0.83 to 1.15)	0.75
Patient adherence is not monitored	1.06 (0.93 to 1.20)	0.39
Intention to treat analysis was reported*	1.19 (1.02 to 1.38)	0.03
Concealment of allocation not reported*	0.85 (0.75 to 0.97)	0.017

* significant at conventional 0.05 level

The interaction between trial size and double-blinding was of interest to the analysis (*Table 3.7*). The interaction term was fitted by adding new variables to the model that were the product of treatment and size, of treatment and double-blinding and of treatment, double-blinding and size. Twenty trials were large single-blinded or open trials. The RORs for these trials compared was 0.46 (0.31 to 0.69, $p < 0.0001$), which meant that the odds ratios in the large open trials were, on average, 54% smaller than in small open trials and large double-blind trials. The large open trials therefore estimated 54% larger treatment effects. The interaction effect between trial size and blinding status varied across the four possible interactions (*Table 3.8*).

There was no interaction effect between trial size and the inclusion of outpatients (ROR = 0.99 (0.64 to 1.55), $p = 0.97$). Neither was there an interaction effect between open-trial design and the inclusion of outpatients. (ROR = 0.90 (0.62 to 1.31), $p = 0.57$).

Those pragmatic trial features with a significant interaction effect with treatment were included in the final model, along with the relevant significant second-order interactions. The trial features of intention-to-treat analysis and treatment allocation not concealed did not bias effect estimates when included with the

other significant variables (*Table 3.9*). The final model included therapy area interaction with treatment, outpatient interaction with treatment, large-trial interaction with treatment, open-trial interaction with treatment and the second-order interaction between treatment, open-trial and large-trial (*Table 3.10*).

*Table 3.7 Interaction effect between blinding and size**

Design feature	ROR (95% CI)	P-value
The trial is large (≥ 300)	2.37 (1.83 to 3.07)	<0.0001
Trial is single blind or open	2.38 (1.62 to 3.49)	<0.0001
Interaction open / large	0.46 (0.31 to 0.69)	0.0001

*In the presence of interaction treatment and therapy area, ROR = 0.41 (0.31 – 0.54) $p < 0.0001$

Table 3.8 Interaction effect between different attributes of blinding and trial size

ROR (95% CI)	Small (<300)	Large (>300)
Double-blind	0.46 (0.31 to 0.69) P = 0.0002	2.17 (1.45 to 3.25) p = 0.0002
Open or single-blind	2.18 (1.46 to 3.26) P = 0.0001	0.46 (0.31 to 0.69) p = 0.0001

Table 3.9 Model of all variables that individually biased effect estimate in randomised controlled trials

Design feature	ROR (95% CI)	p-value
The trial is post-MI	0.39 (0.29 to 0.52)	<0.0001
Intention to treat analysis was reported	0.90 (0.74 to 1.09)	0.27
Concealment of allocation not reported	0.91 (0.79 to 1.06)	0.24
Trial included outpatients	0.78 (0.65 to 0.94)	0.008
The trial is large (≥ 300)	2.37 (1.82 to 3.10)	<0.0001
Trial is single blind or open	2.40 (1.64 to 3.52)	<0.0001
Interaction open / size	0.39 (0.26 to 0.60)	<0.0001

Table 3.10 Final model of pragmatic design features that in combination biased effect estimate in randomised controlled trials

Design feature	ROR (95% CI)	p-value
The trial is post-MI	0.38 (0.28 to 0.50)	<0.0001
Trial included outpatients	0.80 (0.67 to 0.96)	0.015
The trial is large (≥ 300)	2.35 (1.82 to 3.04)	<0.0001
Trial is single blind or open	2.42 (1.65 to 3.54)	<0.0001
Interaction open / size	0.40 (0.26 to 0.60)	<0.0001

3.6 Discussion

Pragmatic trials provide a method for clinical experiment reflecting clinical practice setting. They have been viewed as good vehicles for economic evaluations that seek to evaluate cost-effectiveness in clinical practice for a number of reasons, including their ability to provide resource-consumption data reflective of regular treatment pattern. (Revicki and Frank 1999) The question of clinical trial quality and design continues to be one of great importance to clinical medicine. Empirical research has provided inconclusive evidence as to whether non-randomised, observational studies provide biased effect estimates. The risk of pragmatic trials being ‘contaminated by reality’ (Freemantle and Drummond 1997) and so effecting this same bias was examined in this project.

Methods and main findings

Schultz et al (1995) developed the methodology to investigate the impact of trial quality on effect estimate using multiple regression models. This was replicated to explore the impact of ‘pragmatic’ design features on effect size relative to the corresponding ‘explanatory’ design features. The difference in odds ratios estimated by each design was expressed through the ratio of odds ratio (ROR), which expresses a difference in relative odds of effect rather than absolute odds of effect. The analysis approach adopted here proved valid in previous analyses, accounting for confounding by therapy area by including the interaction term between treatment and area. It also went some way towards controlling for overdispersion by inflating by the scale factor.

Schultz et al (1995) explored the impact of quality features such as concealment of allocation and double-blinding and found that trials that did not adequately conceal treatment allocation provided effect estimates on average 30% larger than trials where treatment allocation was concealed. This finding was confirmed by Moher et al (1998), who found that inadequate allocation concealment was associated with an increased estimate of benefit of 37% compared with trials that used adequate methods. In addition, Schulz et al. found that trials that were not double-blind exaggerated effect size by an average of 17% compared with trials that were double-blind.

The analysis presented in this chapter used a similar multiple regression analysis approach to study the impact of pragmatic design feature on effect size. The only features that overlapped Schulz et al's study was open trial design and lack of concealment of allocation. Results indicated that single-blind or open trials *reduced* the effect estimate by 23% in contrast to the findings of Schulz et al, but in line with their findings, lack of adequate reporting of allocation concealment exaggerated effect size significantly by 15%. This effect did however disappear in the presence of the interaction effect between study size and blinding. This may indicate that study blinding confounds the quality indicator variable of concealment of allocation: an open study does not need to conceal allocation.

Other features of interest to this analysis were primarily related to level of pragmatism in trials. The majority of pragmatic design features did not produce effect estimates systematically different from the corresponding 'explanatory' feature. Long-term and multi-centre trials, trials where an active substance was used as a comparator, or where patient adherence was not monitored failed to provide systematically different effect sizes from those adopting the corresponding 'explanatory' design feature. This analysis does therefore not provide disincentives to those wishing to relax such trial features to achieve an environment for the experiment more reflective of clinical practice.

Trial size (>300), inclusion of outpatients and the reporting of intention-to-treat analysis were all features that *did* produce different effect estimates. Trials that included more than 300 patients in total averaged an effect estimate that was less than half that of trials of total size below 300. A possible explanation for

this may be that trials including more than 300 patient will usually require the contribution of a large number of centres. This might impact upon the level of organisation required (which may itself reduce bias) and dilute the effect of 'enthusiastic' centres that may provide large estimates unrepresentative of treatment effect. In contrast, those large trials that were single-blind or open-label trials reported an average treatment effect that was exaggerated by 60%. Conversely, in the small, double-blind trials included in the analyses, there was an average relative reduction in the odds of an event of 60%, and these trials exaggerated effect size by 60%. Is this empirical evidence of publication bias, indicating that small trials have a higher likelihood of being published if they are positive rather than neutral or negative to the intervention? The results of this analysis are in line with work by Ioannidis et al (1998), who identified that smaller trials may provide more optimistic effect estimates than large trials.

Trials that included outpatients estimated effect size of 20% larger than those that only included inpatients (hospitalised patients). Patients living in the community may be less severely ill than those that are hospitalised, so the inclusion of outpatients that may have a greater potential to benefit from treatment could result in more optimistic results for these trials. In contrast, the use of intention-to-treat analysis may provide more conservative estimates of effect, since by definition this analysis includes all randomised cases in the final analysis, regardless of whether the patients swapped treatment, were lost to follow-up or dropped out completely. The adoption of intention-to-treat analysis was based on whether the trial reports had claimed the use of the analysis, rather than through scrutiny of the methods. Schulz et al (1995) excluded the aspect of intention-to-treat analysis from their trial assessment because of lack of inter-assessor reproducibility on this variable. Indeed, Hollis and Campbell (1999) found that only 55% of trials that stated that they use intention-to-treat analysis in fact analyse the data according to treatment allocated during randomisation. This may have introduced a bias to the analysis, as the analysis is evaluating reporting rather than actual analysis. It is however difficult to assess the direction in which this potential bias would have influenced the results.

A checklist of pragmatism

A series of decisions taken by the trial investigators place the trial somewhere on the continuum between the ideal type explanatory and the ideal type pragmatic trial. The features on the checklist enabled scoring according to criteria determined 'pragmatic' or 'explanatory'. The underlying view conveyed here is that trials may adopt one or more of these design features and therefore achieve a varying degree of pragmatism.

Is the devised checklist appropriate for assessing pragmatism in other clinical trials? The checklist was generic and could be used to score trials according to pragmatism, but there were a number of difficulties in scoring the trials this way. For example, the judgement of many features is limited by the information provided in published trials. Even though the adoption of a certain feature, say monitoring of patient compliance, is not reported, it may well be a characteristic of a trial. Lack of reporting of such features would have confounded the analysis.

Furthermore, what is considered a large trial may vary by therapeutic area and indeed by the frequency of the outcome being measured and the resulting power to detect differences at conventional levels of certainty. For example, a trial of size 300 with a dichotomous outcome of 1% may be a small trial, whereas a trial with a sample size of 300 with an outcome of, say, 40% may be considered a large trial.

There is no absolute definition of what constitutes a large trial, and separating small trials from large is not straightforward. The initial intent was to use the median as the cut-off, but this was strongly influenced by the large number of very small studies in the cardiovascular area. Cappelleri (1996) used a rule based on sample size of more than 1000 patients qualifying as a large trial. As a pragmatic decision the size 300 was determined as the cut-off in this analysis. Sensitivity analysis suggested that the association between size of trial and average effect size was not dependent upon the definition small above a trial size of 200, and this may be because of the large number of small trials (<150 patients) in the post-MI meta-analyses.

Pragmatic trials may adopt other features that were not considered feasible or appropriate to include in this review. For example, pragmatic trials may assess the impact of treatment on a relevant choice of final endpoint, while explanatory trials often use intermediate endpoints. It was not possible to define a generic pragmatic outcome for clinical trials other than mortality, and mortality is not a final outcome of direct relevance in therapy areas such as mental health. Also, pragmatic trials frequently aim to include patients that are representative of the population that will use the medication when it reaches general clinical practice. Characteristics such as severity of disease, gender distribution of the patient sample and age distribution are examples of characteristics that may determine the degree of transferability of trial results between populations. Mortality from cardiovascular disease does, for example, differ between men and women and in different age groups. Similarly, the inclusion of remittent patients in trials in schizophrenia may impact the potential of the sample to demonstrate benefit from the treatment, as such patients are likely to be treatment resistant. It was difficult to develop generic indicators of pragmatism based on these, so they were not included in the checklist.

Similarly, explanatory trials frequently fix the dose of the drug under assessment or titrate and adjust the dose up to a fixed level. Pragmatic trials may leave decision of dose to the treating physician, allowing a flexible dose range. This was a difficult distinction to make, so the adoption of a flexible dose range was not included in the analysis of pragmatism.

Limitations of the study

This analysis was based on a large number of trials and patients from comprehensive literature reviews in two therapy areas, and included more trials than previous authors (318 vs 250 in Schulz et al 1996). However, the study has limitations. Key features that may have introduced bias include the selection of trials and subset of meta-analyses, the choice of binary outcome and the fact that data extraction was undertaken by one researcher only for the purpose of consistency. Firstly, only trials in one therapeutic area were selected, and these had already undergone a 'quality-check' for inclusion in the meta-analyses of the clinical guideline. This may have excluded some relevant trials. Secondly, post-MI trials constitute the vast majority of the trials included in this analysis. A powerful area-effect was detected in terms of effect size, in

that trials in Post-MI on average predicted 45% larger effect size. The outcome of mortality recorded in post-MI trials is objective, but drop out-rate in trials in schizophrenia may not be objective, as the definition of 'drop out' may vary between trials. The choice of outcome for the analysis may have introduced a confounding by area. Finally, decision on several of the pragmatic design features included a judgement on the part of the researcher. For example, some trials did not explicitly state that outpatients were included, and judgement of anecdotal information in the trial reports had to be made.

Because of these limitations, the findings may not automatically be extrapolated to other therapeutic areas, or to other trials adopting pragmatic features.

3.7 Conclusions and recommendations

Some of the results provided in this report were consistent with previous findings, such as the lack of concealment of allocation may bias trial results. However, this effect disappeared in the presence of open-label and large trials, as the latter are more predictive of treatment effect. Evaluation identified a tendency for small double-blind trials to systematically provide more optimistic estimate of effectiveness. It would not be a rational response to the data presented here to abandon meta-analyses of small trials because of the risk of publication bias, or the use of small trials in economic evaluation. But where only small trials are available for meta-analysis and economic evaluations, some caution should be exercised in the interpretation of the results. Meta-analysis of small trials may not be considered sufficient for health policy, although it may provide very useful data to support design of future large-scale randomised trials to evaluate clinical outcomes and resource consumption. The decision to apply the results of meta-analysis of small-scale clinical trials to individual patients is more difficult. Patients, particularly those with life threatening conditions, may only have a single opportunity to benefit from treatment, and meta-analysis of small trials may still provide the best available estimate of treatment effect compared with the alternatives.

The literature does not present any similar analysis to this and judgement of generalisability of findings to other therapy areas is difficult. These results cannot be directly attributed to the treatment areas in the meta-analyses, as this was accounted for through the area treatment interaction. However, other therapy areas may have other traditions for trial design and conduct. For example, small trials may be over-represented in the cardiovascular area, and may provide different associations between design features and effect estimates. Similar analyses of pragmatic design features should be replicated in those therapy areas where extrapolation of these findings does not seem appropriate.

Pragmatic trials adopt design features that make them more reflective of clinical practice and such trials have been judged feasible vehicles for economic evaluation. To ensure that the effects of the intervention are captured by clinical trials, long term follow-up of patients is essential. The cost of RCTs, however, increases with the length of follow-up and this can discourage funding bodies from undertaking pragmatic trials. They may be expensive, lengthy and the data analysis may be difficult due to a tendency of crossover between treatments. (Simon et al 1995) However, pragmatic studies could 'bridge the gap' between the insufficiency of RCTs and the issue of bias in observational studies. The findings of the meta-regression of pragmatic trials presented in this chapter did not provide empirical evidence of bias in pragmatic trial design. Basing an economic evaluation on pragmatic trials increases the external validity of the evaluation without jeopardising the validity of the effect estimate, so economic evaluations should be incorporated in large pragmatic trials.

CHAPTER 4

THE MERITS OF ECONOMIC MODELLING AS A MEANS OF GENERALISING COST-EFFECTIVENESS ESTIMATES: ASSESSMENT OF ECONOMIC MODELS IN OSTEOPOROSIS.

4.1 Introduction

Concerns about the generalisability in economic evaluation of pharmaceuticals across healthcare settings has led to a search for methods to accommodate for setting-specific characteristics, increasing their relevance to local healthcare decision-makers.

Economic models may have the potential to vary parameter estimates over a range of values reflecting both healthcare management systems and clinical characteristics in variety of settings. There is currently little empirical evidence examining the methods by which models may increase generalisability, or indeed whether current methods adopt such methods.

In this chapter, economic models are examined for the degree to which they attempt to accommodate for setting-specific differences and present their findings in a manner that increases their generalisability. A checklist incorporating dimensions of the modelling process that could increase external validity was applied to economic models published in the therapy area of osteoporosis. The merits of economic modelling as a means of increasing generalisability of cost-effectiveness analyses are discussed in view of the findings.

4.2 Background

Decision analytic modelling is widely used in assessing the most cost-effective intervention from mutually exclusive alternatives. In health economic evaluation, models are predominantly used to structure decision problems, to inform treatment decisions under uncertainty. The use of models as vehicles for cost-effectiveness analysis has been extensively described in the literature (Weinstein and Fineberg 1980, Gold et al

1996, Briggs and Schulpher 1998, Buxton et al 1997) and the general advantages and limitations of modelling in healthcare decision-making highlighted by several authors. Sheldon (1996) outlined a number of problems that make models susceptible to bias, including model framing, construction, the reliability of estimates included and the way in which sensitivity analysis is performed. In contrast, based on experience with health technology assessment in the UK National Health Service, Brennan and Akehurst (2000) advocated the following uses of economic models in decision-making: extending results from single trials; combining multiple sources of evidence; generalising results from one specific context to another; informing future research strategies and exploring uncertainty in current knowledge base.

The inclusion of data to populate transition probabilities for clinical and economic events is a key component of economic modelling. Data for economic models are often synthesised from several sources. (Nuijten 1998) Resource use data may be collected within the setting of a clinical trial or in databases where such use is routinely recorded. Estimates of disease prevalence and incidence are frequently based on observational data and assumptions of treatment patterns, such as therapy after response, therapy after treatment failure and second choice therapy, may be based on observational data or 'expert opinion'. Estimates of treatment effect may be derived from one or more randomised controlled trials, while medical records and observational databases may provide information on the incidence of side effects and adverse clinical events.

An evaluation that relies on a compilation of data from many sources, some of which are based on regular clinical practice, may provide an opportunity to increase the generalisability of the evaluation, augmenting the external validity in relation to other patient populations and settings. Variation in healthcare systems, unit costs and patient population base are factors that may impact on estimates of both costs and effects, influencing generalisability of cost-effectiveness estimates. Models carry the potential to accommodate for such factors, by varying their structural and parametric assumptions in a sensitivity analysis. (Manning et al 1996)

Quality assessment of clinical research is developing apace. Clinical scientists and editors of biomedical journals developed the Consolidated Standards of Reporting Trials (CONSORT) statement, listing essential items for reporting of randomised controlled

trials. (Begg et al 1996, Moher et al 2001) Others have developed quality scales for the assessment of published randomised controlled trials. (Moher et al 1996) Researchers in the area of health economics have attempted to develop frameworks for reporting and assessing the quality of economic models. (Eddy 1985, O'Brien 1996, Drummond and Jefferson 1996, McCabe & Dixon 2000) Schulpher et al (2000) derived attributes of good modelling practice within a series of dimensions, including model structure, time horizon, data identification approach and internal validity. But to date, authors have only to a very limited extent incorporated issues of generalisability in such modelling assessment frameworks.

Regulatory bodies acknowledge the importance of local relevance of data submitted for reimbursement and clinical guidance purposes (see section 2.3), but few formal requirements exist regarding generalisability to models that are submitted to jurisdictions requiring such information.

The outcome of full economic evaluations is a collapsed measure of costs and effects in the cost-effectiveness ratio. Inherently, resource consumption data and unit cost values, as well as clinical effectiveness estimates and valuation, may be more or less generalisable. Data determines the generalisability of the model and the applicability of the results depends on the range of inputs that were explored in the model. Models may seek to generalise costs and effects separately, and also to express uncertainty in all parameters simultaneously through probabilistic modelling. (Briggs 2000) What are the merits of modelling as a method for generalising the results of an economic evaluation to a broader healthcare setting?

There is little empirical evidence of the degree to which published models take local circumstances into account when presenting results and a framework is needed for assessing a model's explicit attempt to generalise parameter estimates for clinical effectiveness and costs. Applying this framework to health economic models in one therapeutic area would provide a meaningful comparison of study methodology. The selection of one therapeutic area enables the study of differences across study results that may be explained by location. This chapter outlines the development of such a framework and reports an assessment of published models in the therapy area of

osteoporosis. Several studies published since 1980 on the treatment and prevention of this condition provide a rich source of comparison of methods and results.

4.3 Objectives

The objectives of the review were to

- Describe the methods that may be employed by health economists to increase external validity of economic models.
- Systematically assess economic models in osteoporosis for the degree to which they make use of these methods to increase generalisability of the evaluation.
- Assess the relative impact of the adoption of these methods on the cost-effectiveness estimate of the evaluation.

4.4 Methods

4.4.1 Development of a framework for model assessment

Literature on generalisability and external validity of clinical and economic evaluations was used as background material for the development of a comprehensive framework for assessing aspects of external validity of models. (Mason 1997, Revicki and Frank 1999, O'Brien 1997). The cost-effectiveness ratio is a composite measure of relative resource use and clinical outcomes of healthcare therapies, so a framework to evaluate separately the degree to which cost-effectiveness models attempt to increase generalisability needs to take into account factors that influence costs, effectiveness and external consistency.

Unit costs and clinical management patterns may vary between geographical areas. For example, the average length of stay in hospital for patients with hip fracture is 29.6 days in Aberdeen and 41.7 days in Peterborough. (French et al 1994) In Denmark it is

reported to be 21 days. (Ankjaer Jensen 1996) The level of healthcare resource use and costs of services are likely to differ substantially between settings, both nationally and internationally. (O'Brien 1997) Models that incorporate such regional or national variation in unit cost and treatment patterns may reflect such regional variation, increasing the external validity of the analysis.

Secondly, the clinical effect estimate of the technology evaluated in a model may depend on underlying population and the compliance levels of the target population. Models may undertake adjustments to generalise those assumptions of explanatory trials that might not hold in real life. For example, characteristics relevant to the target population should be specified and where an underlying study has been based on a population not relevant to a particular decision-maker, a model could account for this by adjusting baseline risk estimates. Compliance is generally acknowledged to be higher within the context of controlled clinical trials than in clinical practice, so may contribute substantially to the reduction in efficacy when a drug is used in a different environment. (Revicki & Frank 1999, Bombardier & Maetzel 1999) Patients may choose to decline treatment, accept and discontinue therapy or accept and use therapy intermittently. Reduced compliance in a clinical practice setting is likely to result in reduced effectiveness of the drug, so models that evaluate the population-based impact of a strategy in clinical practice may provide a more representative estimate by factoring in the reduced compliance in the analysis. (Hughes et al 2001) The use of pooled effect estimates based on many trials may help increase external validity, (Pang et al 1999) as may the synthesising of head to head comparisons that have not previously been subject to direct comparison.

Finally, checks can be made for external consistency by validating intermediate outputs to external sources of population-based data and comparing the result of the model with other studies in the literature or economic evaluations of similar technologies. (Schulpher et al. 2000)

The development of a generic checklist (*Table 4.1*) enabled an assessment of the degree to which the economic models adopted the aforementioned techniques, as well as the calculation of summary scores for the studies included in the review. Differences across results may have varied across countries, changed over time or varied according

to the assumed level of compliance, so a quantitative assessment was planned to investigate the impact of 'generalisation techniques' on estimated cost-effectiveness.

Ten aspects of model evaluation were identified (*Table 4.1*), covering cost estimates (Q 1-4), clinical estimates (Q 5-9) and a check for external consistency of the model (Q 10).

Table 4.1 A framework for assessing generalisability of decision-analytic cost-effectiveness models

Question	Dimension of generalisability	Question for appraisal
1	Costs	Did the model attempt to reflect variation in different resource use patterns between healthcare environments nationally (e.g. varying hospitalisation rates and length of stay)?
2	Costs	Did the model attempt to reflect variation in different resource use patterns between healthcare environments internationally?
3	Costs	Did the model attempt to reflect a range of costs relevant to a different healthcare environment nationally?
4	Costs	Did the model attempt to reflect a range of costs relevant to different healthcare environments internationally?
5	Effects	Were the effect estimates based on epidemiological studies, pragmatic trials or meta-analyses?
6	Effects	Were the effect estimates from the underlying studies moderated to reflect the target population of the model?
7	Effects	Did the model accommodate for difference in compliance rates between those observed in clinical trials and those likely in usual care?
8	Effects	Were head to head comparisons synthesised?
9	Effects	Were intermediate outputs of the model compared to external sources?
10	External consistency	Were the results compared to other relevant studies?

4.4.2 Material for the application of the framework

Therapy area

Osteoporosis is a disease characterised by low bone mass density (BMD) and deterioration of bone tissue, leading to enhanced bone fragility and consequent increase in fracture risk. High age and low BMD are important risk factors for osteoporotic fractures (Dennison and Cooper 2001), and therapies aimed at the treatment or prevention of fractures target the mechanism of bone loss. Screening procedures that assess BMD have been developed to identify patients with a low bone mineral density and subsequent high fracture risk. (Hailey et al 1998) Osteoporosis is manifested by the occurrence of hip fractures, wrist fractures and vertebral fractures (Royal College of Physicians 1999) and is most prominent in postmenopausal women. In many western countries the remaining lifetime risk of a hip fracture in white women at menopause is approximately 14%. (Cooper 1996) In the UK, it has been estimated that the cost of fractures occurring in women over the age of 50 is in excess of £700 million per year, of which hip fracture comprises 87%. (Dolan & Torgerson 1998)

Development of fractures is a complex function of osteoporosis, age and other risk factors that evolve over time. The majority of trials have evaluated the impact of hormone replacement therapy (HRT) and other osteoporosis treatments on the intermediate endpoint BMD. (Royal College of Physicians 1999) Only recently have studies evaluated the impact of osteoporosis therapies on the final endpoint of fracture risk. (Torgerson and Bell-Syer 2001, Cummings et al 1998) Historically, there has been a scarcity of literature evaluating final endpoints. (OTA 1995) Decision-analytic models enable final outcomes to be estimated from intermediate outcomes, and also, final outcomes can be assumed from findings in epidemiological studies. The use of economic modelling to assess cost-effectiveness of interventions in this therapy area has therefore been of particular relevance.

Previous reviews of economic evaluations in the osteoporosis area have provided narrative summaries of information and have scrutinised the methods adopted. (Jönsson et al 1995, Torgerson and Reid 1997, Torgerson & Reid 1999, Schulpher et al 1999) The methodological shortcomings of models in this area include the frequent use of non-

empirically derived quality of life weights, inappropriate application of decision rules and shortfall of stochastic models. (Schulpher et al 1999)

Economic model inclusion criteria

Full economic models evaluating therapeutic interventions in osteoporosis were reviewed, and primary studies that considered both costs and benefits of treatment in comparison with one or more alternatives were included. Only studies published in English were included. Economic evaluations that did not describe the construct of a model were excluded from the assessment, as were simple cost analyses and secondary reviews of economic evaluation models.

Computerised searches for published papers reporting economic evaluation models of drugs in osteoporosis were conducted in the bibliographic databases Medline, Cochrane collaboration database and the NHSEED (NHS Economic Evaluation Database). Key words included cost and cost analysis (all subheadings) health economics, economic evaluation, economic*, cost*, osteo*, bispho*, hormone replacement, etidronate, alendronate, HRT, ORT, ERT and English language. Searches were extended to bibliographies of retrieved articles. Reference lists of other review articles in the area were scrutinised and specialists in the field of economics of osteoporosis were consulted. Searches were undertaken in May 1999.

For included studies, information relevant to the review was extracted onto a data extraction form by the author (HU) for each separate study (appendix 2). (CRD 2001) This information was then summarised in data tables, which provided the basis for assessment of the studies.

4.5 Results

Overview

Six publications were excluded from the review (*Table 4.2*). These were either review articles (Delva 1993), postal surveys (Torgerson et al 1995), merely technical reports on

model structures and not used to evaluate specific intervention strategies (Jönsson et al. 1995, Ross et al, 1998, Zaethraeus et al, 1998), or duplicate publications (Jönsson 1996).

Table 4.2 Table of excluded studies

Study	Reason for exclusion
Delva 1993	Review article.
Jönsson 1995	Report a model evaluating cost utility of osteoporosis interventions, however base the analysis on a fictive intervention with a certain efficacy.
Jönsson 1996	Duplicate of Jonsson 1995.
Ross 1998	Describes a model to evaluate cost of osteoporosis, but does not evaluate specific treatment regimes.
Torgerson et al 1995	Reports the result of a postal questionnaire survey and not a full model.
Zaethraeus 1998	Describes a model, but does not evaluate specific treatment regimes.

A total of 19 publications reporting economic evaluation models were identified that satisfied the inclusion criteria (*Table 4.3*). The majority of these which were decision analytic models, including four Markov state transition models. Five were simple decision trees. (Frances (1996), Torgerson et al. (1996) Torgerson (1993), Goddard (1990), Visentin (1996)) A reference list for all the included models can be found in appendix 3.

The first economic model in osteoporosis was published in 1981 (Weinsten 1981). Of the studies published in the 1980s, six were from the United States. Eight studies were from European countries, all of which were published in the 1990s.

Half of the studies were cost utility analyses (*Table 4.4*) and of these, only one used utilities derived by patients. (Daly 1996) The others based quality of life (QoL) weights on those originally assumed by Weinstein (1981). Target populations were predominantly women living in the community, though one specifically evaluated treatment of women in nursing homes. (Torgerson 1993) None of the studies evaluated cost-effectiveness from a broader perspective than that of the health service sector.

Adoption of generalisation techniques

Most studies applied unit cost data to the analysis relevant to the country for which the study was targeted (*Table 4.5*) and three used regional cost estimates (Cheung 1992, Geelhoed 1994 Rosner 1998). None of the models accommodated differences in treatment patterns within regions or across countries. Responses to questions 1 through 4 of the assessment framework revealed that none of the investigators attempted to use cost estimates applicable to a broader audience of decision-makers within or between countries by using a range of costs or treatment patterns representing geographical differences.

A common feature of the models in this review was the use of epidemiological studies to estimate the relative fracture risk reduction in the treated populations (*Table 4.6*). Three models based the hip fracture risk reduction estimates on individual clinical trial data (Francis et al. 1996, Rosner et al. 1998, Torgerson et al. 1995). Ankjer-Jenssen et al. (1996) based the effect estimate for one of the therapies on a meta-analysis of several trials. Of these, only Francis et al. (1996) provided some information about the characteristics of the population from which the estimates were derived. Fourteen studies did not provide any patient information enabling comparison with the target population of the model. The remaining three studies predominantly restricted patient information to age range. (OTA 1995, Rosner et al. 1998 and Tosteson et al. 1990).

Questions 5 through 9 of the assessment framework relate to a clinical effectiveness estimate of the models (*Table 4.1*). For the most part, those studies that did not use estimates of effectiveness from clinical trials based the effect estimate on assumptions supported in epidemiological literature (*Table 4.6*). As a result, the effect estimate in many studies was based partially on methods more reflective of clinical practice than explanatory randomised controlled trials. Two studies (Goddard 1990 and OTA 1995) provided a description of patient characteristics in the studies on which the model's assumption of effectiveness was based. The scope for decision-makers to compare the population on which the original effect estimate was based with that of the population targeted by the economic model (question 6 of the checklist) was, therefore, limited. With the exception of Rosner et al. (1998), there was no incidence of explicit adjustment in risk reduction from underlying study to the target population of the model.

Most studies identified in this review evaluated more than one treatment strategy. Of these, only Rosner et al. (1998) and OTA (1995) analysed incremental cost-effectiveness based on a synthesised head-to-head comparison. The remainder reported average cost-effectiveness ratios from comparing each treatment strategy with no treatment (*Table 4.6*).

Finally, five studies facilitated assessment of the external consistency of results by contrasting their findings with other economic evaluation studies in the area (*Table 4.6*)

Generalisability and the cost-effectiveness ratio

Between them, the studies evaluated different interventions, including a range of clinical outcomes in their calculations of costs and effects of treatment, expressing their results on different currencies from different years (*Table 4.3*). Furthermore, the results of the studies were expressed partly in terms of cost-effectiveness analysis and partly of cost-utility analysis (*Table 4.4*), so comparison of results across studies was not straightforward. Nevertheless, an assessment was planned to investigate the impact of generalisation techniques on the estimated cost-effectiveness.

Not infrequently, assumptions of compliance and duration of treatment were made without adjustment. For example, OTA (1995) assumed 100% compliance in 10, 20, 30 and 40 years respectively. The definition of compliance differed between the eight studies that took this into consideration, but often it meant simply that patients 'declined to accept' therapy (e.g. Clark et al. 1992) or that patients 'accepted but discontinued' (e.g. Daly et al. 1992) (*Table 4.7*).

In general, the cost-effectiveness estimate was found to be sensitive to the assumption of compliance, but the decision recommendation remained unchanged (*Table 4.7*). For example, Tosteson et al. (1990) assumed 100% compliance over 15 years in the base case model, but varied compliance to 30% in the sensitivity analysis and found that cost-effectiveness estimates were sensitive to assumption of compliance.

Table 4.7 Definition of patient compliance and impact on compliance on the results of the cost-effectiveness analysis.

Study	Definition of Compliance	Result of analysis
Ankjaer-Jensen 1996	Compliance was % of patients attending bone mass density (BMD) screening. 'Compliance' of 100% and 50% was explored.	Screening was more cost-effective than population based approach under both assumptions of 'compliance' (attendance).
Clark 1992	Compliance was % of patients completing 15 years of treatment. Model base-case was 90% compliance for high-risk patients and 70% compliance for mid-risk patients. Explored impact of reduced compliance at 10% intervals	Calculation of cost savings were more sensitive to compliance in high-risk group. Break even (cost of savings = cost of intervention) was reached at compliance of 77% in high risk only, or 70% in high risk and 40% in mid-risk group.
Daly 1992	Compliance was % of patients continuously taking the drug. Compliance of 100% in first 5 years falling to 50% (oestrogen only, ORT) and 67% falling to 33% (oestrogen and progestin)	Cost-effectiveness increases with reduced compliance over all treatment strategies. The overall recommendation of the analysis does not change.
Garton 1997	Compliance rate used implied % patients initiating therapy and continuing beyond year 1 to complete a course of 10 years treatment. Compliance of 10, 30 and 50% were explored.	Universal HRT was more cost-effective than screening strategy both under high and low compliance. However if screening could increase compliance then screening could prove more cost-effective.
Geelhoed 1994	Assumed that patients that fill prescriptions but do not take the drug incur costs but gain no benefits. Explored scenario where 70% of prescriptions are filled but only 30% are taken as directed.	The net cost per QALY would increase ('more than double') under a scenario of reduced compliance.
Rosner 1998	Different rates of willingness of patients to initiate (WTI) and continue (WTC) treatment were incorporated in the model ranging from 18.1% to 100%. These were based on epidemiological studies.	The model was "moderately sensitive" to changes in WTI and WTC. One strategy was particularly sensitive but it remained cost-effective under the assumption that public willingness to pay for a QALY is more than \$100,000
Torgerson 1993	Compliance rates were defined as 'willingness to initiate' therapy. Once therapy initiated 100% compliance assumed in 30 years. A range of compliance rates explored.	The outcome of the evaluation was sensitive to the assumption of compliance. More than 50% need to initiate therapy ('be compliant') in order that targeted intervention is cost-effective.
Tosteson 1990	Assumed 100% compliance with treatment varied from 5 year to lifetime use (Baseline model 15-year). Explored 30% compliance in sensitivity analysis.	Results were "sensitive", but compliance did not change the main conclusion of the model: screening remained the more cost-effective than universal treatment. Average CE ratio increased with low compliance.

The main base-case results of the studies did not reveal any systematic differences across results that may be explained by location (*Table 4.8*). Neither systematic variation in cost-effectiveness estimates within countries or systematic changes over time were apparent from the overview. The models compared a range of interventions and presented the result in a variety of outcomes. Ankjaer-Jenssen (1996) reported average cost per hip fracture avoided in screened and unscreened populations comparing three different interventions. In contrast, Tosteson et al. (1991) compared cost-utility of two different interventions in patient populations with different life expectancy. The large range of interventions, patient populations and outcomes recorded in the models prevented quantitative assessment of the impact of model features on the reported cost-effectiveness and with 19 studies, the review was underpowered for the conduct of such an assessment.

Study score

Table 4.9 provides an overview of individual study scores according to each item on the checklist. This information is a synthesis of the information incorporated in *Tables 4.5 and 4.6*, aiming to give an overview of the degree to which models in osteoporosis published prior to 1999 have attempted to generalise the study results.

Table 4.9 Summary table showing overall score for each study on all dimensions of external validity.

Study (year)	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total 'Yes' score
Ankjaer-Jensen 1996	No	No	No	No	In part	No	Yes	No	No	Yes	2
Cheung 1992	No	No	No	No	In part	No	No	No	No	No	0
Clark 1992	No	No	No	No	In part	No	Yes	No	No	Yes	2
Daly 1992	No	No	No	No	In part	No	Yes	No	No	Yes	2
Daly 1996	No	Yes	No	No	In part	No	No	No	No	No	1
Francis 1996	No	No	No	No	No	No	No	No	No	No	0
Garton 1997	No	No	No	No	No	No	Yes	No	No	Yes	2
Geelhoed 1994	No	No	No	No	In part	No	Yes	No	Yes	No	2
Goddard 1990	No	No	No	No	In part	No	No	No	No	No	0
OTA 1995	No	No	No	No	No	In part	No	No	Yes	Yes	2
Rosner et al 1998	No	No	No	No	Yes	Yes	Yes	Yes	No	No	4
Torgerson 1993	No	No	No	No	In part	No	Yes	No	No	Yes	2
Torgerson 1995	No	No	No	No	In part	No	No	No	No	No	0
Tosteson 1990	No	No	No	No	No	Yes	Yes	No	Yes	No	3
Tosteson 1991	No	No	No	No	No	Yes	No	No	Yes	No	2
Visentin 1997	No	No	No	No	Yes	Yes	No	No	No	Yes	3
Weinstein 1980	No	No	No	No	In part	No	No	No	No	NA	0
Weinstein 1983	No	No	No	No	In part	No	No	No	No	No	0
Weinstein 1990	No	No	No	No	In part	No	No	No	No	No	0

4.6 Discussion

Study scope

Previous authors have reviewed and made exhaustive recommendations on the most cost-effective prevention and treatment of osteoporosis (Schulpher et al 1999) and it was beyond the remit of this review to add to them. Attempts have also been made to develop a standardised framework for evaluating the quality of health economic models. (Eddy 1985, Schulpher et al. 2000) The diversity of decision analytic models and the types of data and settings used have made quality assessment a difficult task and it was not within the scope of this review to systematically assess the quality of health economic models in the area of osteoporosis. The purpose of this review was to develop a framework for assessment of model generalisability and to apply this framework on a sample of economic models in which such methodologies are applied.

Model assessment

To increase the external validity of an economic model, the developer could explore the impact of parameter estimates of costs and effects of treatment under circumstances different from those of the data underlying the base-case estimates. The checklist designed for this review considered cost estimates such as healthcare management systems, as well as clinical estimates such as characteristics of patient populations and treatment patterns. Although not exhaustive, it provided a comprehensive framework for evaluating the models in the review.

Cost estimates

None of the 19 economic models of therapies in osteoporosis published prior to May 1999 included a range of costs to reflect a variety of settings. Also, in spite of the fact that therapy traditions and healthcare management differ across units of healthcare providers, both nationally and internationally, none of the studies accommodated for such variation. Neither did the models incorporate variations in treatment patterns in different areas. As a result, though the studies frequently applied at an aggregate national level, they may not have relevance to regional or local healthcare decision-makers or, indeed, national decision-makers concerned about regional variability.

Health and social care systems operate differently. Many factors are likely to influence the resource use, costs and outcome of treatment in the countries involved (O'Brien 1997, Drummond 1992), but none of the studies attempted to extrapolate from their results to another country or continent. Successful, large-scale, international economic evaluation studies have been conducted despite a series of obstacles and an increasing number of publications provide a solid foundation for such research. Some prospective studies have compared trial-wide clinical results with costs based on trial-wide utilisation, while using unit prices for the country in question. (Menzin 1996, Jonsson 1997, Schulman 1996) This method estimates country-specific cost-effectiveness, making some adjustment for gross price differences across countries, but does not allow for differing treatment patterns (Willke et al 1998), which could be captured by applying national costs to national trial data and then pooling the data. (Glick et al. 1998) Modelling studies could provide useful frameworks for exploring the impact of varying unit costs and treatment patterns, but this asset was not exploited in the osteoporosis evaluations included in this review.

Effectiveness estimates

According to drug regulatory requirements, the development of a knowledge base for a pharmaceutical compound progresses from an initial assessment of drug efficacy in phase I to phase III randomised controlled trials, through to assessment of effectiveness in pragmatic randomised trials and observational studies. (Piantadosi 1997) Assuming that economic modellers base their analysis on the best available data, one could expect the model input to be more naturalistic as experience with a drug has accumulated over time. This does not seem to have been the case with hormone replacement therapy (HRT). The knowledge base has predominantly constituted non-randomised, epidemiological studies from the initial stages of routine use in menopausal women. Randomised studies have been carried out only recently. (Royal College of Physicians 1999) This is reflected in the material on which the economic models base their effect estimate. The general absence of randomised clinical trial data for use in the clinical effectiveness estimates in osteoporosis models is striking. Early modellers between 1980 and 1995 based the effect estimate on assumptions supported by epidemiological literature. Later studies based their effect estimates on randomised clinical trials. (OTA 1995, Torgerson 1995, Rosner 1998) Nevertheless, a recent study by Rosner et al

(1998) emphasised that hip fracture was excluded as main outcome because of scarcity of clinical trial data in the literature, so vertebral fractures were used as the main outcome.

Most economic models in the review omitted a description of the patient population characteristics on which the effect assumption was based. This is a key area for decision-makers trying to assess a model's relevance to a specific population and its omission severely restricts the value of information presented.

The concept of clinical compliance with medication is complex. Definitions such as intermittent therapy, omitted tablets, therapy breaks, and lack of adherence to administration regime have been used to describe the phenomenon. (Hughes et al. 2001) This definition may be important to the interpretation of the results. The authors of those studies in this review that take compliance into consideration used different definitions, but for the most part, assumption of compliance was based on the fraction of patients that initiate therapy, 'willingness to start therapy'. (Clark and Schuttinga 1992). Thereafter, patients are assumed to continue for a defined period of time, frequently in excess of 5 years.

An asset of economic modelling studies is that they may synthesise head-to-head comparisons where no direct trial-comparisons exist. (Buxton et al 1997) Standard decision rule methodology is to compare treatments to each other in an analysis of incremental cost and effectiveness. The studies in this review, however, with the exception of OTA 1995, and Rosner 1998, calculated cost-effectiveness compared to no treatment and contrasted treatments by presenting average cost-effectiveness ratios. This practice limited the scope for the synthesis of head to head comparisons, a practice that could have increased the external validity for decision-makers who face therapy choices between treatments in clinical practice.

External consistency

The question of whether intermediate outputs were compared to external sources is relevant to those studies in this review where hip fracture is modelled from a change in bone mass density, and where survival is modelled from hip fracture. There is a trade-

off as to whether available information at the modelling stage should be used in validation or whether it should provide an additional knowledge base. (Schulpher et al. 2000) In this review, Tosteson (1990 and 1991), Geelhoed et al (1994) and OTA (1995) did make this check of external consistency of intermediate outputs by contrasting hip fracture risk modelled from bone mineral density with population-based hip fracture risk stratified by age. Less than half of the studies (7) contrasted the methodology or the results of their evaluation with those of other economic evaluations in osteoporosis. It would be reassuring to see more researchers contrasting their findings to other studies with similar objectives.

The merit of models as tools for generalisability

The checklist developed in this chapter to assess the degree to which economic models seek to increase external validity demonstrates that there is significant potential for these studies to inform decision-makers on generalisability of the results. However, this comprehensive review of published models in osteoporosis treatment and prevention reveals that studies to date perform modestly in terms of adopting these aspects. In stochastic model evaluations, the sensitivity analysis may be a relevant stage in the process to explore alternative data input relevant to larger decision-making audiences. Future use of probabilistic evaluation is likely to become increasingly developed and adopted, and this will enable researchers to use a range of values and reflect the sensitivity of the results in the form of cost-acceptability curves. (Briggs 2000)

An underlying assumption, of this review is that a study attempting to increase the external validity of its findings is one that varies costs over a range of nationally and internationally relevant unit costs and also varies the therapeutic pattern within and between countries. Furthermore, it bases effect estimates on pragmatic trials or meta-analyses, moderates these estimates to the target population of the model and accommodates for reduced compliance rates in clinical practice. Finally, the model performs incremental analysis between the synthesised head to head comparisons, compares intermediate output with external data sources and compares the output of the model with relevant studies in the field.

Are these realistic expectations for published reports of economic evaluation models? Restricted space in scientific journals limits the scope for the researchers to describe even methodology appropriately. As already mentioned, the concept of external validity is meaningless unless it is related to a particular setting, population or time, so the degree to which a model reflects variation in costs and effects should follow from the objectives of the study. If a model intends to inform a decision on treatment strategy by the National Institute for Clinical Excellence (NICE) then the report benefits from evaluating a national range of costs and benefits. Similarly, if the report is intended for an international audience in an international journal, then it would benefit from reflecting a span in costs, treatment patterns and populations between countries. It may not be fair to expect researchers to report additional information from an expanded analysis of range of costs to a huge range of alternative settings, but it would be relevant for models to evaluate the robustness of the results under assumptions from more than one setting.

Study limitations

There are a number of limitations with this model evaluation. Firstly, the small number of studies in only one therapeutic area limits the scope for extrapolation of the findings. The study did not provide enough power to conduct a regression analysis evaluating the impact of increased generalisability on cost-effectiveness estimates. The relatively low number of studies (19) did not provide enough power for such analysis, and furthermore, the use of a range of measures of effectiveness in the analyses reduced the opportunity to generate one uniform expression of cost-effectiveness.

Secondly, only studies published in English were included, and these were predominantly published in international journals aimed at international audiences. This may have introduced a bias to the study that undermines the issues of local adaptability that were explored. Few models assessed in the review attempted to generalise results across settings. Economic models in osteoporosis may nevertheless have been used with the purpose of generalising economic evaluations to specific settings, however reports developed for specific local healthcare decision-makers may not have been published internationally. A comprehensive review of submissions to regulators such as the Pharmaceutical Benefits Advisory Committee in Australia or National Institute for Clinical Excellence in the UK may have provided a different empirical result.

Thirdly, the checklist was limited to evaluating sensitivity analysis on parametric assumptions only. For example, it is a structural assumption that the risk of hip fracture is altered only after a given duration of therapy. Parametric assumptions describe the magnitude of the structural assumption, for example the length of time the risk of hip fracture is altered by treatment. Assessing the sensitivity of results to changes in structural assumptions is more labour-intensive, as it may require reprogramming of computer models and this was not attempted.

Finally, the data extraction recorded only information that was reported in the paper, and the research is restricted by the degree to which models are described in detail. Space restriction in scientific journals may limit scope for presentation of comprehensive sensitivity analyses, and the findings of this review would reflect that rather than a genuine lack of attention to generalisability by the authors of the models.

4.7 Conclusions and recommendations

Economic modelling and data collection requires a series of key decisions on the part of the economist. Modellers can adopt several techniques in order to increase the generalisability of both resource consumption estimates and clinical estimates of economic evaluations. By using a relevant range of data sources or making data adjustments representative to a broad audience of decision-makers, models may increase external validity of the evaluation.

The findings of this review demonstrate that modellers in the area of osteoporosis either omit information that is essential for assessing external validity of a model or do not incorporate aspects specifically aimed at increasing the external validity of the model. There may be good reasons for this, for example limitations on space in scientific journals. Inclusion of such factors in future economic models in this area would however increase the ability of models to reflect a larger span of alternative clinical practice scenarios. Presentation of future economic models in a manner that enables replication of the model would furthermore facilitate the incorporation of local data for local decision-makers.

This review may have revealed a need for more standardised reporting of economic models. A framework analogous to the CONSORT statement advising researchers on how to report the construct, the data search and the findings of models may enable easier assessment of model quality as well as assessment of data aspects and the degree to which the model can be generalised to a population under consideration. In spite of previous efforts (Drummond and Jefferson 1996), no widely established framework exists for the reporting of economic models. In contrast to clinical trial methodology, economic evaluation methods are still in development and establishing of firm quality criteria may provide disincentives for further methodological developments. However, standardised reporting would facilitate quality assessment. Inclusion of the checklist of generalisability that was developed in this chapter in such a statement would facilitate decision-makers assessment of local study applicability.

For the term external validity to have meaning an evaluation needs to be related to a particular setting, patient group or a particular country. Restrictions of generalisability of a model may therefore differ across different decision-making audiences. The framework that was developed in the current chapter provides a comprehensive tool to assess the cost and the effect estimate of economic models. Therefore, it may be used by readers of economic models for the assessment of local applicability of the model.

The findings of this review point to three main areas of future research. Firstly, the same checklist should be applied to a larger subset of models incorporating models from different therapy areas. Increasing number of models would increase the power to undertake a regression analysis of the impact of increased generalisability on model estimate and would enable exploring of features that may influence generalisability. Secondly, the same checklist could be applied to unpublished models submitted to a regulatory authority to assess whether such models to a larger extent attempt to generalise results. This could take an analogous form to the quality assessment of submissions to the Australian Pharmaceutical Benefits Advisory Committee, conducted by Hill et al (2000). Finally, the vast majority of these studies were published before the era of guidelines for the conduct of economic evaluations. Advice by such reference cases for the conduct of economic evaluations that have emerged in recent years (NICE, Australian, Canadian guidelines) would therefore not have influenced the conduct of the

models included in this review. It would therefore be of interest to see whether more recent models to a larger extent take issues of generalisability into consideration in order to meet the need of decision-makers for locally adaptable data.

**PAGE
MISSING
IN
ORIGINAL**

Table 4.3 Details of the models included in the review

Study (year)	Interventions	Design	Time horizon	Perspective	Currency (year)
Ankjaer-Jensen 1996	Screening or not followed by calcium supplementation, etidronate, calcitonin or HRT.	Decision analytic model	Lifetime from 50 years	Health service	1994 Danish Kroner (DKK)
Cheung 1992	Oestrogen alone vs. no treatment and oestrogen and progestin vs. no treatment	Decision analytic model (Replication of Weinstein & Schiff)	Lifetime from 50 years	Health service.	1988 Australian \$
Clark 1992	Screening followed by oestrogen and progestin vs. no screening or treatment.	Decision analytic model	From age 50 to age 90	Health service	1988 US \$
Daly 1992	Oestrogen only vs. no treatment (no uterus) and oestrogen vs oestrogen + progestin (uterus)	Decision analytic model	Lifetime from 50 years.	Health service	1989/90 UK £ Sterling
Daly 1996	Oestrogen only vs. no treatment (no uterus) and oestrogen vs oestrogen + progestin (uterus) vs. no treatment.	Decision analytic model (Replication of Daly 1992)	Lifetime from 50 years.	Health service	1992/93 UK £ Sterling
Francis 1996	Calcitonin vs. no treatment, estrogen + progestin vs no treatment, etidronate vs. no treatment.	Decision tree	2 years	Health service	1994 UK £ Sterling
Garton 1997	A strategy of oestrogen + progestin with screening vs a strategy of oestrogen + progestin without screening (universal)	Decision analytic model	45 – 110 years	Health Service	1993 UK £ Sterling
Geelhoed 1994	Oestrogen vs. no treatment and lifestyle regime (exercise) vs. no treatment.	Markov state transition model	From age 50 to age 100	Health Service	1991 Australian \$
Goddard 1990	Oestrogen and progestin	Decision tree	10 years	Health service	1988 UK £ Sterling

OTA 1995	Estrogen (ERT) vs. estrogen + progestin (PERT) following different screening + treatment strategies	Markov state transition model	From age 50 to age 90	Health Service	1993 US \$
Rosner et al 1998	Estrogen, estrogen + progestin, alendronate, etidronate, calcium combinations and no intervention	Decision analytic model	3 years	Health Service	1996 Canadian \$
Torgerson 1993	Compares a strategy of estrogen only in high risk patients (screening) vs. estrogen + progestin in population based treatment	Decision tree	30 years	Health service	1992 UK £ Sterling
Torgerson 1995	Vitamin D and calcium vs. no treatment and vitamin D vs. no treatment	Decision tree	3 and 4 years	Health service	1994 UK £ Sterling
Tosteson 1990	No intervention, screening followed by HRT (estrogen + progestin) and universal HRT (estrogen + progestin).	Markov state transition model	From age 50 to age 100	Health Service	1987 US\$
Tosteson 1991	Estrogen vs. no intervention (intact uteri), estrogen + progestin vs. no intervention (hysterectomy).	Markov state transition model (replication of Tosteson 1990)	From age 50 to age 100	Health Service	1990 US\$
Visentin 1997	Calcitonin vs. no treatment	Decision tree	1 year	Health Service	1995 US \$
Weinstein 1980	Estrogen vs. no intervention	Decision analytic model	Lifetime from 50 years	Health Service	1980? US \$
Weinstein 1983	Oestrogen vs. no treatment and oestrogen + progestin vs no treatment.	Decision analytic model	Life time from 50 years	Health Service	1982 US \$
Weinstein 1990	Oestrogen vs. no treatment and oestrogen + progestin vs no treatment.	Decision analytic model	Life time from 50 years	Health Service	1988 US \$

Table 4.4 Clinical context of the model

Study (year)	Clinical outcomes considered	Outcome measure	Target population of model	Target setting of model
Ankjaer-Jensen 1996	Hip fractures	Cost per hip fracture avoided (CEA)	Women	Community, Denmark
Cheung 1992	Hip and wrist fracture, myocardial infarction, uterine bleeding, endometrial hyperplasia and cancer.	Cost per QALY (CUA)	Women age 50 (assumed to be in menopause)	Community, New South Wales, Australia
Clark 1992	Hip, vertebral and wrist fractures	Cost savings from reduced number of fractures (CEA)	White women age 50 who have recently gone through menopause	Community, US
Daly 1992	Endometrial cancer, breast cancer, hip, wrist and vertebral fractures, ischaemic heart disease and cerebrovascular disease	Cost per LYG and QALY (CUA)	Women age 50, hysterectomised and non-hysterectomised, respectively.	Community, UK
Daly 1996	Endometrial cancer, breast cancer, ischaemic heart disease, stroke and fractures of hip, wrist and vertebra	Cost per LYG and QALY (CUA)	Women age 50, hysterectomised and non-hysterectomised, respectively.	Community, UK
Francis 1996	Vertebral fractures	Cost per averted vertebral fracture (CEA)	Established osteoporosis in Caucasian women	Community, UK
Garton 1997	Hip, vertebral and wrist fractures.	Cost per fracture averted (CEA)	Women, of which 15% hysterectomised	Community, UK
Geelhoed 1994	Hip fracture, heart disease, breast cancer, ischaemic heart disease and endometrial cancer (and mortality)	Cost per LYG and QALY (CUA)	Healthy perimenopausal Caucasian women in cohorts according to hysterectomy status	Community, Australia
Goddard 1990	Beast cancer, ischaemic heart disease, hip fracture and cerebrovascular disease	Not clear	Women 65 – 74 years old	Community, UK
OTA 1995	Hip fracture and other fractures, heart attack, breast cancer, endometrial cancer and coronary heart disease.	Cost per LYG (CEA)	Women with intact uteri	Community, USA

Rosner et al 1998	Vertebral fractures	Cost of vertebral fracture averted and QALY (CUA)	Women with established postmenopausal osteoporosis (BMD measurement)	Community, Canada
Torgerson 1993	Hip fracture	Cost per averted hip fracture (CEA)	Hysterectomised and non-hysterectomised women	Community, UK
Torgerson 1993	Hip fracture and any other fracture	Cost per averted hip fracture (CEA)	Women in general population and women of high risk (BMD measurement)	Community and Nursing home, respectively, UK
Tosteson 1990	Hip fracture	Cost per QALY (CUA)	Women in the US	Community, US
Tosteson 1991	Hip fracture, ischaemic heart disease and breast cancer.	Cost per QALY (CUA)	Women in menopause	Community, US
Visentin 1997	Hip fracture	Cost per averted hip fracture (CEA)	Women over 50	Community, Italy
Weinstein 1980	Endometrial cancer, hip and wrist fractures	Cost per QALY (CUA)	Postmenopausal women	Community, US
Weinstein 1983	Endometrial cancer, breast cancer, hip fractures and wrist fractures.	Cost per QALY (CUA)	Women in menopause with uterus	Community, US
Weinstein 1990	Hip fracture, Endometrial cancer, Endometrial hyperplasia, Uterine bleeding, Breast cancer	Cost per LYG and QALY (CAU)	Women in menopause	Community, US

Table 4.5 Data sources for unit cost and resource use that underlie the cost estimates and the degree to which the models make attempts to generalise these.

Study (year)	Data source for unit costs	Source for resource use and treatment patterns	Q. 1. Did the model reflect variation in costs nationally (range)	Q. 2. Did the model reflect variation in treatment patterns nationally (range)	Q. 3. Did the model reflect variation in costs internationally	Q. 4. Did the model reflect variation in treatment pattern internationally
Ankjaer-Jensen 1996	Danish prices for drugs and GP visits. Danish cost studies for cost of clinical events (fractures, cancer, coronary heart events)	Danish cost of illness studies and National Patient Register	No Costs that apply nationally were used. However a range to reflect regional variability was not used.	No. Treatment patterns not varied	No	No
Cheung 1992	Australian MIMS for drugs, Australian fees for GP, NSW hospital bed costs. Cost of other costs not stated	Not stated	No Predominantly costs applicable to NSW used. Some Australian costs. No range.	No. As above	No	No
Clark 1992	US drug costs, no source stated but compared to Tosteson 1990. Hospital day cost from American Hospital Association. Other from US cost study	Average of 20 hospital days for hip fracture from American Hospital Association.	No Costs that apply nationally were used. However a range to reflect regional variability was not used.	No As above	No	No

Daly 1992	UK national average costs such as BNF and DoH for outpatient attendance. UK Social services cost data for inpatient days	GP attendance for all diseases from UK survey,	No Costs that apply nationally were used. However a range to reflect regional variability was not used.	No As above	No	No
Daly 1996	UK national average costs such as BNF and UK Social services cost data	Hospitalisation (rate + duration) from a DoH survey + other UK based studies	No	No	No	No
Francis 1996	Drug costs only. British National Formula (BNF) used	Not considered	No Costs that apply nationally were used. However a range to reflect regional variability was not used.	No	No	No
Garton 1997	Drug costs from BNF. Cost of clinical events based on UK costing study, or assumed. Cost of screening from a regional costing study	Based on assumptions	No Costs that apply nationally were used. However a range to reflect regional variability was not used.	No	No	No
Geelhoed 1994	Most hospital costs from Royal Perth Hospital (Western Australia). Drug costs national	Western Australian hospital data for hospitalisation and DoH data for nursing home	No Predominantly regional data	No National nursing home data used	No	No

Goddard 1990	Drug costs: DTB, Cost of GP visit unpublished data and UK DoH cost for cost of admission	UK costing study used for duration of hospitalisation	No Costs that apply nationally were used. However a range to reflect regional variability was not used.	No	No	No
OTA 1995	Apply US costs using sources such as Medicare and primary costing studies from OTA and other authors	Assumptions of treatment pattern of hip fracture incorporated in previous costing study by OTA using Medicare data.	No Costs that apply nationally were used. However a range to reflect regional variability was not used.	No Average US Medicare hospitalisation rate and rate of surgery were used, regional differences stated but no range provided	No	No
Rosner et al 1998	Use Ontario sources of unit cost estimates.	Canadian treatment practice patterns estimated from Delphi panel techniques. Aberdeen cost study	No Regional costs (Ontario) and no range	No	No	No
Torgerson 1993	Drug costs only: British National Formulary (BNF) Cost of screening from a study in Aberdeen, Scotland		No Costs that apply nationally were used. A range to reflect regional variability was not used.	No	No	No
Torgerson 1995	UK drug costs from UK MIMS, cost of a hip fracture assumed	No patterns described	No Costs that apply nationally were used. A range to reflect regional variability was not used.	No	No	No

Tosteson 1990	Unit costs identified both nationally (e.g. Medicare) as well as locally (Boston, Massachusetts).	Used national averages for hospitalisation rate and admission rate to nursing home.	No	No	No	No
Tosteson 1991	Unit costs identified both nationally (e.g. Medicare) and locally (Boston, Massachusetts).	Used national averages for hospitalisation rate and placement in nursing home.	No	No	No	No
Visentin 1997	Italian drug prices and healthcare costs	% hip fracture patients institutionalised from US data.	No	No	No	No
Weinstein 1980	Most costs were based on Medicare charges.	Average length of stay (24 days) from US Medicare data used.	No	No	No	No
Weinstein 1983	Unit costs were based on Medicare charges and drug prices from Boston, US.	Hospitalisation data from US Medicare data used.	No	No	No	No
Weinstein 1990	The unit costs for treatment and hospital costs for Medicare and drug prices averaged over pharmacies in Boston, US.	Hospital and nursing home pattern based on Medicare data.	No	No	No	No

DoH = Department of Health

BNF = British National Formula

NSW = New South Wales

DTB = Drugs and Therapeutics Bulletin

Table 4.6 *Data sources for clinical estimates that underlie the model probabilities and the degree to which the models make attempts to generalise these.*

Study (year)	Data source for main clinical outcome	Population for trial of main clinical outcome	Q. 5. Estimate based on pragmatic trial, epidemiological study or meta-analyses?	Q. 6. Were the effect estimates from the underlying studies moderated to reflect target population?	Q. 7. Did the model accommodate for difference in compliance rates between those observed in clinical trials and those likely in usual care?	Q. 8. Were head to head comparisons synthesised?	Q. 9. Intermediate outputs compared to external sources?	Q. 10. Were the results compared to other relevant studies?
Ankjaer-Jensen 1996	Meta analysis for one therapy (HRT) and selected clinical trials for the other therapies	Patient information not provided.	In part. A meta analysis was used for one effect estimate, individual trials for others.	No. Effect estimates applied to the target population without moderation.	Yes 100% compliance with treatment for 5 / 10 years assumed. 50% compliance with screening explored.	No Effect estimates from studies of the individual drugs compared to no treatment	No. Fracture estimates were not compared to independent external sources of fracture incidence data.	Yes. Compared to a study (review) by Jonsson et al 1995, and differences discussed.
Cheung 1992	Assumptions of effect in Weinstein & Schiff (1983) replicated for QoL and effectiveness.	Patient information not provided.	In part. Effect estimates assumed based on findings in epidemiological studies.	No. As above	No. 100% compliance for 5, 10 or 15 years, respectively, assumed.	No. As above	No. As above	No. No.

Clark 1992	Assume 50% fracture reduction with 15 years HRT treatment based on epidemiologic al studies	Patient information not provided.	In part. Effect assumed at the basis of findings in epidemiologic al studies.	No. As above	Yes 100% compliance for 15 years assumed. 90% and 70% explored.	No. Effect estimates based on assumptions and compared to no treatment.	No. As above	In part. Methodology compared to Tosteson 1990.
Daly 1992	Assume 20% - 60% fracture reduction with >5 years HRT treatment based on epidemiologic al studies. QoL assigned by authors.	Patient information not provided.	In part. Effect assumed at the basis of findings in epidemiologic al studies.	No. As above	Yes Base case 100% compliance in 5, 10 and 15 years. Explore 100% compliance for first 5 years for all patients, falling to 50% for ORT* and 33% for CRT* after 5 years.	No. Effect estimates based on assumptions and compared with no treatment.	No. As above	Yes. Results compared to those of Weinstein 1983 and Goddard 1990
Daly 1996	Assume 20% - 60% fracture reduction with >5 years HRT treatment based on epidemiologic al studies. QoL based on population survey.	Patient information not provided.	In part. Effect assumed at the basis of findings in epidemiologic al studies.	No. As above	No. 100% compliance in 5, 10, 15 and 20 years assumed.	No. Effect of HRT compared with no treatment in women with and without uterus.	No. As above	No. Results not compared.

Francis 1996	5 RCTs	Age range in the trials varied between 45 and 75	No	No	No	No	No	No	No
Garton 1997	50% risk reduction with HRT assumed based on a previous review article	Patient information not provided.	No	No	Yes Compliance with therapy for 10 years. Compliance with therapy (patients continuing) of 10, 30 and 50 % explored.	No	Yes Compliance with therapy for 10 years. Compliance with therapy (patients continuing) of 10, 30 and 50 % explored.	NA	Yes Tosteson 1990
Geelhoed 1994	Risk reductions assumed based on epidemiologic studies. QoL assigned by authors.	Patient information not provided.	In part Risk reductions based on epidemiologic studies or meta-analyses of such.	No. Base case was based on epidemiologic data from different parts of Australia, risk reductions were not.	Yes 100% compliance in 50 years and 15 years assumed. 70% compliance explored.	No. Effect of oestrogen and exercise were compared individually to no treatment.	Yes. Model's estimate of hip fracture risk was compared to an Australian population-based study.	No. Results of the cost-effectiveness analysis not compared.	No
Goddard 1990	Assumption (e.g. 50% fracture risk reduction) based on a review of epidemiologic studies	Patient information not provided.	In part	No Age range was restricted, perhaps because of underlying studies.	No	No	No	No	No

OTA 1995	Fracture risk reduction assumed based on epidemiologic studies.	Review of epidemiologic studies, ranging from Swedish cohort of women above 35 to US cohort of women over 65.	No. Treatment effect on bone loss assumed and model output not explicitly compared to independent studies.	In part. The baseline risk function and effect estimates were based on observational studies covering different populations.	No. 100% compliance assumed in 10, 20, 30 and 40 years respectively.	No. The outcomes of different treatment durations were compared to no treatment and compared with and without prior screening.	Yes. Model's estimate of hip fracture risks from BMD was compared to population based studies.	Yes. A broad discussion of methodology, costs, clinical assumptions and results in relation to other cost-effectiveness analyses.
Rosner et al 1998	Fracture risk reductions based on randomised clinical studies. QoL from 'expert panel' of clinicians.	Patient information not provided.	Yes. Effect estimates were identified from literature reviews of trials. Partly individual trials and partly combinations of trials were used.	Yes. An average was taken of two studies in populations considered to have higher and lower risks of fracture compared to target population	Yes. Different rates of willingness of patients to initiate and continue treatment was varied from 18.1% to 100% and based on epidemiologic studies.	Yes. The outcomes if the different treatments were combined in different synthesised treatment strategies for comparison in an incremental analysis.	No. The model estimates were not compared to independent external sources.	No. The findings are compared to recommendations of clinical guidelines, but not to findings in other cost-effectiveness studies.
Torgerson 1993	Assumption of 50% hip fracture risk reduction based on epidemiologic studies	Patient information not provided.	In part	No	Yes Compliance was 'willingness to initiate' therapy. Once therapy initiated 100% compliance assumed in 30 years	No	No	Yes Weinstein 1980, Weinstein and Schiff 1983

Torgerson 1995	1 trial, 1 study, 1 epidemiologic al study	Patient information not provided.	In part 1 epidemiologic al study	No	No Assume 100% compliance for 4 (4) years	No	No	No	No
Tosteson 1990	Effect estimate based on assumption of no bone loss while under treatment. Hip fractures are modelled from bone density.	Fracture risk function from survey of 45- 54 year old women in Rochester, Minnesota	No Effect assumed (no bone loss) and compared to epidemiologic al studies of HRT.	Yes. Risk function of hip fracture compared to fracture rates of other populations	Yes. Assume 100% compliance with treatment varied from 5 year to lifetime use (Baseline model 15- year). Explore 30% compliance in sensitivity analysis.	No. The comparison between 'treatment following' screening' and 'universal treatment' were compared with no therapy.	Yes. Model's estimate of age at first hip fracture was compared to estimates from a cohort study. Model's estimate of hip fracture risk was compared to population- based studies.	No. Cost- effectiveness ratios were not explicitly compared to those of other studies.	
Tosteson 1991	Model hip fractures from bone density and assume no bone loss while under treatment.	Same as above	No. Effect assumed (no bone loss) and compared to epidemiologic al studies of HRT.	Yes. As above	No. Assume 100% compliance with 10 and 15 years of HRT therapy. Do not explore compliance.	No. The comparison between 'treatment for 10 years' and 'treatment for 15 years' was compared to no treatment.	Yes. Model's estimate of age at first hip fracture, hip fracture risk and breast cancer compared to estimates from population based studies.	No. Cost- effectiveness ratios were not explicitly compared to those of other studies.	

Visentin 1997	Hip fracture risk from an observational study	Study based on women in Mediterranean countries, 77% of which Italian.	Yes Effect based on one epidemiologic al study	Yes The study was predominantly Italian women	No 100% compliance for 1 year assumed.	No Calcitonin compared to no treatment	No. The model estimates were not compared to independent external sources.	Yes. Compare to Tosteson 1991 and Geelhoed.
Weinstein 1980	Fracture risk reduction assumed based on epidemiologic al studies. QoL assigned by author.	Patient information not provided.	In part. The assumption of treatment effect was based on two population-based studies.	No. Data from studies assumed to apply to model population	No. Assume 100% compliance with 10 and 15 years of HRT therapy.	No. Treatment age 50-60 and age 55-70 were compared to no intervention	No. The model estimates were not compared to independent external sources.	NA. This was the first cost-effectiveness analysis published.
Weinstein 1983	Fracture risk reduction assumed based on epidemiologic al studies. QoL assigned by authors.	Patient information not provided.	In part. Effect estimates assumed based on findings in epidemiologic al studies.	No.	No. 100% compliance for 5, 10 or 15 years, respectively, assumed.	No. Assumptions of effect estimates compared to no treatment.	No. The model estimates were not compared to independent external sources.	No. Cost-effectiveness ratios were not explicitly compared to those of other studies.
Weinstein 1990	Fracture risk reduction assumed based on epidemiologic al studies. QoL assigned by authors.	Patient information not provided.	In part. Effect estimates assumed based on findings in epidemiologic al studies.	No.	No. Assume 100% compliance for 5 and 15 years, respectively.	No. Treatment for 5 and 15 years was compared to no intervention	No. The model estimates were not compared to independent external sources.	No. Cost-effectiveness ratios were not explicitly compared to those of other studies.

LYG= Life years gained

QALY= Quality adjusted life years

*ORT= ERT= oestrogen only

*CRT= PERT= oestrogen + progestin

Table 4.8 Main base-case results of the cost-effectiveness analyses studies included in the review. Determining the impact of location and time of publication on cost-effectiveness estimate

Study (year)	Country	Cost-effectiveness ratio	Cost per QALY
Ankjaer-Jensen 1996	Denmark	Cost per hip fracture avoided: Population based vs. screening: Calcitonin: 1.136.200 DKK Etidronate: 77.567 DKK Calcium: 522.300 DKK	NA
Cheung 1992	Australia	NA	All cases 10 years treatment ORT: \$ 9.500 – 17.500, symptomatic women ORT: \$ 45.800 – dominated, asymptomatic women ORT: \$ 6.510 – 1 020.000, hysterectomy CRT: \$ 26.100 – 1 450.000, asymptomatic women CRT: \$ 9.820 – 34.700, symptomatic women
Clark 1992	US	Did not provide cost-effectiveness ratio	NA
Daly 1992	UK	Cost per life year gained (10 years treatment): ORT: £ 2.900 in women with no uterus ORT: £ 8.300 in women with uterus CRT: £ 14.400 in women with uterus	10 years treatment in mildly symptomatic women ORT: £ 1.700 in women with no uterus ORT: £ 4.400 in women with uterus CRT: £ 6.200 in women with uterus
Daly 1996	UK	NA	Mildly symptomatic women ORT: £ 310 (5 years treatment) to £ 660 (20 years treatment) CRT: £ 550 (5 year treatment) to £ 1.250 (20 years treatment)
Francis 1996	UK	Cost per vertebral fracture avoided, treatment vs. no therapy: HRT: £ 138 – 680 Etidronate: £ 1.880 Calcitonin: £ 9.075 – 25.013	NA

Garton 1997	UK	Cost per fracture prevented HRT vs. no therapy: Screening: £ 1.000 – 4.200 Universal treatment: £ 1.200	NA	NA
Geelhoed 1994	Australia	NA	Intervention vs. no treatment: ORT for life from 50 years: \$ 8.810 ORT for 15 years: \$ 16.500 ORT from age 65: \$ 8.450 Calcium and exercise: \$ 28.480	
Goddard 1990	UK	Did not provide cost-effectiveness ratio	NA	
OTA 1995	US	Mean cost per Life Year Gained ORT (Screening): \$ 22.431-151.392 depending on duration of therapy and screening threshold. ORT (Population-based): \$ 23.334- 126.876 depending on duration of therapy. Life-long therapy most cost-effective.	NA	
Rosner et al 1998	Canada	ICER: Cost per averted vertebral fracture for four non-dominated strategies in order: (i) calcium – no therapy (ii) HRT – calcium – no therapy: ICER = \$166 (iii) HRT – etidronate – calcium – no therapy: ICER = \$ 2.331 (vi) HRT – alendronate – calcium – no therapy: ICER = \$ 40.965	NA	
Torgerson 1993	UK	Cost per averted hip fracture Universal treatment for 10 years: £ 40.080 Screening then treatment: £ 34.971	NA	
Torgerson 1995	UK	Cost per averted hip fracture Oral vitamin D + Calcium: £ 17.379 in community (low BMI); £4.735 in nursing homes; cost saving overall in nursing homes (low BMI). Vitamin D injection: all options save costs overall.	NA	

Tosteson 1990	US	NA	Screening and then treatment (15 years): \$ 4.200 – 37.800 depending on treatment threshold Universal treatment (15 years): \$ 144.000
Tosteson 1991	US	NA	ORT (in women without uterus): US\$ 7.010 – 9.020; without quality of life side effects 9.930 – 14.940 CRT (in women with uterus): US\$ 32.660 – 33.780; without quality of life side effects US\$ > 150.000
Visentin 1997	Italy	Cost per averted hip fracture: Calcitonin vs. no treatment: Population based approach: US\$ 2.367.987 Screening and treatment of high risk patients US\$ 838.120	NA
Weinstein 1980	US	NA	ORT in women with uterus: \$ 7.420-18.160 ORT in women with osteoporosis: \$ 5.460 – 15.100
Weinstein 1983	US	NA	ORT: \$ 130.000 vs. CRT CRT: \$ 42.000 (5 years treatment) to 24.000 (10 years treatment) vs. no treatment
Weinstein 1990	US	NA	ORT for 5 years: US\$ 72.100 in asymptomatic women ORT for 5 years: US\$ 12.600 – 33.100 in symptomatic women depending on symptom relief. CRT for 5 years dominates ORT for 5 years assuming quality of life differences are minimal CRT for 15 years: US\$ 22.650

CRT = oestrogen only therapy

HRT = ORT = oestrogen and progestin therapy

NA = Not applicable

CHAPTER 5

DEVELOPING AND DEMONSTRATING METHODS FOR ECONOMIC ANALYSIS IN AN OBSERVATIONAL PATIENT RECORD DATABASE.

5.1 Introduction

To health economists, patient record databases are currently most commonly known to supply partial data for transition probabilities in decision-analytic models. (Nuiten 1998) The search for methods by which economic evaluations may achieve a higher external validity has pointed to the use of clinical data that are reflective of clinical practice circumstances such as pragmatic trials (see chapter 3). Another potential source for such data are electronic patient record databases, that link patient demographic data and clinical data at patient level and contain information on the treatment of patients under local healthcare delivery conditions. Such databases may represent larger study populations than is normally available for clinical trials, and as a result, computerised patient record databases have been used as tools in epidemiological evaluations of drug safety. (Jick 1998) Such databases may also have the potential to constitute vehicles for full economic evaluation of costs and outcomes of pharmaceutical interventions, however this remains to be established.

At their best, routinely collected electronic data may have the potential to provide easily accessible data that can be identified and analysed in a short timeframe at reasonable costs relative to a prospective clinical trial. A formal feasibility study in the context of health economic research may provide valuable insight into whether such databases should be given further consideration as data source for full economic evaluation, either in the present format or with some enhancements. A study evaluating such feasibility should evaluate whether the database meets fundamental data requirements at satisfactory data quality.

This chapter outlines the process of evaluating the UKCPD database (UK Clinical Practice Database, Mediplus®) for provision of data to economic evaluation. A critical appraisal of the UK Primary Care Database (MediPlus® hereafter) is provided in section 5.1 before the aim and research objectives for a feasibility study are outlined. Section 5.2 outlines the study designs for two pilot studies that were developed to evaluate data quality and feasibility of the Mediplus® database in the therapeutic area of osteoporosis. Section 5.3 presents the results of the studies. Finally, the findings are discussed in section 5.4 in light of strengths and weaknesses of the database approach to increasing economic evaluation generalisability.

5.2 Background and study material

5.2.1 On observational data

Observational databases in the UK

There are many sources of observational data. Examples are national and regional registers and national statistics, such as activity data from Prescriptions Pricing Authority and clinical surveys, such as the morbidity survey by the Royal College of General Practitioners. Large healthcare administrative databases are common in the US and within other private, insurance-based healthcare systems for billing purposes. Generally, socialised healthcare systems, such as those in Europe, has offered less incentive for the healthcare provider to record activity at a detailed patient level, routinely collecting longitudinal patient data. To be of potential use in the economic evaluation of healthcare interventions, data should be in a form where exposure to health technologies can be identified at patient level, and where patient health can be followed over time. A number of databases constructed of computerised patient records with routinely recorded patient-level data are available in UK. (Lis et al 1997)

The vast majority of UK citizens are registered with a personal general practitioner (GP), who provides primary care and acts as a “gatekeeper” to secondary care. Increasingly, the GP uses a computer to record patient level data in healthcare, enabling commercial organisations, public health bodies and research centres in the UK to construct databases of computerised longitudinal information on large patient numbers. Examples of comprehensive UK-based patient record databases include the General Practitioners Research database (GPRD), the UK Primary Care Database (Mediplus®) and the Medicines Monitoring Unit (MEMO). These databases code clinical and administrative information systematically and store data electronically in a formatted record. Data from these sources are available for academic or commercial research.

The coverage of these three databases differs on parameters such as patient numbers and geographical area. Patient records are anonymously supplied to the GPRD and Mediplus® databases from selected practices across the whole of the UK. Data collection for GPRD has been operating since 1987 and currently, the patient records of approximately 2.7 million patients, equivalent to about 4.7% of the UK population, are recorded in the database (<http://www.gprd.com/>). The Mediplus® database has been collecting since 1992, and it contains the primary care records of 1.8 million patients. About 600 GPs in about 150 general practices in England, North-Ireland, Scotland and Wales contribute data and there are records of 1.8 million patients in the database. Needless to say, there is a constant turnover of patients, as some die or move area and so change their GP. Similarly, new patients join the practices on the data collection panel. So approximately 1 million patients in the database are alive and registered with the GP. (IMS Health 1998) MEMO (based in Scotland) accesses patient records of GPs from the county of Tayside with a patient population of about 400,000 patients (<http://www.dundee.ac.uk/memo/>). The MEMO database differs somewhat from the others in that it combines information from a variety of population-based datasets and is therefore complete for one particular geographical area. Furthermore, MEMO provides the additional opportunity to link primary care records with secondary and specialist care records.

The GPRD was set up primarily as a source for medical and public health research

purposes. (Rodriguez and Gutthann 1998) In contrast, Mediplus® was set up by a commercial agency to provide information on market research to the pharmaceutical industry. (IMS Health 1998) Common to the GPRD and Mediplus® databases is the routine assessment of the quality of information provided by the GPs. Only practices that satisfy given quality criteria are included in the research databases. The GP recording of information in the GPRD has been extensively validated (van Staa et al 1994, Jick et al 1991, Jick et al 1992) and the database has served as material for an extensive published literature on drug safety (Jick et al 1998). Although none of these databases were constructed with the purpose of supplying data for economic evaluation of pharmaceuticals, cost-effectiveness evaluations have been published using the MEMO database. (e.g. MacDonald et al 1995)

Longitudinal collection of demographic data, treatment choice and clinical diagnoses give opportunity to evaluate associations between interventions and patterns of care with clinical outcome. In the development of further methodologies to generalise economic evaluation adopting current methodology in the areas of economic evaluation, epidemiology and medical statistics are key to the evaluation of the merits of an observational database for this purpose. The Mediplus® database is the focus for this particular evaluation and its structure will be described in detail in section 5.1.2. However, prior to evaluating details of this specific database, general research methodology for the conduct of evaluations using observational data is provided.

Observational study designs and measure of association

In a randomised controlled trial (RCT) the process of randomisation to treatment or control group ensures that patients differ only by the play of chance, and the treatment allocation. This enables attribution of observed effects and provides good foundation for the statistical analysis of the trial results. Fundamentally, randomisation contributes to the development of an unbiased effect estimate of treatment.

The process of randomisation is not a feature in clinical studies with an observational design. There are two basic types of observational analytic investigation: the case-control and the cohort study. In a case control study, a case group or series of patients who have a disease of interest and a comparison group of individuals without the disease are selected for investigation, and the proportions with the exposure of interest in each group are compared. This design looks backward from a disease to a possible cause. In contrast a cohort study looks forward from an exposure to an outcome. Subjects are classified on the basis of the presence or absence of exposure to a particular factor and then followed for a specified period of time to determine the development of disease in each exposure group. Case control studies are always retrospective, whereas cohort studies may be conducted both retrospectively and prospectively. (Hennekens and Buring 1987)

The alternative design strategies offer advantages and disadvantages and the choice of either depends on the nature of the disease under investigation, the type of exposure and the available resources. The case-control design is particularly efficient for investigation of a relatively rare disease. The cohort design is best suited to investigations of relatively common outcomes that will incur in sufficiently large numbers over a period of follow-up of reasonable duration. (Hennekens and Buring 1987)

Clinical studies, whether randomised or observational, may report odds ratio (OR) or relative risk (RR) as the measure of association between exposure and risk of outcome. The OR and the RR express a comparison of disease frequency in the treatment and control group combined into a single summary parameter. The relative risk indicates how much more likely one group is to develop a disease than the comparison group by expressing the ratio of the proportions either population. *Table 5.1* shows the general layout of the 2 X 2 table that arises in this situation.

Table 5.1 General presentation of the results from a prospective study as a 2 X 2 table

	Exposure	Yes	No	Total
Outcome present	Yes	a	b	a+b
	No	c	d	c+d
	Total	a+c	b+d	n

The risk in the exposure group is $a/(a+c)$ and the risk in the control group is $b/b+d$. Relative risk is therefore:

$$RR = \frac{a/a+c}{b/b+d}$$

In case-control studies, where patients are selected on the basis of disease status (outcome), it is usually not possible to calculate the rate of development of disease given the presence or absence of exposure. The relative risk is not a valid estimate for this design, as any given value can be achieved by varying the number of cases and controls. The alternative method is based on calculations within each group. The ratio a/c is the odds of outcome in the exposure group. Equally, the ratio b/d is the odds of outcome in the control group. Combining these provides the odds ratio:

$$OR = \frac{a/c}{b/d}$$

If the outcome of interest that defines the cases is rare, then a will be small and $a/(a+c)$ will be approximately equal to a/c and similarly, b will be small and $b/b+d$ will be approximately equal to b/d . The relative risk will be approximately equal to the odds ratio. For case-control studies, the outcome of interest is usually rare, so the odds ratio offers a method of getting an approximate relative risk. (Altman 1997)

Analysis of epidemiological studies

The validity of observed associations between an exposure and a disease in clinical research can be assessed through consideration of whether the association is due to alternative explanations such as chance, bias or confounding. (Hennekens and Buring 1987) Statistical tools to assess this include hypothesis testing, estimation and statistical modelling.

Hypothesis test, estimation and the play of chance

The play of chance may always affect the results observed in a study, simply because of random variation from sample to sample. Hypothesis testing involves conducting a test of statistical significance and quantifying the degree to which

sampling variability may account for the results observed in a particular study. Denoted as H_0 , the null hypothesis represents the assertion that there is no relationship between exposure and outcome. The alternative hypothesis, denoted as H_1 , is the assertion that there is some relation between the exposure and the disease. The test statistic is a function of the difference between the values that were observed in the study and those that would have been expected if the null hypothesis were true as well as the variability in the sample. The significance test leads to a probability statement, or P-value, that indicates the likelihood of obtaining the study result by chance alone assuming that there is truly no association between the exposure and outcome under consideration (i.e. H_0 is true). The level of the p-value that indicates an association is statistically significant is arbitrarily determined, but by convention in medical research this is set at 0.05. Thus, the H_0 is rejected if the P-value is less than 0.05, as this is taken to mean that chance is an unlikely explanation of the findings. The expected value of the RR and OR is 1 under the null hypothesis. The most frequently encountered hypothesis tests in medical literature are the chi-square test for discrete data and the t-test for continuous data. (Altman 1997)

The P-value reflects both the magnitude of the difference in effect between the groups and the sample size. One problem inherent in the interpretation of the P-value is that even a small difference may be statistically significant if the sample size is sufficiently large, and conversely a larger effect may not achieve statistical significance if the sample size is insufficient. Increasingly therefore, estimation and confidence intervals are used to draw inference in clinical research. The confidence interval, or the range within which the true magnitude of effect lies with a given degree of certainty, is a related but more informative measure of the role of chance than the P-value. (Gardner and Altman 1989) The confidence interval provides the information of the P-value in terms of whether the association is statistically significant at a specified level. For example, if the null value (e.g. 1.0 for the relative risk estimate) is included in a 95 percent confidence interval, then the corresponding P-value is by definition greater than 0.05. The advantage of estimation over hypothesis testing is that the width of the confidence interval indicates the amount of variability inherent in the estimate and therefore expresses the precision of the estimate in addition to the effect of sample size.

Statistical modelling and the role of bias and confounding

In a randomised clinical trial, observed and unobserved influences (prognostic factors) are randomly allocated to trial participants and potential differences in trial outcome between the exposure group and the control group may therefore be attributed to the exposure (i.e. treatment) or the play of chance. This is not true for observational data, where factors beyond the control of the researcher influence the selection to the groups. Basic hypothesis tests may therefore be insufficient and indeed misleading for analysis of data produced by observational studies. The use of statistical modelling techniques that take observed differences between the study groups into account and that may bias or confound the findings is therefore essential when drawing inference of comparative effectiveness.

The aim of statistical modelling is to build a parsimonious mathematical representation in which an outcome variable is expressed as a combination of one or more variables that may modify the effect of exposure. Variables are included to a model to explore whether they can explain the observed variation in outcome between individual observations in a study sample. The analysis provides coefficients and standard errors for the coefficients from which the statistical significance of the explanatory variable may be calculated. Statistical modelling, or multivariate analysis, enables the exploring of the impact of several potential explanatory variables individually and in any combination. (Altman 1997) A large number of multivariate models have been developed for specialised purposes, each with a particular set of assumptions underlying its use. The choice of model is based on underlying study design, the nature of the variables and assumptions of interrelationship between the exposure and outcome.

The simplest approach to analysing the relationship between the values of explanatory variables and continuous outcomes is by the use of linear regression models. To fit a regression line the standard method is to minimise the sum of squares of the distances of the observations to the line and obtain an ordinary least squares regression (OLS). The general equation of a regression line is

$$Y = a + bx$$

Where Y is the outcome variable, a is a constant corresponding to outcome when explanatory variables take the value of zero, b is the regression coefficient and X is the explanatory variable. The regression coefficient produced by the analysis expresses the variability in the response variable that can be attributed to different values of the predictor variable, e.g. the increase in outcome for each unit increase in the explanatory variable. (Altman 1997) Underlying such analyses are the assumptions that the values of the outcome variable have a normal distribution for each value of the predictor variable, and that the relationship between the two variables should be linear.

The simplest approach to modelling binary outcomes is logistic linear regression. For such outcomes an appropriate regression model would predict the proportion of individuals with the characteristic for a combination of explanatory variables. The probability of individuals having a certain binary outcome (characteristic) is transformed on to a logit scale for the reason that impossible probabilities outside of the range 0 to 1 can not be predicted on this scale. Here, the proportion of individuals with the characteristic is p , and $1 - p$ is the probability that they do not have the characteristic. The odds of the characteristic is therefore

$$\text{Logit}(p) = \log_e [(p) / (1-p)]$$

When comparing the odds of having a particular characteristic in two groups (p_1 and p_2) the general equation is

$$l_1 - l_2 = \text{logit}(p_1) - \text{logit}(p_2) = \log [p_1 / (1-p_1)] - \log [p_2 / (1-p_2)]$$

which is the log of the odds ratio (OR) of an outcome. Therefore, the odds ratio for an explanatory variable in logistic regression (i.e. the degree to which the risk of outcome changes as a consequence of having the attribute) can be obtained directly from the regression coefficient. An underlying assumption of the logistic regression is that the relation between the dependent variable and each continuous explanatory variable is linear. Furthermore, it assumes that the effects of each variable are independent of each other.

Some studies aim to compare the time to outcome between two patient groups. The data from such a study will define a time origin and the time of occurrence of some particular outcome or each patient. The observations for those patients that do not experience outcome by the end of follow-up are censored, and the patients have varying lengths of follow-up time during which an outcome may occur. The analysis of data that describes the time to a particular event is called survival analysis, or Cox proportional hazards regression analysis (Collett 1994). The hazard ratio provided by is Cox regression analogous to the relative risk, however it incorporates the aspect of time. A strong assumption underlying this analysis is that of proportional hazards, which means that the impact of the different variables on outcome is constant over time. No particular distribution is assumed for the survival times, but a strong assumption is made that the effects of the different variables on survival are constant over time and additive in a particular scale. (Collett 1994)

Linear regression models for binary and continuous outcomes are based on assumptions that may not hold for many types of data. For example, it may not always be reasonable to assume that data are normally distributed when modelling counts or proportions. The generalised linear models (GLMs) extend the traditional linear models and are applicable to a wider range of data analysis problems, particularly when dealing with multiple explanatory factors simultaneously. Generalised linear models constitute a family of linear models that allows the outcome (the “linear predictor”) to be calculated from a non-linear link function and allows the response probability distribution to be any member of an exponential family of distributions (McCullagh and Nelder 1989). Many widely used statistical models fit into the family of generalised linear models. These include classic linear models with normal errors, logistic and probit models for binary data and log-linear models for multinomial data. Binary data, continuous data and counts can all be analysed by the same model class.

The GLM is specified by selecting an appropriate link function, which defines the relationship between the observations and the linear predictor, and a response probability distribution. For binary outcomes where the linear predictor (response) is a probability, the link function is a logit link. For data where the linear predictor

presents means (i.e. the variable is continuous) the link function is expressed as an identity function. For data where the linear predictor is counts, the link function is a log link. The error structure expresses the distribution of deviation from expectation of the model. For binary outcomes the error structure is binomial. For continuous outcomes the error structure is normal, and for counts the error structure is specified as a Poisson distribution.

Prognostic variables can be included in models in a mechanistic way by using a forward or backward selection approach. (Altman 1997) In *forward regression* the relationship between each potential explanatory variable and the outcome variable is examined. The variable that has the strongest association with the dependent variable will be entered into the model as a factor. Subsequently, the variable amongst the others that explains the largest amount of the remaining variability is then added to the model. Variables are added until the addition of an extra variable is not statistically significant at a chosen level (usually $p < 0.05$). In *backward regression* all those variables that *a priori* were considered potentially important explanatory variables are included in the model to start with and then unimportant variables are removed one at a time until all those variables remaining in the model contribute significantly at a chosen level. Backward regression approach has the advantage that it increases the patient material for which it bases its calculation for each variable that it removes. Consequently the power to detect significant variables increases as the model is reduced. The disadvantage comes when the impact of several potential explanatory variables is explored and the potential for multiple significance testing and false positives arises. To minimise this problem, only factors that are plausible confounders should be included in a model selection exercise.

The aim of modelling is to arrive at a regression that fits the data reasonably well. The model fit is described by the residual deviance between the observations and the regression line. A modelling exercise therefore always seeks to reduce the residual deviance to find the best predictive model. An integrated part of a modelling exercise is to assess how well a statistical model fit the data by considering the proportion of the variation in outcome that can be explained by the variables in the model. Model fit in regression models can be explored by

examining two values, the statistic R^2 and the residual deviation. R^2 is the proportion of the total sum of squares of all included covariates that can be explained by the regression. This can be expressed as a percentage and will crudely assess the degree to which the model fits the observations in the dataset.

The value of R^2 is expected to increase as the number of variables in the model increases, even though each added variable does not relate to the dependent variable (Altman 1997). Exploring the residual deviation is therefore essential. A residual is the difference between an observed and a fitted value and each observation will have a residual. The variance of the residuals produced by a regression line expresses the amount of variability ('random variation') that is unexplained by the covariates in the fitted model. The variance and standard deviation of the residuals is therefore a measure of 'goodness-of-fit' of the line produced by the model. (Altman 1997)

The residuals are normally distributed for a well-fitting model of continuous outcome. Therefore, the residuals produced by generalised linear models with normal error may be explored visually by producing a plot of the residuals, for example against the fitted values or against each of the explanatory variables. For models with binary and Poisson error however, there is no requirement for the residuals to be normally distributed. When assessing model fit for such models the issue of overdispersion becomes more important. A measure of overdispersion is obtained by dividing the value of the residual deviance by the degrees of freedom. When the model fits well then these are expected to have the same value.

5.2.2 Critical appraisal of the UKPCD (Mediplus®) database

The European Health Outcomes Research team at Eli Lilly and Company Ltd. (Windlesham, Surrey) funded the project over the course of which the Mediplus® database was evaluated. There was increasing interest in the conduct of observational studies to evaluate real-life cost-effectiveness of the drugs in clinical practice. The Health Outcomes team in the UK affiliate of Eli Lilly (Basingstoke,

Hampshire) had access to the Mediplus® database through a continuous subscription with IMS Health. Combined with an academic interest in exploring the use of observational databases for economic evaluation, this provided joint research interest in evaluation of Mediplus® for such analysis. The evaluation process included choice of a research area of interest, a critical review of published literature on the database to assess a potential for evaluation and finally exploring of practical features of the data that would eventually feed into a specific study design.

Identifying a therapeutic area worthy of further research

At the time when the study was initiated, Eli Lilly and Company had recently launched a pharmaceutical treatment in the therapy area of osteoporosis. The company was interested in planning a future evaluation of the drug in a clinical practice setting in the UK. Two studies were planned. The first study would be a pilot study evaluating the ability of the database to provide data to a full economic evaluation in the therapeutic area of osteoporosis. The second study would include the drug marketed by Eli Lilly and Company in the same therapeutic area.

The chosen subject of the study was the cost and consequences of the use of osteoporosis medications in UK clinical practice. A number of economic evaluations in the therapeutic area of osteoporosis have been published. (Torgerson et al 1997) These economic evaluations have to date predominantly been based on decision-analytic models, and these were reviewed in Chapter 4 of this thesis. (Sculpher et al 1999) Health economic models are constructed in order to predict the impact of a therapy under circumstances of uncertainty. Clinical and economic assumptions incorporated in such models may or may not reflect the clinical practice reality in which the pharmaceutical eventually will be used. For example, lower compliance rate in clinical practice may adversely affect the effectiveness of the therapy when routinely used outside of a trial. The review in Chapter 4 revealed that many models assume full compliance with therapy over several years, and omit exploring the impact of reduced compliance on the cost-effectiveness ratio. Furthermore, costs that are relevant to the population of patients in clinical practice, such as cost of side effects of treatment, was not

incorporated in the previous economic evaluations. Finally, in previous economic evaluations, the fracture risk reduction has been modelled from the change in BMD at fracture-susceptible sites from trials, or based on assumptions with basis in epidemiological studies (see chapter 4). (Tosteson 1991, Weinstein 1980)

Three issues were therefore of particular interest to the use of clinical practice data in this therapy area. Firstly, the primary clinical outcome in registration trials of osteoporosis therapies is frequently bone mass density (BMD), a measurement that is widely recognised as a risk factor for osteoporotic fractures. (Marshall et al 1996) For the purpose of economic evaluation however, hip fracture is a more relevant outcome because of the costs and human suffering associated with them (Dolan and Torgerson 1998). It is therefore of interest to an economic evaluation to assess the relative effectiveness of osteoporosis medications to preventing hip fractures as manifested in clinical practice.

Secondly, the effect of preventive medications for osteoporosis may rely upon the duration of treatment. Economic models are frequently based on the assumption that medication is used over a period of several years, for example 15 years in a study by Tosteson et al. (1990). In clinical practice, however, some patients cease to take medication before the preventive effect is experienced. It was therefore of interest to identify the degree to which patients cease to take pharmaceutical therapies in osteoporosis in clinical practice.

Finally, the use of the bisphosphonates in treatment of osteoporosis (alendronate and etidronate) has been associated with gastrointestinal conditions (Ettinger et al 1998^a, Ettinger et al 1998^b). Gastrointestinal conditions such as dyspepsia and perforation are frequently treated with drugs such as the H2 antagonists (H2s) and the proton pump inhibitors (PPIs). The costs of treating these side effects have not previously been assessed in the UK clinical practice setting.

Together, these provided rationale for further study and good background for the design of a feasibility study for which the intention was to explore these issues and arrive on recommendations for future use of the database for economic evaluation studies.

Data review

The availability of observational databases and evaluation methodology provide material and methods on which to build when aiming to develop further methodology to increase the external validity of economic evaluations. There is indication that the Mediplus® database potentially provides a wealth of clinical practice information. What are the fundamental data requirements for an economic evaluation in primary care, and is Mediplus® likely to meet these data requirements of contents, quality and structure? Prior to analysing data from a specific study in the database, literature covering the contents and construct of the database was reviewed. Through this review, the coverage of the data was assessed, the risk of contamination during data entry procedure was appraised, the data structure was reviewed, and the data retrieval procedures were evaluated.

Specific requirements for economic evaluation data depend on the type of analysis undertaken and its sophistication. The problem in question, the institutional framework and the perspective of the analysis will determine the data requirements. Fundamentally, patient level resource consumption in the sectors relevant to the perspective of the analysis must be identified, measured and valued. Similarly, clinical consequences at patient level must be identified and measured. (Drummond et al 1997) Furthermore, data requirements vary profoundly according to the disease area, whether the condition is prevalent in primary or secondary care, whether the condition is chronic or based on treatment episodes and whether it is a physical or mental disorder. A study undertaken from the perspective of primary care should collect costs that occur in the primary care sector. These include cost of study medication and relevant concomitant medication, healthcare activity costs such as GP visits and nurse visits and cost of clinical tests and assessments. The relevant clinical outcome needs to be identified, and the timing of outcome may be of importance. Finally, a key principle of clinical and economic evaluation is the presence of a control group either in the form of 'no treatment' or alternative intervention. (Pocock 1984)

Fundamental to a study aiming to evaluate context specific effectiveness of a therapy is the inclusion of subjects who reflect the population to which results intend

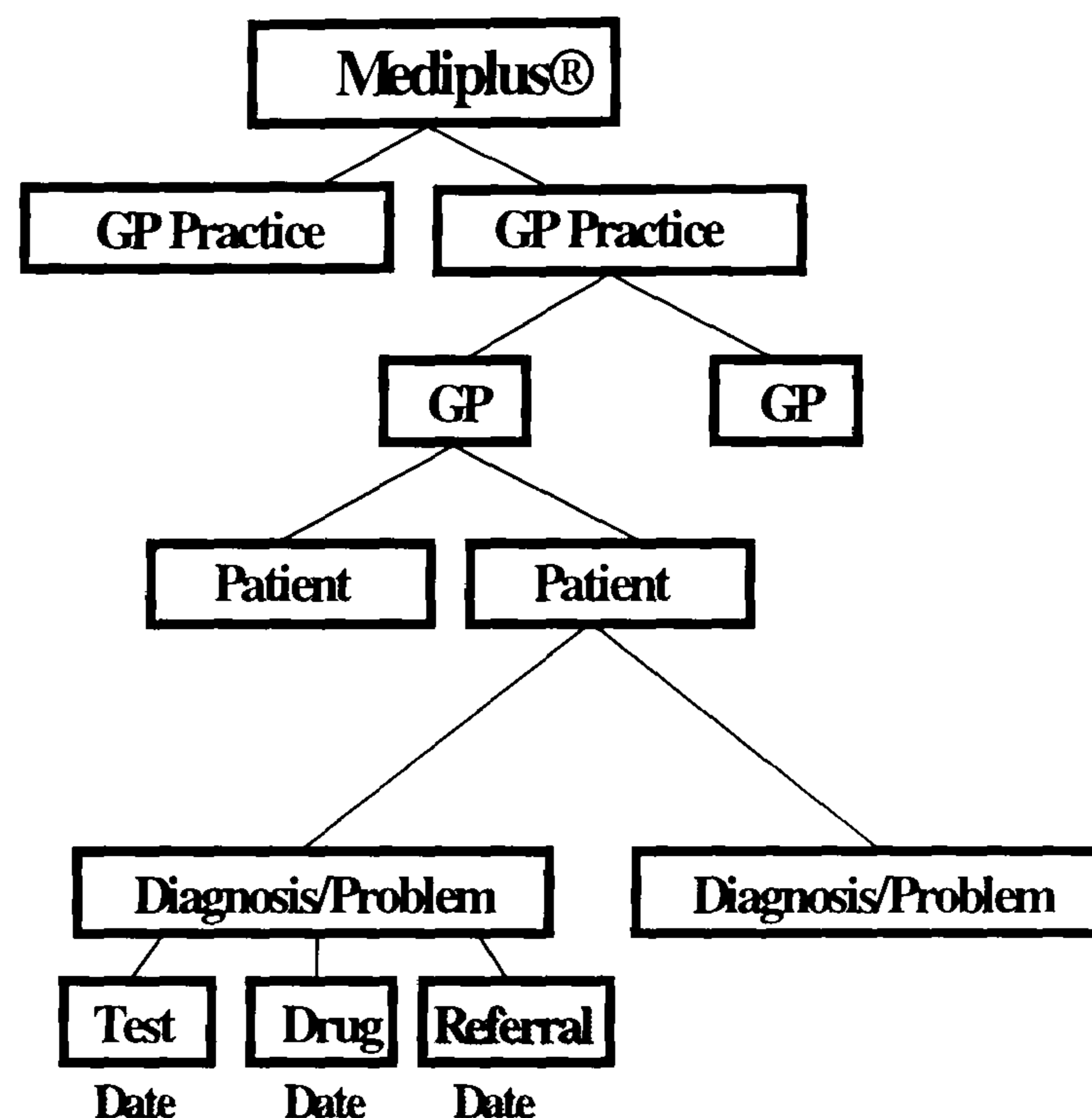
to be extrapolated. Judging by demographic information, the Mediplus® database covers a population of GPs and patients reasonably reflective of UK general practice. Deviations include age distribution of the GPs on the data panel, geographical representation and size of the GP practices. A larger proportion of younger GPs (<45 years) and fewer older GPs than in the UK as a whole are included on the panel. The areas of Northern and Yorkshire and South and West are over-represented and Scotland and Northern Ireland underrepresented. A greater proportion of participating practices are large (6-9 partners) compared with the UK average. (IMS Health 1998)

Is the recording of data likely to provide valid and accurate information? Monitoring of data recording is a key part of the conduct of randomised clinical trials. (ICH 1996) Data recording regular activity in clinical practice, where monitoring procedures do not exist, may be more susceptible to contamination by human error. The general practices on the data collection panel for Mediplus® use computers with software provided by IMS Health. A data file of the daily clinical record entries is created in each practice and transmitted by modem to IMS daily. General practitioners were included on the Mediplus® panel after a trial period of data entry and quality training. There is a regular audit of recording at practice-level with feedback at quarterly intervals and requirements to meeting predetermined standards of data quality. The data quality standards cover areas such as patient coverage and prescription details. For example, the IMS minimum requirement for percentage of prescriptions entered onto computer is 80%. An audit in 1998 revealed that 98% of practices met the requirement. (Lawrenson et al 1998) Practices who consistently fail to reach the minimum standards are eliminated from the panel. (IMS Health 1998) The quality audits are in place to assure a minimum quality standard. However, contrary to conditions of an ideal epidemiological study, (Jick et al 1998) the Mediplus® system does not enable source validation of the outcomes against written clinical records or original hospital records.

Does the construction of the data in Mediplus® enable identification of the essential costs and outcomes for economic evaluation study, such as demographic data, resource consumption and clinical conditions? The recording of gender and age is complete in the database. For reason of confidentiality, a virtual birthday is

assigned at January 1st in the year of birth for all patients in the database. This enables identification of patients by age and sex. Indicators for resource consumption items in the database include medication use and cost, GP activity and referral and admissions. Drug prescriptions are recorded by generic name and trade name as well as by code in the Anatomical Therapeutic Chemical (ATC) international system for classification of drugs. The cost of the issued amount of medication is calculated in the database as a function of pack size and MIMS price (Monthly Index of Medical Specialities). The date at which the prescription was issued, either by the GP or the practice nurse through repeat prescription, is recorded and this may provide indication of GP activity. Clinical conditions are recorded in the Read code system. (Chisholm 1990) The construction of Mediplus® patient records is based around clinical conditions and all activity undertaken is linked to a condition. The providers call this “problem based” record keeping, where “problem” equals clinical condition, provided valid and accurate recording. (IMS Health 1998) In theory this would enable evaluation of all aspects of care being evaluated as a function of a clinical condition. For example, a patient with a diagnosis of asthma would have a series of prescriptions issued, clinical assessments undertaken and perhaps also referrals to specialist care. All these activities should be linked to the condition of asthma. Similarly, all activity undertaken in association with a condition of hypertension on the same patient would be linked to hypertension rather than asthma. The construct of the database and the ‘problem based’ records is illustrated in *Figure 5.1*.

Figure 5.1 A simplified diagram showing the construct of the Mediplus® database and how levels of information can be linked.



Data collection, the quality audit, and the management and upkeep of the Mediplus® database is administered by IMS Health, a commercial organisation that specialises in the provision of market information to healthcare industries, such as the pharmaceutical industry. (<http://www.imshealth.com>) Customers can buy data or full analyses from the Mediplus® database prepared by researchers at IMS Health for a fee. The data fee is calculated on the basis of the number of patients and number of variables included in the datafile. A researcher at IMS Health conducts the process of extracting data from the masterfile of the whole database to a datafile containing the information specified for any specific study. Once the datafile is extracted from the main database there is limited opportunity to go back and seek more information or review whole patient records. Similarly, there is limited opportunity for the external customer wishing to undertake the study to gain insight in the data extraction step.

Approval to conduct an analysis for publication must be gained prior to the work starting through submitting a protocol for review by an Independent Scientific and Ethics Advisory Committee. Any manuscript to be submitted for publication should

be made available for review by IMS Health for comments at least 4 weeks prior to submission. IMS Health then comments on citation but does not 'approve' or disapprove' publications. Seventy-five titles have been published from the Mediplus® database; 27 of these are study results published in scientific journals, and the rest were predominantly abstracts and posters. A publication list from IMS Health 2001 can be found in appendix 4. There are examples of drug utilisation studies (Jones et al 1995) and epidemiological studies (Farmer et al 1997, Farmer et al. 1998). The Mediplus® database has not previously provided data for a full economic evaluation.

Towards an evaluative design

The access to full patient records of more than a million patients appear to be a rich opportunity for the study of clinical effectiveness and cost-effectiveness of pharmaceutical interventions in clinical practice. The review of literature and exploring of data construct and recording revealed that the Mediplus® database may have a potential for generating data for a full economic evaluation of osteoporosis drugs. A feasibility study adopting a rigid design and clear objectives would clarify remaining issues surrounding data contents and quality and enable assessment of whether the Mediplus® database satisfies basic data requirements for the conduct of economic evaluation of pharmaceuticals and explore whether the available data are appropriate for such analysis.

The Mediplus® database provides, in theory, opportunity for both retrospective cohort study and case-control design. For economic evaluation, where the cost-effectiveness of one or more therapies is evaluated, the cohort design is more appropriate as this is based on inclusion according to exposure rather than outcome. Therefore, the conduct of a retrospective cohort study was planned.

The decision to include patients who received the six most prevalent treatments for postmenopausal osteoporosis in the UK was made. The bisphosphonates Alendronate (Fosamax™) and cyclical etidronate (Didronel PMO™) were included. Two hormone replacement therapies, the so-called non-bleed HRT combinations, were included. These contain conjugated oestrogen/medroxyprogesterone

(Premique™) and oestrogen/norethisterone (Kliofem™). For simplicity the HRT therapies were called HRT(P) and HRT(N) throughout the report. Finally, tibolone (Livial™) and raloxifene (Evista™) were included, where the indication for prescription at the time of study was prevention of osteoporosis. (BNF 2000)

Two studies were planned. The first study was a feasibility study of the potential for Mediplus® to enable economic evaluation of therapies in the treatment of osteoporosis. The second study was a reiteration of the feasibility study at a later time to evaluate effectiveness of raloxifene at a time when a larger number of patients had initiated therapy on the drug. Aspects of the Mediplus® database of strategic importance to the design was explored at the database terminal at the UK affiliate of Eli Lilly and Company at Basingstoke, Hampshire and IMS Health headquarters in Pinner, London.

Patients

The Mediplus® database enables identification of patients by gender, age and medication use. The purpose of the pilot studies was to evaluate an association between exposure to selected pharmaceutical interventions and clinical effectiveness and resource use. Therefore, patients were included in the studies on the basis of the medication that they had received. It was possible to identify medication use through searching by relevant Anatomical Therapeutic Chemical (ATC) codes.

The inclusion of patients in the control group according to the condition of osteoporosis was done in a similar database analysis in the GPRD database. (van Staa et al 1998) As outlined earlier, the “problem based” recording in Mediplus® should in theory enable the identification of a subset of women that have the diagnosis of osteoporosis but are not prescribed any of the study interventions. It did however become apparent in the exploring of the data that this was unfeasible. This was illustrated by two findings in an exploratory analysis at the terminal. Firstly, the bisphosphonates are almost exclusively prescribed in treatment and prevention of osteoporosis. However, for many women prescriptions of the bisphosphonates were issued under different clinical problem heading, for example

“Asthma” or “Hypertension”. Secondly, it is reasonable that in an activity driven record only actions taken against a condition would trigger the recording of the conditions. In other words, a patient may well be at risk or suffer from a condition, but unless any action is taken by the GP no recording of the condition is likely to be made.

Clinical outcomes

The incidence of hip fracture was primary clinical endpoint for the economic analysis. A Read code (Chisholm 1990) for hip fracture was available to indicate the presence of outcome in the patient records. Duration of treatment in clinical practice was also evaluated. Attrition is a key element of treatment success and difference in patients’ adherence to the drug regime may be an indication of which drug is more tolerated by the patients. The event of premature cessation of therapy is not routinely recorded in the Mediplus® database. Therefore, a stop date was imputed based on the duration of the medication supplied at last prescription of study drug and the date of prescription. This enabled estimation of treatment duration by subtracting start date from stop date in the data analysis step.

Resource use

Acquisition cost for the prescriptions of index drug can be extracted directly from the database. The cost of medication quoted in Mediplus® is the manufacturers listed price as published in MIMS (IMS Health 1998) and is identified at chemical substance level (ATC level).

There is a formula incorporated in the database design that estimates number of GP visits based on number of entries in the patient record. Entries to the patient record can be made by telephone contact to the surgery, and at repeat prescriptions, when results from tests reach the practice, etc. All entries are dated, but more than one entry can be made on one day. Therefore, the number of dates entered can not be used as a direct estimate for GP visits. In order to see whether the Mediplus® formula gives a realistic presentation of GP visits it was decided to collect information on the number of dates entered and compare this with the estimated visits.

Referrals to specialities and hospital admissions are recorded in Read codes and are identified in the notes attached to patient entries (see *Figure 6.1*). There is a date attached to every entry, and consequently it is possible, in theory, to quantify the number of referrals in a patient record for the purpose of economic evaluation.

5.3 *Aim and objectives*

The decision to undertake an evaluative study in the Mediplus® database developed from the research need to develop further methodologies to provide economic evaluation data with high external validity. At the same time, Eli Lilly and Company was interested in making best use of current databases to provide information of real-life cost-effectiveness of one of their pharmaceuticals. The clinical rationale for study combined with the availability of data led to the formulation of specific research objectives for a project aimed at assessing the feasibility of Mediplus® for economic evaluation.

The feasibility project was planned in August 1999. Raloxifene (Evista™) was registered in the UK in September 1998 and only very few patients had been exposed to the therapy at the time that the study was planned. Therefore, two studies with separate objectives were planned. The **first** study would evaluate the feasibility of the database for full economic evaluation of pharmaceutical interventions and explore this through a pilot study in the therapy area of osteoporosis. The **second** study was developed after the analysis of the first study was completed at a time that more patients had been exposed to Raloxifene. The second study would include those components of a cost-effectiveness evaluation that the first study deemed feasible, and apply those to raloxifene and a comparison group

As section 5.5 will show, only certain variables in the Mediplus® database satisfied the information need and quality standards for the conduct of an economic evaluation. Design of the second study was dependent on the first study.

Consequently, only those variables from the first study that met the data requirements were used as input to the design of the second study, and this is reflected in the objectives and the study design of the second study.

Aim

To evaluate whether the potential wealth of information contained in the UK Primary Care Database (IMS Mediplus®) make it a feasible tool for a full economic evaluation, including evaluation of treatment patterns as well as clinical and economic outcomes of treatment.

Objectives for the First study

- To evaluate the incremental costs and effectiveness associated with the use of therapy with tibolone, alendronate, etidronate, HRT(P) or HRT(N) (study drugs) in women over the age of 55 in UK general practice setting (study population) between January 1st 1996 and December 31st 1998 (study period).
- To assess resource consumption in primary care in the study population associated with the use of the study drugs during the study period.
- To assess the risk of hip fracture and its determinants associated with initiation of treatment with the study drugs in the study population during the study period.
- To assess the extent to which the study population discontinue treatment with the different study drugs during the study period.

Objective for the Second study

- To assess and compare drug use patterns and costs associated with the use of raloxifene and the bisphosphonates in UK clinical practice.

5.4: Methods

5.4.1 Design of first study: Pilot study of osteoporosis drugs in Mediplus®

Data extraction

The UK Primary Care Database (UK MediPlus®) was used to derive information from computer records of 1 million active patients from approximately 140 general practices throughout the UK. A SAS analyst and researchers at the IMS Health (UK) main offices in Pinner, London, developed the analytical files according to a detailed design document. Five analytical files were created and linked by a unique patient number. The files contained observations on predetermined variables describing demographic data, therapy progress, concomitant medication, referrals and clinical events respectively. The statistical variables for the first study are provided in appendix 5.1 through 5.5.

Patients

Inclusion and exclusion

Patients were identified from the computer records according to the following inclusion criteria:

- Patients received a first prescription ('index prescription') of either of the study drugs (see table 1) during the period of January 1st 1996 through December 31st 1998.
- Patients were women at the age of 55 years or older when the index prescription was issued.
- GP practice to which the patient belonged was active in the database in December 1999.
- Patients were active in the database in December 1999.
- Patients were included in the study provided that they were active in the database 9 months or more before the index date. This criterion was developed to ensure that all patients had a previous recording history in the database of at least 9 months.

- The patients had a period of 6 months prior to the index date during which no drug with an impact in osteoporosis was issued.

Patients were excluded based on the following criteria:

- Patients who had a prescription of two different study drugs on the index date.
- Patients who appeared in more than one study cohort (1996, 1997 and 1998).

The patients included in the control group were matched to the exposure group by age and gender and were included on the basis that they were not exposed to the selected medications. Two patients were included in the control group per patient initiating therapy with study drug. Two reasons explain the choice of a 2:1 ratio. Firstly, the providers of the database, IMS, provide a data charge that is dependent on the number of patients included in any given study. Limited budget would therefore encourage an economical use of the data. Secondly, there is little statistical efficiency to be gained by including more patients in the in the control group. The variance of the effect estimate, i.e. the probability of events in the intervention group less the probability in the control group, cannot be reduced by reducing the variance of the effect estimate in the control group only. The general formula for the standard error for the difference between proportions is

$$SE_{diff} = \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

For a study with 1000 patients in the intervention group, the probability of event in intervention group 0.7 and the probability of event in the control group 0.3, the SE_{diff} would be 0.022 for a 1:1 ratio inclusion of patients to the control group. When including 2000 patients in the control group in a 2:1 ratio, the SE_{diff} would drop to 0.019. The value of the SE_{diff} would be 0.017 for 5000 patients in the control group, and indeed stay 0.017 for inclusion of 10.000 patients in the control group. Thus, the return in terms of greater precision in the estimate by including more patients in the control group diminishes rapidly beyond the 2:1 ratio.

Study medication

Hormone replacement therapy (HRT) is more frequently used in a younger population of women for treating menopausal symptoms rather than as prevention strategies for osteoporosis. The inclusion of women over the age of 55 aimed at excluding those women that were prescribed HRT for postmenopausal symptoms. Indeed, for women older than 60 years, the main reason for using HRT is for fracture prevention. (Ettinger 1999) The bisphosphonates are on the other hand indicated in a population of women with a high risk of osteoporosis or where the condition already is manifested (BNF 2000).

The Mediplus® database is constructed such that GPs can record drug prescriptions either by trade name, generic name or by chemical substance. Identification of medication was therefore based on both product name and substance name to ensure inclusion of all eligible patients (see *table 5.2*). Etidronate and HRT(N) were exemptions as the same substances are shared with different products. Drug dose and form of administration was not considered.

Table 5.2. Study drugs of interest to the evaluation

Substance	Product	Basis for selection
Tibolone	Livial™	Substance + Product
Alendronic acid	Fosamax™	Substance + Product
Conjugated oestrogen and medroxyprogesterone (HRT(P))	Premique™	Substance + Product
Etidronate	Didronel PMO™	Product (multi-substance) ¹
Oestrogen and norethisterone (HRT(N))	Kliofem™	Product ²
¹ The same substance is used in a formulation used in a different therapeutic area. ² The same hormone combination is shared by a different product, Elleste Duet Conti™.		

Women in the HRT group and the bisphosphonate group were analysed separately throughout. Bisphosphonates are indicated as treatment of manifest osteoporosis, whereas the HRT drugs and tibolone raloxifene are likely to be prescribed as osteoporosis prevention. These cohorts are likely to differ in terms of prior susceptibility to fracture and comparing fracture risk between these cohorts would therefore be inappropriate. To simplicity analysis the controls of the tibolone patients were grouped with those of the HRT therapies. Importantly however, tibolone have different pharmacological effects, side effects and clinical benefits

than the HRTs. (Royal College of Physicians 1999) The control patients were grouped in one control cohort matched with the HRT patients (including tibolone) and one control group matched with the bisphosphonate patients.

Baseline characteristics

The Read code system by which clinical conditions are recorded in Mediplus® was used to develop definitions for a number of baseline characteristics. A set of Read codes to indicate selected conditions was developed by the researcher (HU) and a researcher at IMS Health UK (see appendix 5.2). The patient records were screened for the presence of a number of clinical conditions recorded until 9 months prior to the date at which the first prescription of the study medication was issued. These were Asthma, Pulmonary disease, Respiratory system disease, Rheumatism, Arthritis and Gastrointestinal (GI) disease. Some chronic conditions were searched 5 years back in the patient records. These were recording of Osteoporosis, Hysterectomy, Breast cancer, Ovarian cancer, CHD, Stroke, MI, Cerebrovascular disease and Hypertension.

Follow-up

Follow up was 1, 2 or 3 years depending on which time the first prescription was issued. If the prescription was issued in 1996 the follow-up time was 3 years, if the prescription was issued in 1997 the follow-up time was 2 years and similarly the follow-up time was 1 year for those included in 1998.

Outcomes

Definition of clinical outcome

A set of Read codes to categorise clinical outcome was developed by the researcher (HU) and an analyst at IMS Health (UK) (See appendix 5.2). The patient records were searched for the recording of either of the indicators for each clinical outcome.

Hip fracture

A set of clinical Read codes was developed to identify the recording of hip fractures in the database. These are provided in appendix 5.2. Occurrence of hip fracture was recorded in the original datafile as **number of** fractures recorded in the patient records and **date** at which the fracture was recorded. If the code occurred at several dates then this was extracted into the datafile as a new fracture. If two dates recording the same code for outcome were within 3 months of each other they were considered duplicate recording of the same fracture and only counted once.

Treatment pattern

The day at which the first prescription was issued was assigned the 'Index date'. This date indicates inclusion in the study. "Stop date" was defined as the date at which no further study drug issued in the 6 months following the end date of the duration of the last prescription of drug. A sensitivity analysis was undertaken to assess the impact of changing the definition of "stop date" to a drug-free time period of 3 months.

Resource consumption

The preliminary exploring of the Mediplus® database described in section 5.1 indicated that it may be possible to generate resource consumption data on GP visits, drug acquisition costs and cost of concomitant medication relevant to the therapy. Furthermore, items that were collected to describe the resources consumed in secondary care were referral rates to hospital and referral to various specialists. Two alternative ways of expressing GP activity at patient level were extracted from the database:

- The number of GP visits computed by the providers of the database
- The number of dates entered in the patient records

These two indicators were compared in an attempt to identify whether the database providers' definition is an appropriate estimator of GP activity. The results of this comparison indicated the plausibility of the computed measure of GP activity, and therefore the reliability with which GP visits can be estimated in Mediplus®.

The cost of drug treatment was extracted directly from the database. The ATC code for the relevant study medications were identified. The cost for the quantity of drug that was issued to the patients in the 6 months period before the first prescription of study medication was extracted. Similarly, the cost of the quantity of study medication in the first, second and third year after the first prescription of study medication was extracted.

The costs of prescriptions issued for other medications were extracted in a similar manner to the acquisition cost of study medication. Information was extracted on the following drugs: antacids, H2 antagonists, proton pump inhibitors, multivitamins and minerals (ATC A11A and A11B only), mineral supplements (Calcium, vitamin C and D), laxatives, sex hormones, corticosteroids, antirheumatic agents, non-narcotic analgesics (N02A and N02B only) and bronchodilators.

Referrals to the specialists listed below were identified through the Read code system and quantified by the number of dates entered for each referral type. Referrals to hospital were identified and quantified the same way. The referrals recorded were hospital referral (A&E) gynaecologist, orthopaedic, rheumatologist, gastrointestinal investigation, physiotherapist and radiologist.

Statistical analysis

All analysis was conducted using SAS for windows release 6.12 and 8.0 (SAS Institute, Cary, NC). A number of simple checks of the data were done prior to analysis to investigate the presence of outliers, implausible values and missing values. This included creation of frequency tables of class variables and calculation of means, standard deviation and range of numeric variables. The datasets were also viewed on screen for assessment of face validity. The design protocol was compared with the datafiles to confirm the presence of all variables. It was not possible to undertake source verification of the occurrence of outcomes such as hip fracture through access to written patient records. The Mediplus® system does not enable such checks.

Modelling hip fracture risk

The primary outcome of the pilot study was hip fracture. Overall time to hip fracture was defined as

Date of fracture – Date of inclusion in the study (Index date)

An important feature of the data resulting from this design is the fact that patients have varying follow-up time during which hip fracture may occur. The timing of fractures after initiating therapy was therefore of interest, as well as the frequency of fractures. The observations for those patients that do not experience outcome by the end of follow-up are censored.

This censorship was the key reason why Cox regression is the most appropriate model structure in this evaluation (see section 5.2.1). The analysis of data that describes the time from an origin to the occurrence of a particular event is called survival analysis, or Cox proportional hazards regression analysis (Altman 1997). The hazard ratio provided by Cox regression is analogous to the relative risk, however it incorporates the aspect of time. A strong assumption underlying this analysis is that of proportional hazards, which means that the impact of the different variables on outcome is constant over time. No particular distribution is assumed for the survival times in Cox regression. (Altman 1997)

A number of factors that practically and intuitively were considered to be plausible predictors of outcome were explored on their own and in combination, and those factors that came out significant were included in the final model. Powerful explanatory variables may be distributed unevenly across the patient cohorts because of the non-random treatment assignment. The impact of the medication is therefore at risk of being outweighed by other risk factors for fracture. It was essential to analyse the association of medication effect on fracture risk in the presence of such potential covariates on fracture risk. The patient records were screened for the presence of the following variables, which *a priori* were considered to be plausible risk factors for fracture: smoking status, previous hip fracture,

previous wrist fracture, previous vertebral fracture, previous other fracture, history of osteoporosis and age. Patient age was divided into two strata: <75 and ≥75.

Those variables that were deemed to have prognostic value (at conventional significance level 0.05) were used in a model selection analysis to assess their importance in the presence of other variables. The combination of those factors that were significant in the univariate analysis was explored through creating interaction terms and that were fitted into the Cox models. Drug exposure (index drug) was finally added to the model to assess the impact of osteoporosis drug on the risk of hip fracture in the presence of observed covariates. Patients in the bisphosphonate treatments (etidronate and alendronate) was analysed and compared separately from the patients receiving the HRTs, raloxifene and tibolone.

The cumulative hazard of hip fracture was plotted using the Kaplan-Meier method. The goodness-of-fit of the Cox regression model was assessed through the plotting and examination of the residuals. Examining that the curves in the Kaplan-Meier plots did not cross tested whether the assumption of proportional hazards was violated.

Modelling attrition rate

A secondary outcome of the analysis was the patients' duration of therapy in clinical practice. Of particular interest was whether or not there was a difference in duration of treatment between the cohorts of women initiating therapy on the various intervention drugs.

Attrition rate is the rate at which the GPs cease to issue prescriptions of the study drug. Overall time to stop was defined as

Stop date – Date of inclusion in the study (Index date)

The proportion of patients remaining at each quarter after index date was compared using chi-square test. The estimated time to stop was compared between the

treatments using Cox regression. Kaplan Meyer plot was used to present graphically the rate at which GPs cease to issue prescriptions of the study drugs.

Analysing resource consumption data

Unit costs for the resource items were published by Personal Social Services Research Unit, University of Kent (see table 5.3). (PSSRU 2000)

Table 5.3. Unit costs of resource consumption in secondary healthcare (source PSSRU 2000).

Item	Unit cost
GP consultation	£ 18
Practice nurse	£ 9
Hospital referral (A&E)	£ 282
Rheumatologist	£ 78
GI investigation	£ 73
Physiotherapist	£ 21
Radiologist	£ 23

The intention was to cost GP visits and analyse average cost increase prior to and after exposure to index drug between appropriate pairs of study drugs. The assumption of normality underlying generalised linear modelling techniques of continuous data, is frequently violated in the case of economic evaluations. Cost data are frequently positively skewed, many patients often incur no cost and a small proportion of patients may account for a large proportion of the resource use and hence costs. Comparison of distributions on a log scale data provides a comparison of geometric means of the sample. Geometric means are however always less than the arithmetic mean (average) and therefore the comparison of geometric mean through log-transformation of the data may be inappropriate. (Thompson and Barber 2000) Using non-parametric Mann-Whitney U test is also inappropriate as this is based on the comparison of medians rather than means and the results are difficult to interpret in the context of pound value. (Thompson and Barber 2000)

The non-parametric bootstrap procedure (Efron and Tibshirani 1986) has been proposed as an alternative to investigating differences in costs between groups as it does not make parametric assumptions about normality or symmetry on the distribution of cost the difference. (Briggs and Gray 1999) The method involves re-

sampling with replacement from the study sample and computing of cost differences in each of the multiple samples. In this analysis the non-parametric bootstrap procedure with 10.000 repetitions was used to compare the difference in arithmetic mean costs between treatment and control groups. To express uncertainty in the cost difference estimate the 25th and 75th percentiles of the ranked bootstrapped cost differences represent the boundaries for the 95% confidence interval (i.e. the 251st and 9.750th ranked differences). The difference between costs in the groups is significant if the 95% confidence interval does not include 1 at conventional level of significance ($p < 0.05$).

The drug class non-steroid anti-inflammatory drugs (NSAIDs) are known to induce gastrointestinal side effects similar to those experienced with bisphosphonates. The change in average cost of gastrointestinal (GI) drug use associated with the use of bisphosphonates was assessed by comparing the costs incurred before initiation of study medication with those incurred after first use of bisphosphonate. This analysis assessed whether the observed association between bisphosphonates and GI drugs could be explained by an inter-relationship between GI drug use (outcome) and NSAID. The analysis of GI drug costs was therefore carried out incorporating both variables (bisphosphonate and NSAIDs). The use of NSAID was defined as a binary variable.

Referrals and admissions were tabulated and compared between the cohorts. In those cases where the difference in absolute number of referrals differ significantly between the cohorts between year 0 and 1 the referral rates were costed and analysed by using bootstrap t-test.

Cost-effectiveness analysis

The aim of cost effectiveness analysis is to compare the costs and effects of one treatment compared to some relevant alternative. Known costs and effects of treatment versus control interventions may therefore be summarised in a cost-effectiveness ratio. The incremental cost-effectiveness ratio (ICER) provides a summary of the cost-effectiveness of one intervention relative to a comparison therapy or no therapy by dividing the difference in the mean costs of the two

alternatives by the difference in mean effects of the alternatives. The ICER of one treatment compared to an alternative is therefore given by:

$$ICER = \frac{C_t - C_c}{E_c - E_t}$$

Where C_T is the aggregate cost of treatment, C_C is the aggregate cost of control, E_T is the effectiveness of the therapy, and E_C is the effectiveness of the control therapy. Incremental cost effectiveness ratios are relevant in situations where one intervention is more costly and produce greater health benefits than the intervention being compared. In such situations the ICER facilitates a judgement of about whether the additional costs of the more expensive therapy is justified by the additional effects associated with that therapy. In situations where one intervention is more costly and produces less health benefits than the other, one intervention is said to dominate the other, and the recommendation from the economic analysis would be to adopt the less costly and more effective therapy.

The analysis of this study planned to aggregate the costs incurred in each intervention group and estimate a mean total cost per patient. The main analysis would take the perspective of the primary care sector and therefore include the following cost items: GP visits, acquisition costs and cost of concomitant medication. Where relevant, an ICER would be calculated for those study drugs that were more costly and produced greater effects than the no-intervention alternative in the age-matched control groups.

The non-parametric bootstrap method would be used to calculate 95% confidence intervals around the estimates of cost differences. Cox regression analysis was used to analyse adjusted hip fracture risk difference and 95% confidence intervals for hip fracture. Unit of analysis for effectiveness was patients with hip fracture. The cost effectiveness would therefore be expressed as “cost per patient avoiding hip fracture”. Number of patients avoiding hip fracture would be modelled from the adjusted fracture risk reduction. Both methods of confidence interval estimation are restricted to handling uncertainty related to the sampling variation.

5.4.2 Design of second study: Raloxifene and bisphosphonates in Mediplus®

This second study was designed with the purpose of replicating those aspects of the Mediplus® database that were judged feasible through the evaluation of data from the pilot study outlined in chapter 5.4.1. As will become apparent in section 5.5.1, the evaluation of drug use patterns and costs associated with medication use appeared feasible in the database. This second study evaluated raloxifene and the bisphosphonates (etidronate and alendronate) in UK clinical practice.

Data extraction

An analyst at IMS Health (UK) conducted the data extraction from the Mediplus® database onto an analytical file according to a design protocol. The statistical variables for the second study are provided in appendix 5.6.

Patients

Patient inclusion was based on the following criteria:

- Patients received a first prescription (index prescription) of any of the exposure drugs (index drug) during the period of 1st August 1998 to 31st January 2000.
- Practice was active in the database in December 1999.
- Patients were active in the database 6 months prior to the index prescription.

Patients were excluded based on the following criterion:

- Patients who had a prescription of two different study drugs on the index date.

The only demographic characteristics collected were age and the history of osteoporosis in the patient records.

Table 5.4 lists the interventions included in the study. GPs record drug prescriptions in the Mediplus® database either by product trade name, generic name or by chemical substance. Selections from the database were therefore based on both product name and substance name. Drug dose and form of administration was not considered.

Table 5.4. Study drugs considered

Product name	Substance	Basis for selection
Evista®	Raloxifene	Substance + Product
Fosamax®	Alendronic acid	Substance + Product
Didronel (PMO)®	Etidronate/calcium carbonate	Product (multi-substance)

Two control patients were identified for each study patient and matched by age. Follow-up was 6 months.

Outcomes

Concomitant drug use

A record was made of the use of the following drug therapies 6 months before and 6 months after the index date (i.e. date of inclusion in the study):

- GI Drugs (H2 blockers and proton pump inhibitors)
- Analgesics
- NSAIDs

Treatment pattern

The day at which the first prescription was issued was assigned the 'Index date'. This date indicates inclusion in the study. "Stop date" was defined as the date at which no further study drug issued in the 6 months following the end date of the duration of the last prescription of drug. A sensitivity analysis was undertaken to assess the impact of changing the definition of "stop date" to a drug-free time period of 3 months.

Statistical analysis

Modelling attrition rate

A secondary outcome of the analysis was the duration of patients' therapy in clinical practice. Of particular interest was the existence of a difference in duration of treatment between the cohorts of women initiating therapy on the various intervention drugs.

Overall time to stop was defined as:

Stop date – Date of inclusion in the study (Index date)

The proportion of patients remaining at each quarter after index date was compared using the chi-square test. Estimated time to stop between the treatments was compared using Cox regression and Kaplan Meyer plotting was used to give a graphic presentation of the rate at which GPs cease to issue prescriptions of the study drugs.

Cost of gastrointestinal drugs

The change in average cost of gastrointestinal (GI) drug use associated with the use of the bisphosphonates was assessed by comparing the costs incurred before initiation of study medication with those incurred after first use of bisphosphonate, using the non-parametric bootstrap procedure with 10.000 replications (see section 5.4.1). The drug class Non-Steroidal Anti-Inflammatory drugs (NSAID) is known to induce gastrointestinal side effects similar to those experienced with bisphosphonates, so the analysis assessed whether the observed association between bisphosphonates and GI drugs could be explained by an inter-relationship between GI drug use (outcome) and NSAID. The analysis of GI drug costs incorporated both variables (bisphosphonate and NSAIDs), with the use of NSAID defined as a binary variable (i.e. NSAID use or no NSAID use).

5.4.3 Summary of design differences between the two studies

The studies differed on several parameters. *Table 5.5* summarises the main design differences of the studies.

Table 5.5 Study aspects that differed between the first and the second study designed to evaluate the feasibility of the Mediplus® database for economic evaluation.

Study aspect	First study	Second study
Interventions	Alendronate Etidronate Tibolone HRT(P) HRT(N)	Raloxifene Alenronate Etidronate
Study period	January 1 st 1996 to December 31 st 1998	August 1 st 1998 to January 31 st 2000.
Follow-up	12 months to 3 years	6 months
Costs	GP visits Referrals to secondary care Accident and emergency admissions Drug acquisition costs Concomitant medication costs	Drug acquisition costs Concomitant medication costs (GI drugs only)
Clinical outcomes	Hip fracture Other clinical events (e.g. other fractures and cancer) Attrition rate	Attrition rate

5.5 Results

Section 5.3.1 reports the results of the **first** study of HRTs, bisphosphonates and raloxifene in clinical practice, exploring the feasibility of the database to provide variables for analysis of the cost-effectiveness of drugs in clinical practice. Section 5.3.2 reports the results of the **second** study, reiterating the feasible aspects of the first study, but including bisphosphonates and raloxifene only. The purpose of the second study was to reiterate the first, focusing on those variables of the first study that provided plausible results and applying those to a subset of patients using raloxifene and the bisphosphonates. During analysis it became apparent that the results of the overlapping variables differed substantially. Section 5.3.3 highlights these differences in the findings of the two studies.

5.5.1 Results of the First study

Patients

Overall, 8463 patients were included in the first study (*Table 5.6*). A total of 2838 patients were observed over three years, 2877 over two years and 2748 over one year. Two thirds of these patients were age-matched control patients. The number of patients exposed to each therapy in each year was similar; only etidronate differed, with a doubling of patient numbers from 1996 to 1998.

The indications for the prescription of drugs included prevention or treatment of osteoporosis. (BNF 2000) The prescription of study drug was issued under a problem heading (i.e. diagnosis) of osteoporosis for an average of 18.5% of patients included in the study (*Table 5.7*). Alendronate and etidronate users had the highest proportion of prescriptions issued for diagnosis of osteoporosis with 57.3% and 46.3%, respectively.

Table 5.6. Number of patients included in the study in study drug cohort by year.

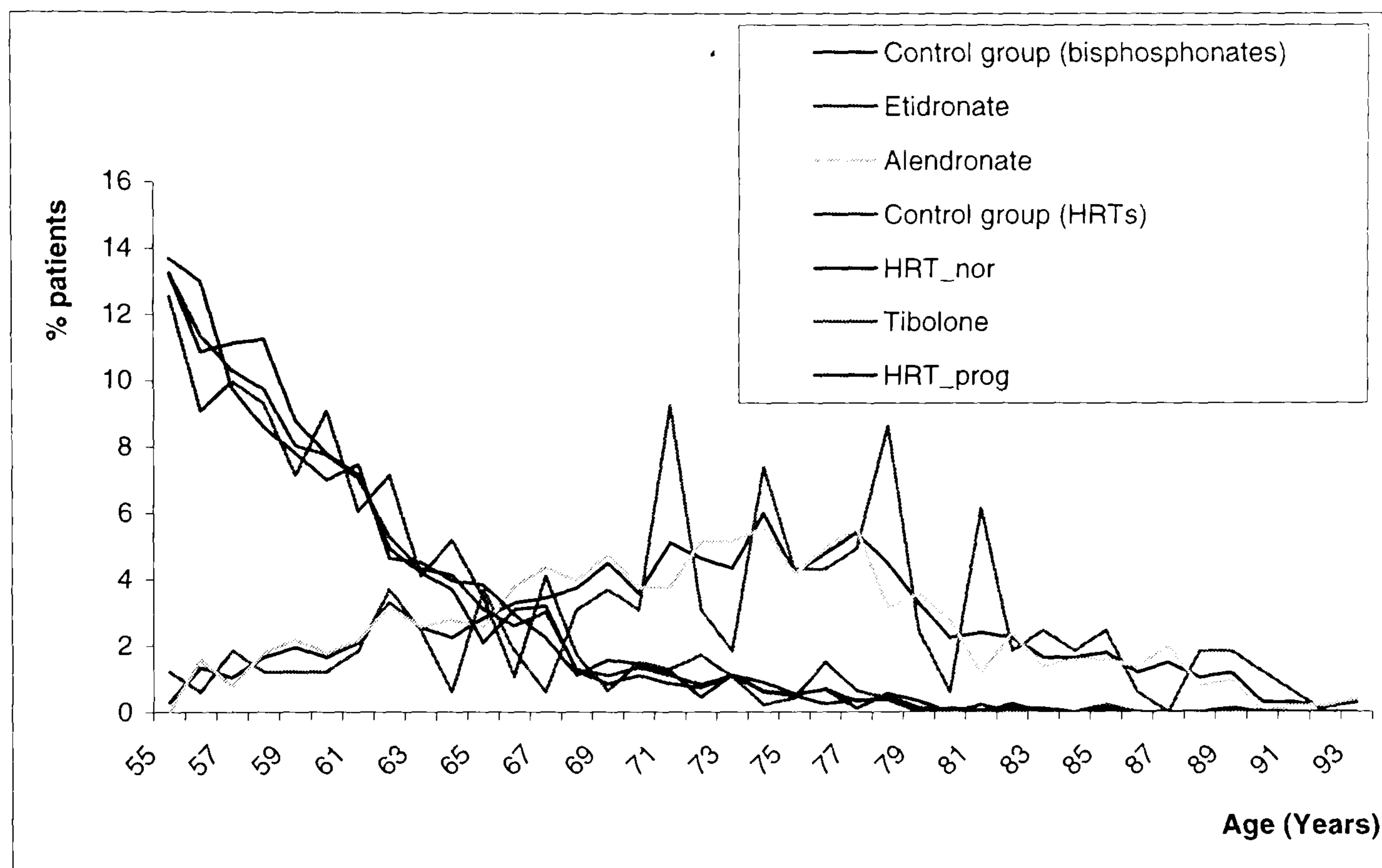
Study drug	1996 (3 years follow-up)	1997 (2 years follow-up)	1998 (1 year follow-up)	Total number of patients	Total years of follow- up
HRT(P) ¹	249	286	274	809	1593
HRT(N) ²	358	290	236	884	1890
Tibolone	151	167	144	462	931
Etidronate	34	60	68	162	290
Alendronate	154	156	194	504	968
HRT Control Group	1516	1486	1308	4310	8828
Bisphosphonate Control Group	376	432	524	1332	2516
Total	2838	2877	2748	8463	17016
¹ Conjugated oestrogen and medroxyprogesterone (Premique™)					
² Oestrogen and norethisterone (Kliofem™)					

Table 5.7. Percent of the prescription of study drug that was linked to a diagnosis of osteoporosis by study drug.

Study drug	Diagnosis of Osteoporosis	Other diagnosis or diagnosis missing	Total number of patients
Etidronate	46.3 (n=75)	53.7 (n=87)	162
Alendronate	57.3 (n=289)	42.7 (n=215)	504
HRT(N)	7.9 (n=70)	92.1 (n=814)	884
Tibolone	7.8 (n=36)	92.2 (n=426)	462
HRT(P)	6.2 (n=50)	93.8 (n=759)	809
Total	18.5 (n=520)	81.5 (n=2301)	2821

Patient characteristics of age, height body mass index (BMI) and smoking status were recorded. Patient height was not routinely recorded in the database – indeed, was missing for between 70% and 90 % of the patients. Extreme values were recorded, for example, one patient had a height of 1.20 meters recorded and one patient had a height of 1.90 meters recorded. The mean heights were between 1.58 metres and 1.62 metres (*Table 5.8*). Age was negatively skewed for the patients using hormone replacement therapies (HRTs) and their corresponding control group (*Figure 5.2*). Patient age was symmetrically distributed around the mean in the bisphosphonate cohorts and their corresponding control group. The age distribution was similar between the bisphosphonates as a group and between the different HRTs.

Figure 5.2 Age distribution for the patients included in the study by study drug cohort.



Patients that initiated therapy with the bisphosphonates (etidronate and alendronate) were more likely to have clinical conditions recorded in their patient records than those that initiated on HRT therapy (*Table 5.9*). For example, between 34.5 % and 38.3 % of women on bisphosphonates suffered from arthritis, while between 16.9 % and 19.5 % of women on the HRT drugs suffered from the same condition. The presence of gastrointestinal (GI) disease was an exception, as it had similar frequency in all study drug cohorts, with between 9.3% and 13.0% for the bisphosphonates and between 5.7% and 14.3% in the HRT cohorts.

Osteoporosis was recorded as a clinical condition in 59.9% of the patients receiving etidronate and in 73.4% of patients receiving alendronate (*table 5.10*). In comparison, osteoporosis was recorded in only 1.8% of the patients in the bisphosphonate control group. Between 9.9 % and 12.3% of the HRT patients had a recording of osteoporosis in their patient records.

Table 5.11 reports the number of patients, by drug cohort, that had a fracture history prior to inclusion. The patient records were searched for the 5-year period

prior to the prescription of study drug for the presence of various fractures. Overall, 63 women had a previous fracture recorded in their patient records, representing 0.7% of the total number of women included in the study. Of the etidronate patients, 2.5% and of the alendronate patients 4.4 % had hip fractures recorded in their clinical records prior to being prescribed a bisphosphonate for the first time. Only about 0.2% - 0.7% of the women initiating HRT therapy had previously experienced a hip fracture.

Table 5.8. Patient characteristics by drug cohort

Patient characteristic	Bisphosphonate group				HRT group			
	Etidronate	Alendronate	Bisphosphonate Control Group	HRT_Nor	HRT_prog	Tibolone	HRT Control Group	
Age mean (sd)	73.4 (8.23)	72.1 (8.0)	72.4 (8.0)	60.6 (5.3)	60.3 (5.1)	61.0 (5.5)	60.5 (5.3)	
Height mean (sd)	1.58 (0.1)	1.58 (0.1)	1.59 (0.1)	1.62 (0.1)	1.61 (0.1)	1.61 (0.1)	1.61 (0.1)	
BMI mean (sd)	25.7 (4.4)	24.5 (4.2)	26.3 (5.3)	25.7 (4.4)	25.8 (4.8)	26.0 (4.6)	27.3 (5.7)	
Smokers n (%)	26 (16.1)	58 (11.5)	145 (10.9)	150 (17.0)	146 (18.1)	90 (19.5)	549 (12.7)	

Table 5.9. Number and frequency (%) of patients with clinical conditions entered in the records in the 9 months prior to index date.

Clinical condition	Bisphosphonate group				HRT Group			
	Etidronate	Alendronate	Bisphosphonate Control Group	HRT_nor	HRT_prog	Tibolone	HRT Control Group	
Asthma	20 (12.3)	40 (7.9)	70 (5.3)	59 (6.7)	45 (5.6)	35 (7.6)	207 (4.8)	
Pulmonary disease	20 (12.3)	58 (11.5)	115 (8.6)	52 (5.9)	55 (6.8)	45 (9.7)	234 (5.4)	
Respiratory system disease	9 (5.6)	36 (7.1)	38 (2.9)	53 (6)	53 (6.55)	26 (5.6)	128 (3.0)	
Rheumatism	35 (21.6)	86 (17.1)	133 (10.0)	89 (10.1)	74 (9.2)	50 (10.8)	310 (7.2)	
Arthritis	62 (38.3)	174 (34.5)	265 (19.9)	149 (16.9)	152 (18.8)	90 (19.5)	548 (12.7)	
GI disease	21 (13.0)	47 (9.3)	99 (7.4)	50 (5.7)	63 (7.8)	32 (6.9)	195 (4.5)	

Table 5.10. Number and frequency (%) of patients with clinical conditions entered in the records during the 5 years prior to index date.

Clinical condition	Bisphosphonate group			HRT Group			
	Etidronate	Alendronate	Bisphosphonate Control Group	HRT_nor	HRT_prog	Tibolone	HRT Control Group
Osteoporosis	97 (59.9)	370 (73.4)	24 (1.8)	102 (11.5)	80 (9.9)	57 (12.3)	48 (1.1)
Hysterectomy	6 (3.7)	21 (4.2)	24 (1.8)	4 (0.5)	10 (1.2)	9 (2.0)	16 (2.7)
Breast cancer	1 (0.6)	11 (2.2)	32 (2.4)	2 (0.2)	0	4 (0.9)	90 (0.2)
Ovarian cancer	0	1 (0.2)	5 (0.4)	2 (0.2)	2 (0.3)	1 (0.2)	16 (0.37)
CHD	29 (17.9)	78 (15.5)	163 (12.2)	58 (6.6)	50 (6.2)	26 (5.6)	232 (5.4)
Stroke	4 (2.5)	5 (1.0)	17 (1.3)	3 (0.3)	4 (0.5)	2 (0.4)	24 (0.6)
MI	11 (6.8)	18 (3.6)	34 (2.6)	9 (1.0)	12 (1.5)	3 (0.7)	41 (1.0)
Cerebrovascular disease	9 (5.6)	27 (5.4)	68 (5.1)	13 (1.5)	10 (1.2)	4 (0.9)	57 (1.3)
Hypertension	36 (22.2)	136 (27.0)	339 (30.0)	163 (18.4)	149 (18.4)	88 (19.1)	861 (20.0)

Table 5.11. Number and frequency (%) of patients with fractures entered in the records during the 5 year prior to index date.

Fracture	Bisphosphonate Group			HRT Group			
	Etidronate	Alendronate	Bisphosphonate Control Group	HRT_nor	HRT_prog	Tibolone	HRT Control Group
Hip	4 (2.5)	22 (4.4)	15 (1.1)	6 (0.7)	2 (0.3)	1 (0.2)	13 (0.3)
Vertebral	4 (2.5)	27 (5.4)	1 (0.1)	7 (0.8)	1 (0.2)	1 (0.2)	7 (0.2)
Wrist	5 (3.1)	38 (7.5)	31 (2.3)	32 (3.6)	9 (1.1)	14 (3.3)	88 (2.0)
Other	14 (8.6)	57 (11.3)	64 (4.8)	41 (4.6)	35 (4.3)	14 (3.0)	148 (3.4)

Table 5.12. Number and frequency (%) of patients with fractures entered in the records after index date.

Fracture	Bisphosphonate Group				HRT Group			
	Etidronate	Alendronate	Bisphosphonate Control Group	HRT_nor	HRT_prog	Tibolone	HRT Control Group	
Hip	2 (1.23)	18 (3.57)	14 (1.05)	1 (0.11)	3 (0.37)	1 (0.22)	10 (0.23)	
Vertebral	1 (0.62)	7 (1.39)	1 (0.08)	2 (0.23)	2 (0.25)	1 (0.22)	3 (0.07)	
Wrist	4 (2.47)	7 (1.39)	14 (1.05)	10 (1.13)	9 (1.11)	4 (0.87)	45 (1.04)	
Other	7 (4.32)	22 (4.37)	1 (1.73)	20 (2.26)	8 (0.99)	8 (1.73)	85 (1.97)	

Table 5.14. Frequency (%) of patients with potential prognostic variables for hip fracture.

Variable	Bisphosphonate Group				HRT Group			
	Etidronate	Alendronate	Bisphosphonate Control Group	HRT_nor	HRT_prog	Tibolone	HRT Control group	
Smoking	16.1	11.5	10.9	17.0	18.1	19.5	12.7	
Previous hip fracture	2.5	4.4	1.1	0.7	0.3	0.2	0.3	
Previous wrist fracture	3.1	7.5	2.3	3.6	1.1	3.0	2.0	
Previous other fracture	8.6	11.3	4.8	4.6	4.3	3.0	3.4	
History of Osteoporosis	59.9	73.4	1.8	11.5	9.9	12.3	1.1	
Age >75	42.0	34.1	36.0	2.0	1.5	3.0	2.0	

Risk of Fractures

The variables describing the presence of fractures were defined as the number of fractures recorded on each patient. Six patients had two hip fractures recorded, three patients had four hip fractures recorded and one patient had six hip fractures recorded in her record. Similarly, one patient had four wrist fractures and one had four vertebral fractures recorded. This is likely to have been repeated recording of the same injury, so the fracture data was analysed according to the number of patients experiencing a fracture rather than by number of fractures which occurred (*Table 5.12*).

A total of 49 patients experienced a fracture after the index date in the whole sample of 8463 patients, representing 0.6% of the population, or about 6 per 1000. More patients in the bisphosphonate treatment groups experienced a hip fracture after index date compared to those receiving HRT therapies (*Table 5.13*).

Table 5.13. Frequency of hip fracture in the index drug cohorts after index date.

Drug cohort		Patients with fracture	Frequency (%)
Bisphosphonate group	Etidronate	2	1.2
	Alendronate	18	3.6
	Bisphosphonate Control Group	14	1.1
HRT group	HRT(N)	1	0.1
	HRT(P)	3	0.4
	Tibolone	1	0.2
	HRT Control Group	10	0.2
Total		49	0.6

A number of clinical conditions were, *a priori*, considered to be likely predictors of the risk of fracture, and the predictive power of the following variables was explored: smoking status, previous hip fracture, previous wrist fracture, previous vertebral fracture, previous other fracture, history of osteoporosis and age >75. The prevalence of these attributes differed by study drug cohort (*Table 5.14*). When analysed individually, the predictive value of previous hip fracture, previous wrist fracture, previous other fracture, history of osteoporosis and patient age (> 75) were significant, at $p < 0.05$ (*Table 5.15*).

The significant variables were tested in the presence of each other. Recording of wrist and other fracture and of osteoporosis in the patient records were insignificant in the presence of age>75 and previous hip fracture (*Table 5.16*). First order interactions between the significant variables were not considered clinically plausible, and so were not explored. The final model contained previous hip fracture and age>75 (see *Table 5.17*).

Table 5.15. Cox proportional hazards analyses of time to hip fracture among patients with clinical conditions potentially having predictive power.

Variable	Property	Number of hip fractures by variable	p-value (SE)	Hazard ratio (95% CI)
Smoking Status	Non smoker	41 / 7264 (0.56%)	0.650 (0.39)	1.19 (0.56 to 2.54)
	Smoker	8 / 1157 (0.69%)		
Previous Hip Fracture	No fracture	26/8381 (0.31%)	0.0001 (0.29)	171.84 (97.66 to 302.35)
	Fracture	23/63 (36.51%)		
Previous Wrist Fracture	No fracture	45 / 8208 (0.55%)	0.020 (0.52)	3.365 (1.21 to 9.36)
	Fracture	4 / 213 (1.84%)		
Previous Vertebral Fracture	No fracture	49 / 8373 (0.58%)	1 (774.84)	0.00
	Fracture	0 / 48 (0%)		
Previous Other Fracture	No fracture	43 / 8054 (0.53%)	0.0099 (0.44)	3.077 (1.31 to 7.23)
	Fracture	6 / 367 (1.61%)		
History of Osteoporosis	Diagnosis	43 / 8054 (0.53%)	0.0313 (0.37)	2.213 (1.07 to 4.56)
	No diagnosis	6 / 367 (1.61%)		
Patient age (> 75)	<70	25/7474 (0.33)	0.0001 (0.29)	7.816 (4.46 – 13.69)
	≥ 75	24/947 (2.47%)		

Table 5.16. Cox proportional hazards analyses of time to hip fracture among patients with clinical conditions having predictive power individually.

Variable	p-value	SE	Hazard ratio (95% CI)
Previous wrist fracture	0.694	0.59	0.8 (0.3 to 2.5)
Previous hip fracture	<0.0001	0.32	104.7 (55.9 to 197.4)
History of osteoporosis	0.950	0.39	1.0 (0.5 to 2.1)
Previous other fracture	0.142	0.50	2.1 (0.8 to 5.4)
Age >75	<0.0001	0.31	3.7 (2.0 to 6.8)

Table 5.17. Final Cox proportional hazards model of time to hip fracture*

Variable	p-value	SE	Hazard ratio (95%CI)
Previous Hip Fracture	<0.0001	0.31	108.3 (58.4 to 200.8)
Age >75	<0.0001	0.32	3.5 (1.9 to 6.5)
* -2 Log L = 671.515			

Finally, the index drugs were included in the model to investigate drug impact in the presence of the explanatory model of prognostic variables. None of the drugs produced a significant reduction in the relative risk of hip fracture compared with the control groups. Indeed, the analysis indicated that alendronate compared to the bisphosphonate control group actually *increased* the likelihood of experiencing a fracture compared with the age matched control groups (Table 5.18).

Further analysis to compare fracture rates between the HRTs was not considered meaningful, as patient numbers and recorded fractures were low in these groups.

Table 5.18. Adjusted Cox proportional hazards analyses of time to hip fracture among patients using bisphosphonates and HRTs.

Comparison	Hazard ratio (95% CI)	P-value	-2 Log L
Etidronate vs Bisphosphonate control group	0.77 (0.17 to 3.49)	0.74	187.45
Alendronate vs Bisphosphonate control group	1.95 (0.94 to 4.07)	0.07	371.37
Etidronate vs Alendronate	0.42 (0.10 to 1.82)	0.25	199.69
HRT(N) vs. HRT control group	0.23 (0.03 to 2.03)	0.18	104.60
Tibolone vs. HRT control group	0.91 (0.11 to 7.59)	0.93	101.78
HRT(P) vs. HRT control group	1.07 (0.27 to 4.28)	0.93	137.58

Treatment pattern

The attrition rate for those patients who initiated therapy with the bisphosphonates was analysed separately from those initiating therapy with the HRTs, raloxifene and tibolone. The proportion of patients remaining on etidronate was significantly lower than that of alendronate takers at every point in time for the main analysis using the 3-month definition of stop date (*Table 5.19, Figure 5.4*). Using the 6-month definition of stop date provided a different result (*Table 5.20 and Figure 5.3*).

*Table 5.19. Proportion of patients retaining treatment with the bisphosphonates at quarterly intervals of the first year**

	Etidronate	Alendronate	p-value
0 months	100	100	NA
3 months	69.8	94.1	<0.05
6 months	61.1	92.9	<0.05
9 months	54.9	91.5	<0.05
12 months	51.9	89.9	<0.05
*The 3-month of drug free period was used as definition of stop date.			

*Table 5.20. Proportion of patients retaining treatment with the bisphosphonates at quarterly intervals of the first year**

	Etidronate	Alendronate	p-value
0 months	100	100	NA
3 months	90.7	97.0	<0.05
6 months	88.3	96.4	<0.05
9 months	85.2	96.23	<0.05
12 months	84.0	95.4	<0.05
The 6-month of drug free period was used as definition of stop date.			

Figure 5.3 Attrition rates for the study drugs at one year using 6 month definition for stop date.

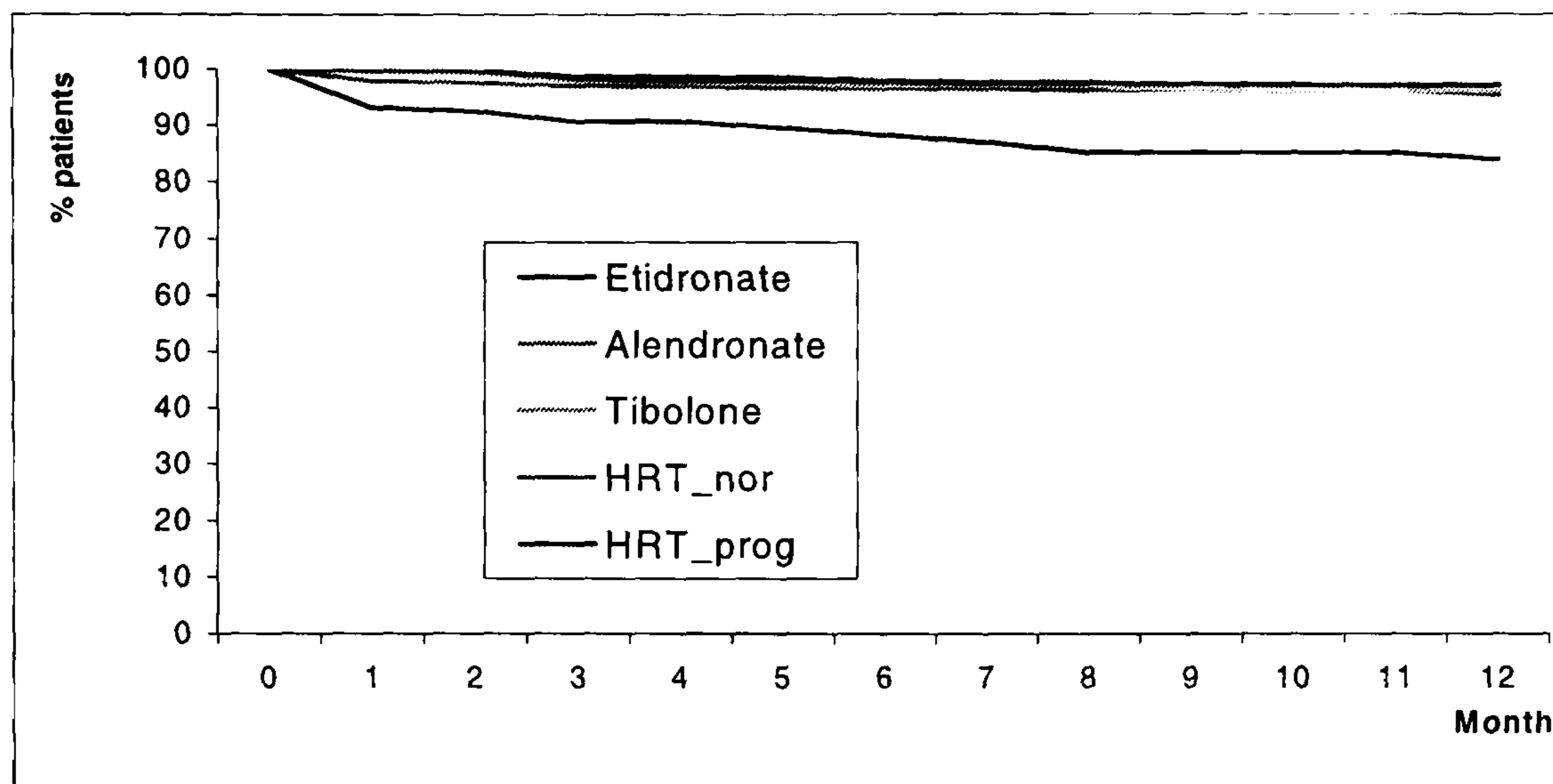
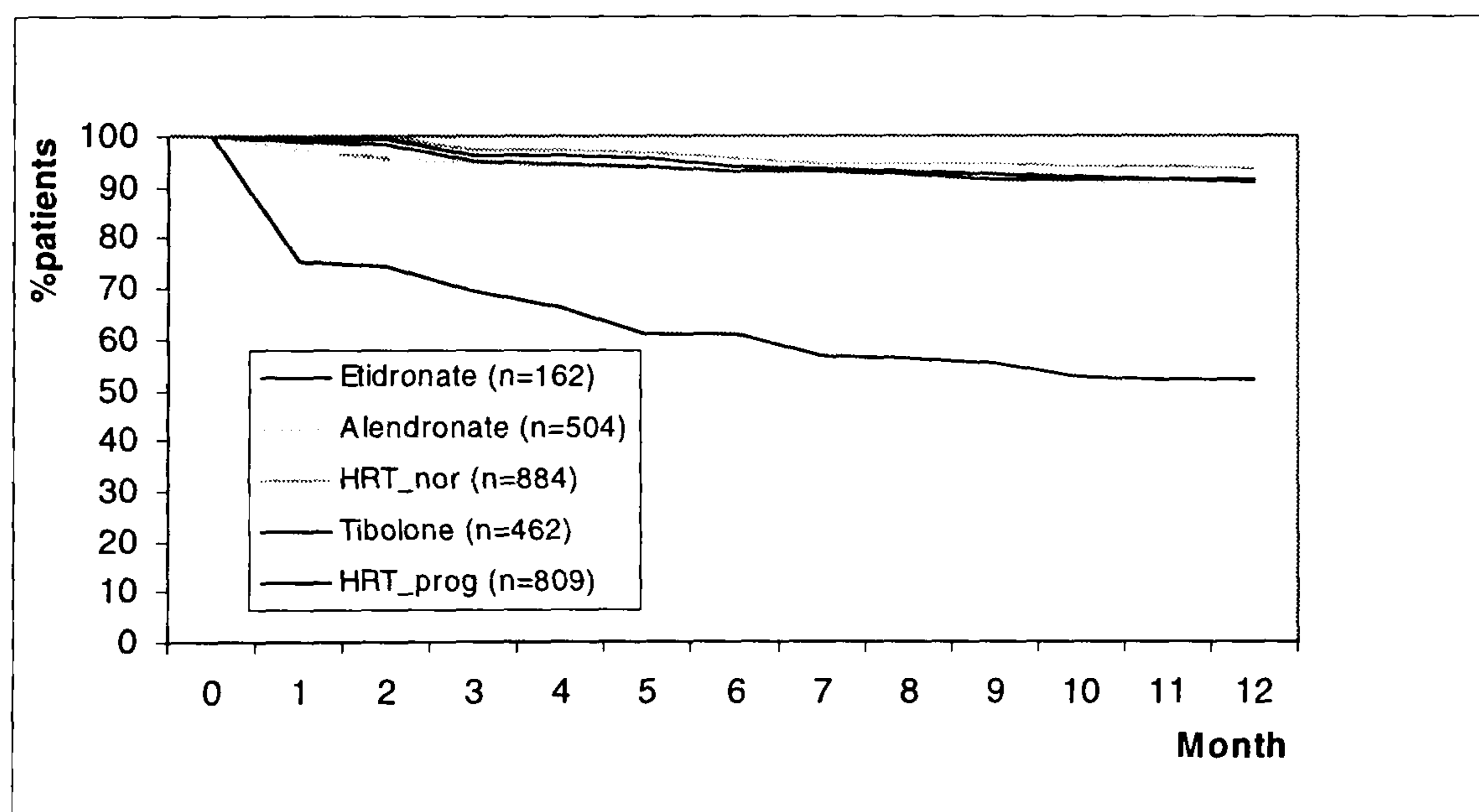


Figure 5.4 Attrition rates for the study drugs at one year using 3 month definition for stop date.



The hazard of stopping therapy with etidronate was 5.3 (3.8 to 7.4) times the hazard of stopping therapy with alendronate when using the 3-month definition of stop date (*Table 5.21*). This difference was highly significant ($p < 0.0001$). When using the 6-month therapy stop definition the hazard estimate differed substantially, as the hazard of stopping therapy with etidronate was 2.1 (1.7 to 2.6) times that of stopping therapy with alendronate – still a highly significant difference ($p < 0.0001$).

Table 5.21. Cox proportional hazards analyses of average time to therapy stop in the first year among patients using the bisphosphonates.

	Etidronate	Alendronate
Total number of patients	162	504
Stopped overall n (%)	83 (51.2)	68 (13.5)
Average time to stop, months (95% CI)	3.7 (1.9 - 5.6)	8.00 (6.9 - 9.1)
The 3-month of drug free period was used as definition of stop date.		

There was no statistically significant difference between any of the remaining proportions of women initiating therapy with the HRTs at any point in time (*Table 5.22*). The pattern was confirmed when using the 6-month definition of stop date, so Cox proportional hazards analyses for the HRTs was not pursued.

*Table 5.22. Proportion of patients retaining treatment with HRTs at quarterly intervals of the first year**

Interval	HRT(N)	HRT(P)	Raloxifene	Tibolone
0 months	100	100	100	100
3 months	97.4	96.4	100	95.5
6 months	95.5	94.1	85.7	93.1
9 months	94.5	92.3	85.7	91.6
12 months	93.6	91.6	85.7	90.9
The 3-month of drug free period was used as definition of stop date.				

Resource consumption

Table 5.23 illustrates the average number of GP consultations as reported by the formula inherent in the Mediplus® database to calculate GP visits from number of dates in the patient record. The average number of GP visits was between 9 and 26 per year in the patient cohorts, according to the calculation by this formula. Standard deviations were large and the reported range was exceptionally high, with extreme values of estimated 132 GP visits per year. In comparison, the number of dates entered in the patient records was remarkably similar to the number of GP visits as estimated by the Mediplus® system (*Table 5.24*). The Mediplus® formula does not seem to provide a realistic estimate of GP visits, so further analysis of cost of GP activity was not pursued.

Table 5.23. Average (std / range) GP visits by drug by year as estimated by the formula inherent in the Medipus® database.

Study period	Bisphosphonate Group			HRT Group			
	Etidronate	Alendronate	Bisphosphonate Control Group	HRT_nor	HRT_prog	Tibolone	HRT Control group
6 months prior	12.5 (0 - 43)	10.6 (0 - 47)	6.4 (0 - 56)	5.8 (0 - 42)	5.9 (0-46)	10.1 (2-22)	4.3 (0-59)
1 st year	26.2 (4-132)	25.0 (1-116)	13.6 (0-100)	14.7 (1-69)	14.8 (1-89)	17.1 (2-77)	9.2 (0-124)
2 nd year	24.8 (0-110)	23.0 (0-112)	14.5 (0-128)	13.9 (0-69)	13.5 (0-68)	15.2 (0-62)	9.6 (0-103)
3 rd year	24.3 (0-95)	23.6 (0-97)	14.3 (0-102)	14.1 (0-58)	14.1 (0-68)	16.4 (0-88)	9.8 (0-89)

Table 5.24. Average (std / range) number of dates entered in the patient records

Study period	Bisphosphonate Group			HRT Group			
	Etidronate	Alendronate	Bisphosphonate Control Group	HRT_nor	HRT_prog	Tibolone	HRT Control group
6 months prior	12.7 (0-34)	10.7 (0-47)	6.5 (0-56)	5.9 (0-42)	5.9 (0-46)	6.3 (0-35)	4.4 (0-59)
1 st year	26.6 (4-135)	25.3 (1-116)	13.8 (0-100)	14.9 (1-69)	15.1 (1-89)	17.4 (2-77)	9.4 (0-124)
2 nd year	25.1 (0-112)	23.37 (0-112)	14.8 (0-128)	14.2 (0-69)	13.8 (0-70)	15.6 (0-65)	9.8 (0-103)
3 rd year	24.6 (0-95)	23.9 (0-97)	14.6 (0-102)	14.35 (0-59)	14.3 (0-68)	16.8 (0-91)	10.1 (0-89)

Average costs of the study drugs are reported in *Table 5.25*. These were accumulated over the period of time that patients took medication, whether they stopped prematurely or whether they continued until the end of follow-up. The average cost of etidronate was about £106, and of alendronate £340. The standard deviation was very large in the etidronate group, and the maximum observed cost was, indeed, more than £300 higher than that of alendronate. The HRT drugs were, on average, less costly than the bisphosphonates. Tibolone differed from the HRT drugs in that the average cost was higher and the maximum observed cost higher.

Table 5.25. Acquisition costs for study drugs by drug cohort. Costs are in UK £ Sterling (1999).

	Bisphosphonate group		HRT group		Tibolone
	Etidronate	Alendronate	HRT(N)	HRT(P)	
Average	105.96	340.22	125.91	122.53	186.73
Standard deviation	154.24	315.07	109.62	98.45	160.30
Minimum value	8.78	25.69	8.65	6.46	13.20
Maximum value	1467.18	1156.05	441.15	678.30	856.92

The costs of gastrointestinal (GI) drugs in the year prior to index date differed between the bisphosphonate group and the corresponding control group. The average cost in the previous year for H2-antagonists was £5.12 in the control group and £11.94 in the bisphosphonate group, and similarly £10.08 and £30.10 for proton pump inhibitors (*Figure 5.5*). During the first year of bisphosphonate use, the average cost of H2 antagonists increased by £0.96 (95% CI -1.45 to 3.38) when compared to the controls (*Table 5.26*). In comparison, the average cost of proton pump inhibitors increased significantly by £10.29 during the first year (95% CI 4.47 to 16.46) when compared to controls. Similar increases were observed in subsequent years. The relative magnitude of the difference in the use of H2 antagonists and proton pump inhibitors was similar, at about 30%, though the absolute difference was greater in the case of the proton pump inhibitors, as these are more expensive drugs.

Figure 5.5. Increase in the average cost of GI drugs between previous year and first year after inclusion in the study

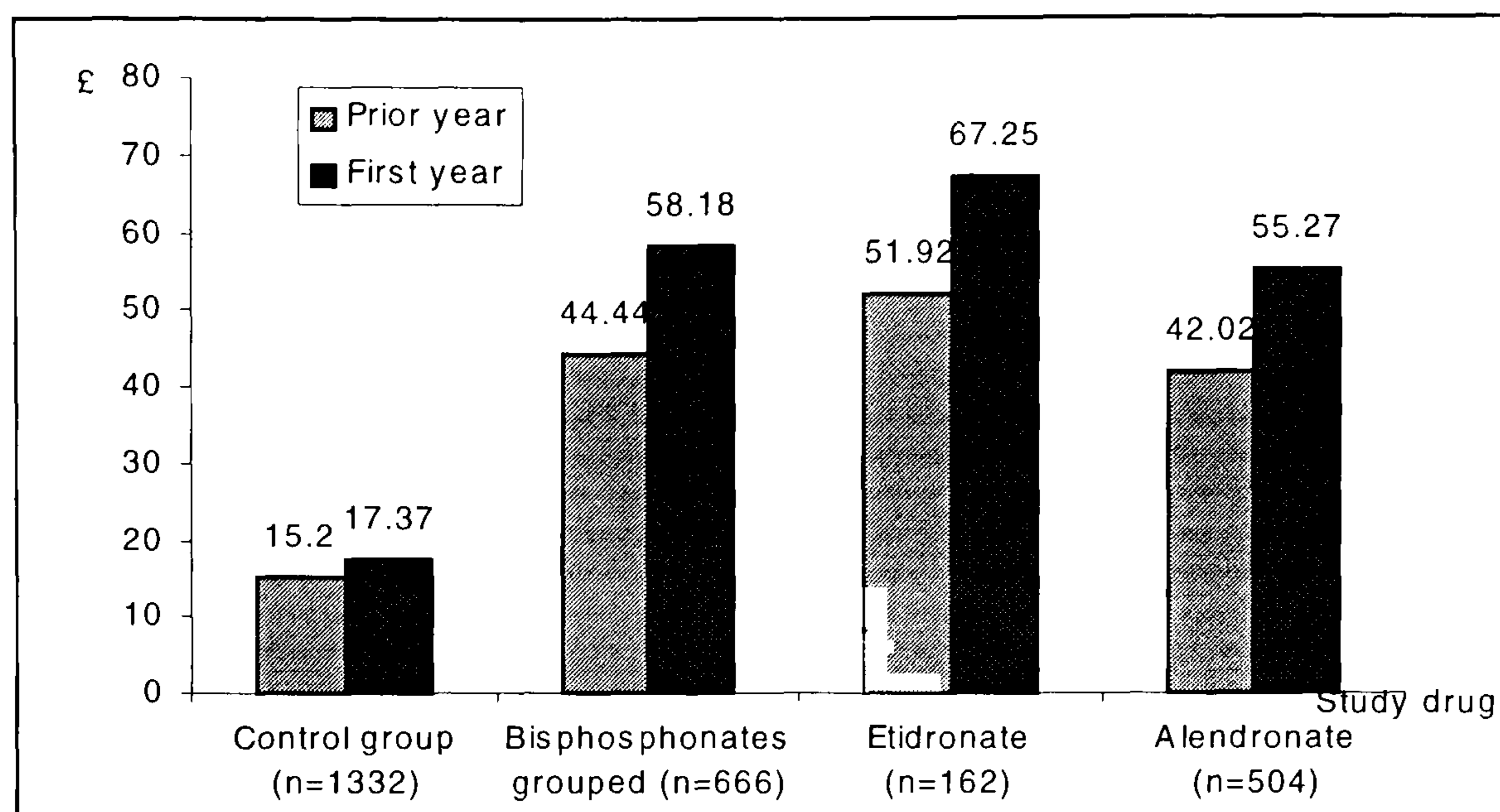


Table 5.26. Increase in GI drug costs for patients using the bisphosphonates between prior and first year, and the subsequent follow-up years. The results are stable in the presence of GI condition and NSAID use.

Drug	Year	Mean difference from control group (£)	95% CI
H2 antagonists	0-1	0.96	-1.45 to 3.38
	1-2	3.11	0.31 to 6.04
	2-3	3.30	0.72 to 6.27
Proton pump inhibitors	0-1	10.29	4.47 to 16.46
	1-2	11.48	4.01 to 19.25
	2-3	9.47	3.71 to 15.48
H2 antagonists and Proton pump inhibitors grouped	0-1	11.26	5.25 to 17.51
	1-2	14.62	6.78 to 22.93
	2-3	12.73	6.36 to 19.34

Both concomitant NSAID use and the presence of a GI condition in the patient records individually influenced the cost of GI drug use. The interaction effect between treatment and GI condition and treatment and NSAID use respectively, was small and the effect of treatment on GI costs was stable in the presence of both factors.

The absolute number of people taking GI drugs increased between the prior and the first year of treatment with the bisphosphonates (*Table 5.27*). H2 antagonists were issued to 35 (5.3%) new patients and proton pump inhibitors to 49 (7.3%) new patients after treatment with alendronate and etidronate. In the same time period,

30 patients stopped taking GI drugs after initiating bisphosphonate therapy. In subsequent years the increase was not remarkable, in spite of the fact that the average cost of GI drugs continued to increase.

Table 5.27. The number of patients on bisphosphonates that take GI drugs and patients that stop and how many of those that initiate therapy on GI drugs in the follow-up years. (%)

Interval	Patients	Patients using H2 antagonists	Patients using Proton Pump inhibitors
Year 0-1 (666 patients)	Using GI drug prior to study	78 (11.7)	82 (12.3)
	Using GI drug in first year	94 (14.1)	102 (15.3)
	Stop taking therapy	19	11
	Initiate therapy	35	49
Year 1-2* (404 patients)	Using GI drug in first year	66 (16.3)	72 (17.8)
	Using GI drug in second year	52 (12.9)	79 (19.5)
	Stop taking therapy	33	12
	Initiate therapy	19	24
Year 2-3* (188 patients)	Using GI drug in second year	28 (14.9)	37 (19.7)
	Using GI drug in third year	22 (11.7)	39 (20.7)
	Stop taking therapy	13	7
	Initiate therapy	7	9

* The baseline number of patients having accumulated follow-up times of more than a year is reduced by year. That is why the absolute number of patients using GI drug in first year is reduced when comparing to use in second year. Similar for year 2-3.

The difference in GI costs for etidronate and alendronate were also analysed separately (*Table 5.28 and Table 5.29*). Significant changes observed in GI costs before and after therapy differed in direction and magnitude.

*Table 5.28. Increase in GI drug costs for patients using **Etidronate** between prior and first year, and the subsequent follow-up years.*

Drug	Year	Mean difference to control group (£)	95% CI
H2 antagonists	0-1	-1.09	-5.70 to 3.17
	1-2	7.91	1.83 to 14.78
	2-3	2.67	-1.79 to 7.53
Proton pump inhibitors	0-1	12.51	0.05 to 24.95
	1-2	16.65	-0.49 to 32.70
	2-3	16.64	3.62 to 29.69
H2 antagonists and Proton pump inhibitors grouped	0-1	13.49	-1.08 to 26.30-
	1-2	24.45	6.64 to 42.00
	2-3	19.50	6.05 to 33.16

Table 5.29. Increase in GI drug costs for patients using Alendronate between prior and first year, and the subsequent follow-up years.

Drug	Year	Mean difference to control group (£)	95% CI
H2 antagonists	0-1	- 0.93	-3.71 to 1.79
	1-2	1.59	-1.49 to 4.82
	2-3	3.43	0.53 to 7.10
Proton pump inhibitors	0-1	9.63	3.28 to 16.33
	1-2	10.02	1.81 to 18.54
	2-3	7.05	1.00 to 13.52
H2 antagonists and Proton pump inhibitors grouped	0-1	10.54	- 4.18 to 17.43
	1-2	11.65	2.77 to 20.29
	2-3	10.58	3.62 to 17.61

Table 5.30 contains the recording of referrals and admissions within each cohort. Overall, 11.3% (n=982) of patients had two or more accident and emergency (A&E) admissions in the six months before treatment and 21% (n=1777) had two or more A&E admissions recorded in the first year after treatment. Eighty-six patients had 10 or more A&E referrals recorded in their first year. Similar proportions were seen in subsequent years, with many patients having 10 or more admissions in their record. Recording generally seems to overestimate the number of referrals, so further costing and analysis of these data was not pursued.

The recording of referrals to specialists such as physiotherapist, radiologist and gastrointestinal (GI) investigation seemed more realistic than that of A&E admissions (not shown). One to three GI investigations were recorded for 33 patients in the 6 months prior to initiating the study drug, and 69 patients had one to three GI investigations in the first year after treatment. No patient had more than four investigations in any year. Recording of referrals to physiotherapist was more frequent, with an overall 0.48% (41) of patients having two or more referrals recorded in the previous six months. Of these, one patient had 10 admissions recorded in six months. Seventy-eight patients (0.92%) had two or more referrals recorded in the first year after initiating the study drug and of these, 13 patients had between 5 and 15 referrals in the first year. Referrals to radiologist were recorded for 27 patients (0.04%) who had two to three radiologist referrals in the six months prior to treatment, and 37 (0.05%) patients had two to four referrals in the first year after initiating treatment. The proportion of patients in each cohort with referrals

recorded was very similar over all categories in each year, so formal analysis of the average difference in cost was not pursued.

Incremental cost-effectiveness ratio

An analysis of incremental cost effectiveness ratio was planned for the first study for each study drug compared to the control group. This would have been appropriate in those circumstances where the study drug was more costly and produced greater effects than the no-treatment alternative (see section 5.4.1). The costs comprised estimates of GP visits, drug acquisition costs and costs associated with increased use of concomitant medication. Hip fracture was primary endpoint intended for the economic evaluation.

The results presented in this section did however show that the analysis of the data from the first study did not provide plausible estimates of neither resource use nor clinical effectiveness. Realistic estimates of GP visits were not obtainable, and in spite of the use of multivariate techniques, the analysis of hip fracture did not arrive at unbiased estimates of hip fracture risk reduction associated with treatments. This negated the calculating of incremental cost effectiveness ratios for the interventions under study.

The results of the second study are reported in the next section. Due to the revelations of shortcomings of the first study of the Mediplus® data to provide plausible and valid information on GP visits, referral rates and clinical outcomes, the recording of these outcomes were not pursued in the second study.

Table 5.30. Average (std / range) number of hospital accident and emergency (A&E) admissions by drug by year as recorded in the Medipus® database.

Study period	Bisphosphonate Group		Bisphosphonate Control Group		HRT Group		Tibolone	HRT Control group
	Etidronate	Alendronate	Bisphosphonate	Control Group	HRT_nor	HRT_prog		
6 months prior	1.63 2.22 (0-11)	1.36 1.86 (0-12)	0.35 0.91 (0-9)	0.65 1.29 (0-14)	0.71 1.65 (0-16)	0.59 1.31 (0-13)	0.37 1.04 (0-12)	
1 st year	2.22 3.07 (0-19)	2.17 3.44 (0-39)	0.72 1.56 (0-17)	1.24 2.14 (0-24)	1.18 2.13 (0-26)	1.54 2.84 (0-31)	0.76 1.73 (0-25)	
2 nd year	2.52 3.36 (0-17)	1.85 3.51 (0-33)	0.80 1.65 (0-13)	1.07 1.80 (0-14)	1.20 2.23 (0-22)	1.16 2.09 (0-19)	0.78 1.66 (0-21)	
3 rd year	1.85 2.11 (0-7)	1.75 2.57 (0-19)	0.70 1.67 (0-16)	1.12 1.96 (0-16)	0.98 1.78 (0-10)	1.42 2.29 (0-16)	0.69 1.47 (0-19)	

5.5.2 Results of the second study

The second study was developed to replicate those components of an economic evaluation that the first (pilot) study indicated was feasible for the purpose. Therefore the second study had a more limited scope, and was not designed with the objective of arriving at a full cost-effectiveness analysis. The study evaluated patients using raloxifene, etidronate and alendronate according to the outcomes of attrition rate and cost of concomitant GI medication.

Patients

Two-thirds of the 2076 patients included in the **second** study belonged to the age-matched control group (see *Table 5.31*). By design, all patients had a follow-up time of six months. None of the patients in the study drug cohorts had a clinical history of osteoporosis recorded in their patient records in the six months prior to inclusion in the study, but 22 patients (1.59 %) in the control group had this history.

Table 5.31. Number of patients and average patient age in each cohort (sd)

	Number of patients	Average age (sd)
Alendronate	409	69.22 (12.1)
Etidronate	59	68.22 (12.4)
Raloxifene	224	61.79 (9.8)
Control group	1384	66.73 (11.9)

Treatment pattern

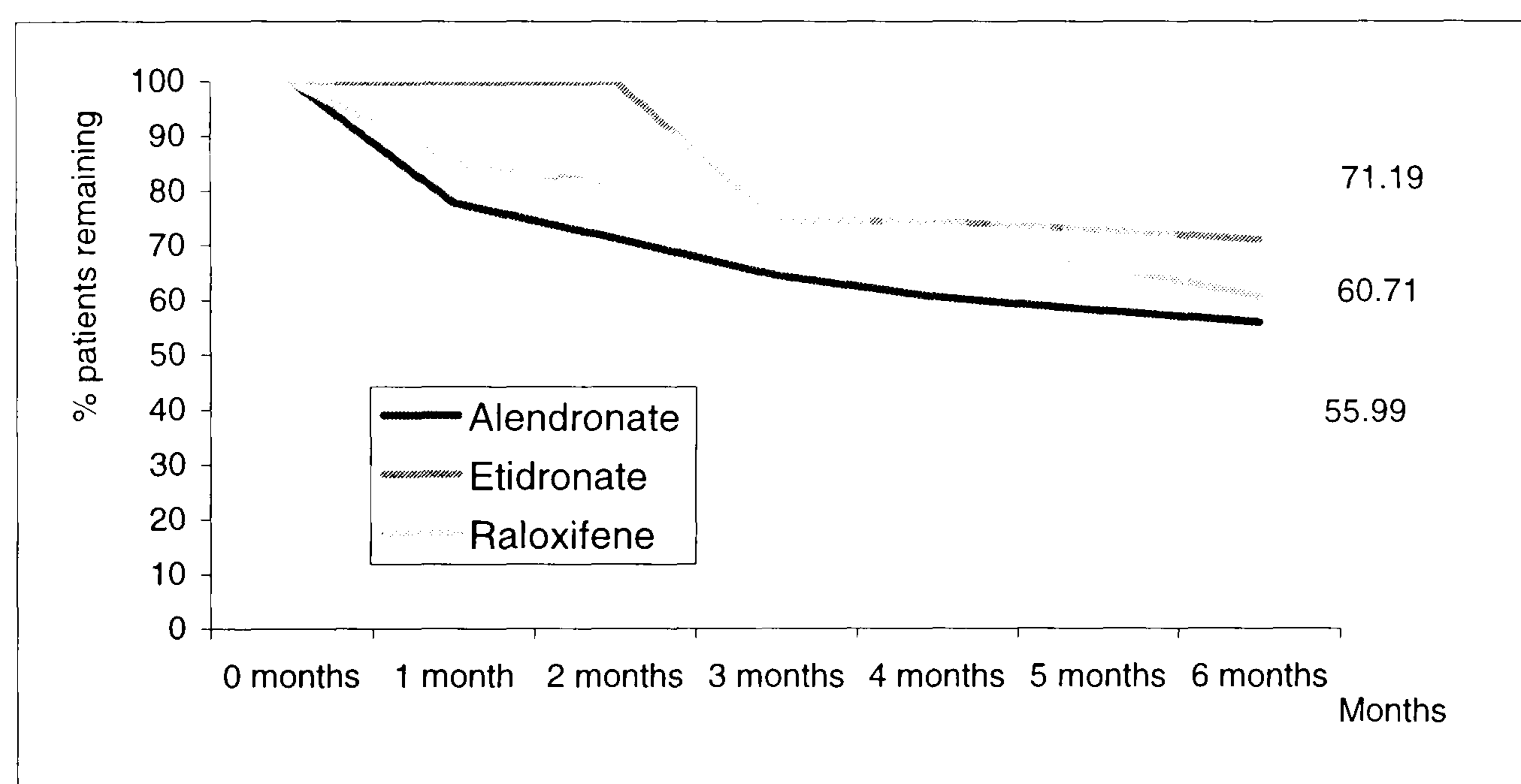
The proportion of patients remaining on alendronate six months after first prescription was 56%, significantly lower than the 71% remaining on therapy with etidronate (*Table 5.32, Figure 5.6*). The risk of stopping therapy with alendronate was 38% (-7 % to 63%) higher than that of raloxifene, though this difference was not significant at a conventional level.

Table 5.32. Cox proportional hazards analyses of time to stop among patients using raloxifene or bosphosphonates.*

Variable	Risk Ratio (95%CI)	p-value
Alendronate versus Raloxifene	1.12 (0.94 to 1.60)	0.135
Etidronate versus Raloxifene	0.63 (0.37 to 1.07)	0.085
Etidronate versus Alendronate	0.53 (0.315 to 0.87)	0.013

*The 3-month of drug free period was used as definition of stop date.

Figure 5.6 Attrition rates for the study drugs at six months using 3-month definition for stop date.



Resource consumption

The average costs of gastrointestinal (GI) drugs in the six months prior to index date were somewhat higher in the bisphosphonate group compared with the corresponding control group (table 5.33). The average cost in the previous six months for H2-antagonists was £46.50 in the control group, £66.50 in the alendronate group and £77.35 in the etidronate group. The average cost in the previous six months for proton pump inhibitors was £118.37 in the control group, £137.37 in the alendronate group and £128.87 in the etidronate group.

During the first six months of bisphosphonate use, the average cost of H2 antagonists increased by £1.48 (95% CI 0.13 to 3.01) when compared with the controls. The average cost of proton pump inhibitors saw a significant increase of £5.48 during the first year (95% CI 2.15 to 8.79) when compared with controls.

Figure 5.7 Costs for concomitant medication with GI drugs.

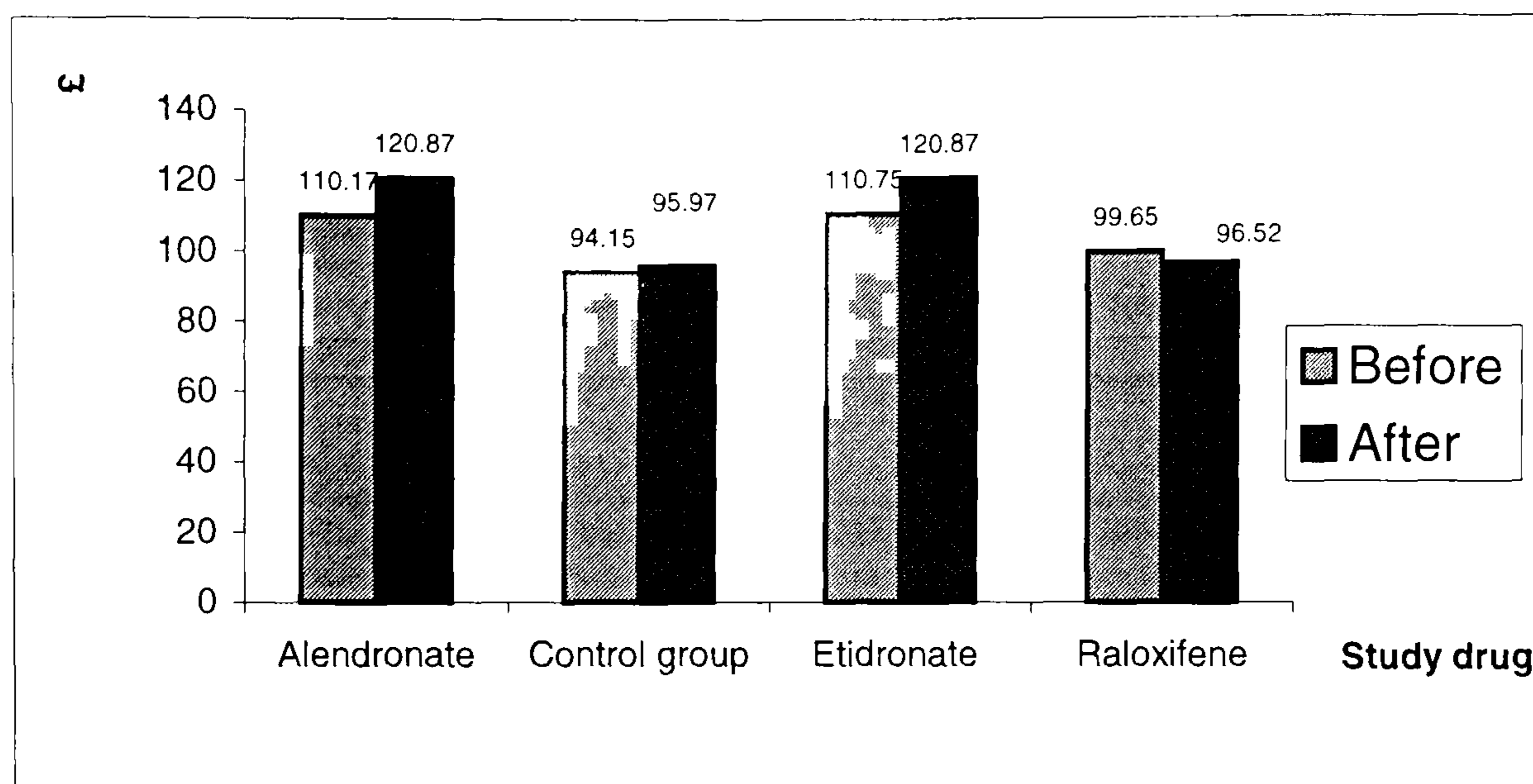


Table 5.33. Increase in GI drug costs for patients using the bisphosphonates between prior 6 months and following 6 months.

GI drug	Mean difference to control group (£)	95% CI
H2 antagonists	-1.48	-3.01 to 0.13
Proton pump inhibitors	5.48	2.15 to 8.79
H2 antagonists and Proton pump inhibitors grouped	3.96	0.66 to 7.48

The differences in GI costs for etidronate and alendronate were also analysed separately (*Table 5.34 and Table 5.35*). No significant changes in GI costs could be observed between the six months before patients initiated either of the bisphosphonates.

Table 5.34. Increase in GI drug costs for patients using **Etidronate** between prior and first year, and the subsequent follow-up years.

GI drug	Mean difference to control group (£)	95% CI
H2 antagonists	-4.59	-8.85 to 0.08
Proton pump inhibitors	4.54	-7.48 to 15.84
H2 antagonists and Proton pump inhibitors grouped	0.02	-11.35 to 10.60

Table 5.35. Increase in GI drug costs for patients using **Alendronate** between prior and first year, and the subsequent follow-up years.

GI drug	Mean difference to control group (£)	95% CI
H2 antagonists	-1.04	-2.61 to 0.39,
Proton pump inhibitors	5.62	-2.25 to 8.94
H2 antagonists and Proton pump inhibitors grouped	4.54	-1.09 to 8.10

The average acquisition costs of the study drugs are reported in *Table 5.36*. The costs were based on all patients, whether they stopped prematurely or whether they continued until the end of follow-up. The average cost of etidronate was about £81, and of alendronate £126. Standard deviation was very large in both groups. Patients initiating therapy with raloxifene generated an average cost of therapy of about £108.

Table 5.36. Average acquisition costs for the study drugs.

Drug	Average cost (std)
Alendronate	125.63 (74.57)
Etidronate	81.08 (31.22)
Raloxifene	108.00 (57.08)

5.5.3 Summary of differences between the two studies

The studies reported in section 5.3.1 and 5.3.2 included patients over two different time periods. The study period of the pilot study (the first study) was January 1st 1996 to December 31st 1998 and the study period of the second study was August 1st 1998 to January 31st 2000. The data extraction for the first study was done in March 2000, and the extraction for the second study was done in September 2000. The second study was undertaken in order to analyse patients on raloxifene according to aspects of the Mediplus® database that proved of value in the first (pilot) study. The aspects included in the second study were treatment pattern and cost of concomitant gastrointestinal (GI) medication for the treatment of suspected side effects. Patients initiating therapy on raloxifene, etidronate and alendronate were included in the second study. By design, patient inclusion criteria were virtually identical. The outcomes investigated in the two studies overlapped, so the results of the analysis of patients on bisphosphonates were expected to be comparable, within a reasonable range of variation due to chance.

However, the findings from the studies outlined above revealed large discrepancies. Firstly, the estimate of the degree to which GPs ceased to issue prescriptions of the bisphosphonates (attrition rate) differed profoundly between the studies. *Figures 5.3* and *5.6* highlight this finding. The proportion of patients remaining on therapy with etidronate was 61.1% after six months in the first study and 71.2% in the second study, both when using the 'three month' definition of stop date in the first study. In contrast however, the proportion of patients remaining on therapy with alendronate

was as high as 92.9% in the first study, but dropped to 56.0% in the second study. Secondly, the estimate of cost of GI drugs differed between the studies, both at baseline and in terms of change over the period during which patients had been exposed to the study medications. *Figures 5.5 and 5.7* highlight these differences. For example, the estimated average cost of GI drugs in the 12 months prior to prescription of etidronate in the first study was £51.92, whereas the estimated average cost of GI drugs in the six months prior to prescription of etidronate was £110.75. The estimated baseline level differed by more than 300% - a difference that was unlikely to be attributable to chance.

Action was taken to verify the origin of these unexplained discrepancies. A comprehensive description of the datafiles, analysis and results was presented to IMS Health in March 2001. Upon presentation of the discrepancies, the consultancy attempted to compare the SAS code that was used for the two extractions. Two different analysts at IMS Health had conducted the data extraction from Mediplus® for the datafiles for the first and the second study. The analyst that conducted the first data extraction had since left the company and the remaining staff was unable to verify the SAS code that had been developed by that individual. After reviewing their internal data extraction notes, the consultancy confirmed that the data for the **first** study was faulty (oral communication with Mary Thompson at IMS Health on August 21st 2001). The consequences of this finding are discussed in the next section.

5.6 Discussion

Health care systems increasingly aspire to provide cost-effective health care to its populations. A challenge for researchers is to develop methodologies for adoption by health economists wishing to undertake economic evaluations where the findings have a high external validity. Retrospective observational databases score in theory highly on speed and convenience of conduct and external validity of the results. Such studies can therefore be carried out quickly compared to the time it takes to set up a prospective study or a trial. However, the use of observational data for the purpose of economic evaluation introduces a series of challenges. Specifically these include provision of relevant clinical and resource variables, provision of valid recording of clinical events and resource use, and enabling the use of appropriate analysis methods to arrive at unbiased estimates of costs and effectiveness.

This study set out to evaluate a patient record database for its feasibility to providing data for a full economic evaluation in primary care setting. This aim was met through the development of design and evaluation of two studies of therapies in osteoporosis using the UK Primary Care Database (UK Mediplus®). Specific objectives of the first study described in section 5.2.1 were to assess the risk of hip fracture, discontinuation rate and primary care resource use amongst women in general practice setting in the UK who initiated therapy on individual osteoporosis drugs. The objective of second study was described in section 5.2.2 and this sought to evaluate a second subset of patients by outcomes deemed feasible by the first study.

Section 5.4.1 reviews whether the Mediplus® database in its current form fulfilled basic data requirements for a full economic evaluation study design. Section 5.4.2 discusses the study results in view of whether they provide plausible and valid information on costs and outcomes of medication use in clinical practice. Finally, the discussion turns to answer the bigger question of overall feasibility of Mediplus® for economic evaluation. The results of the two studies, the first pilot study and the second replication study, are discussed parallel throughout, and implications are drawn of the fact that the data for first study was flawed on delivery from IMS Health.

5.6.1 The merits of the database for economic evaluation study design

The first phase in the evaluation of Mediplus® involved an assessment of whether the database fulfils data requirements to economic evaluation. For the conduct of a relevant and comprehensive economic evaluation in a database, the GPs and patients need to be representative of the general population; patient level resource consumption in the sectors relevant to the perspective of the analysis must be identified, measured and valued; clinical consequences at patient level must be identified and measured; and finally, it must be possible to identify an appropriate control group.

There were no substantial demographic differences between the GP panel and the general UK population of GPs. (IMS Health 1998) At first glance this may indicate that the database is representative of the general pool of GPs. However, there may be reasons that the GPs on the panel are not typical of UK GPs after all. Firstly, the

panel was initially set up based on practices that found the transition to a computerised record keeping acceptable. Before entering the panel supplying data to Mediplus® the GPs undertook a trial period. The GP practices were then selected on the basis of their adherence to the entry regulations of the database. (IMS Health 1998) Secondly, the GPs contributing to the Mediplus® database are younger than the average GP population. There may be reason to believe that the practice of younger GPs deviate from that of more senior colleagues. Perhaps these would pick up new therapy traditions more easily and may also be more computer skilled. Finally, GPs on the Mediplus® panel are regularly monitored and subject to regular quality audits. This may increase completeness of the data, but it introduces a selection bias in that only those GPs participate that are likely to be compliant, computer skilled and younger than average. Taken together, these factors suggest that the recorded information in Mediplus® may systematically differ from the care that is provided to the general population. Whether this influences the content and quality of care is unsure, but it may limit the scope for extrapolation of study results.

The designs of the feasibility studies were based on patient inclusion by drug exposure, age and gender. The database enabled collection of patient-level information on variables covering relevant clinical conditions at baseline, incidence of relevant clinical outcomes and indicators for use of resources such as GP visits, referral rates and cost of medication. The database enabled valuation of resource use directly by providing wholesale prices for issued prescriptions.

In the context of an observational study, where exposure is assigned through non-random selection, it is essential to match by certain features or adjust for the presence of explanatory factors that may impact the effect of treatment. In a study of osteoporosis drugs it would be desirable to identify a control group by matching with the clinical condition of osteoporosis. In preliminary exploring of the Mediplus® data it became apparent that very few patients had the diagnosis of osteoporosis recorded if they were not treated for the condition. This may not necessarily be a shortcoming of the database as such but may rather be a result of clinical practice being 'activity driven'. For example, why would a GP record the presence of osteoporosis if the patient does not receive treatment for it? In addition, findings of the first study indicated that relevant clinical conditions may indeed not be routinely recorded even when action towards the condition is taken. For example, only 46.3% of the prescriptions of etidronate and 57.3 of the prescriptions of alendronate were linked to a problem of osteoporosis in spite of the indication for these drugs are prevention

and treatment of osteoporosis. For etidronate and alendronate patients, only 59.9 % and 73.4 % respectively, had a diagnosis of osteoporosis in their records. Studies undertaken in a database where recording is activity driven, and where relevant clinical information may be omitted, there is a risk that the measurable is made important rather than the important being measured.

In summary, the database may present a biased GP population, it enabled identification of key clinical and resource use indicators in the patient records, but activity-driven patient records may result in the data being less comprehensive than anticipated. This latter point had particular consequences for the identification of covariates for the modelling of effect estimate, as discussed in the next section.

5.6.2 Assessment of the study findings

The further analysis of the individual data components in the first study indicated that some aspects of the data seemed implausible and therefore limited the scope for a full economic evaluation. This was particularly evident for the evaluation of clinical outcomes, such as hip fractures. Therefore, only concomitant medication and attrition rates were reiterated in the second study.

Evident from the results of the analysis of the two studies was the profound differences between the two studies on those outcome variables that were common to the studies. Both the attrition rates and cost of gastrointestinal (GI) drugs differed by more than should be expected by chance and they were considered sufficiently large to contact IMS Health for verification of the findings. IMS Health, who undertook the data extraction for both studies from the Mediplus® system, assigned these differences to problems in the extraction of the first study, and it was confirmed that the extraction of the datafile according to design of the **first** study may have been flawed. The consultancy was however unable to confirm exactly which aspects of the study was flawed, as the SAS code that was used to extract the first dataset had been produced by a researcher who had left the organisation since. It was emphasised that the contents of the **second** study were reliable.

All data from the first study had been analysed and written up at this stage, and parts of the data had been presented. (Urdahl et al 2000, appendix 6) The second study

had been undertaken to reiterate feasible aspects of the first study. In addition to providing information on a second subset of patients, the second study served to highlight shortcomings in the procedures to ensure the quality of the data extraction from the Mediplus® database. The discrepancies between the studies and the implications are discussed parallel to the assessment of the study findings in this section.

Fracture risk and effect estimate

During the analysis of the first study, multiple recording manifested itself as a problem with the database. This was documented by the fact that six patients had two hip fractures recorded, three patients had four hip fractures recorded and one patient had six hip fractures recorded in her record. In an attempt to overcome this problem, the number of patients with a hip fracture rather than number of fractures in the sample was used as unit of analysis. Overall, 49 patients experienced a fracture after the index date in the whole sample of 8470 patients, this was 0.6% of the population, or about 6 per 1000.

Fracture was more frequently recorded in the records of those patients in the bisphosphonate group, where 3.6% of alendronate patients and 1.2% of etidronate patients had hip fracture recorded in their records. In comparison, 2.5% of the women on etidronate have hip fracture prior to inclusion in the study, and prior fracture was present in 4.4% of the women on alendronate. Hip fracture was less frequently recorded for those patients receiving the HRT therapies, with proportions of recorded fractures between 0.2% and 0.3% prior to index date and between 0.11% to 0.37% after index date. The indications for these drugs may to some extent explain these differences, as bisphosphonates are indicated for both treatment and prevention of osteoporosis, whereas hormone replacement therapies are indicated for osteoporosis prevention.

In clinical practice, one would expect treated patients to differ from untreated patients, as the former would have an explicit indication for the treatment. To the extent that the indication is related to the outcome variable as well, the indication can function as a confounding variable. Matching by indication and stratification may be used in observational studies to control for this confounding by indication. This was however impossible in this study due to the 'activity driven' recording resulting in diagnosis having a higher likelihood of being recorded in the database when action is

taken on the condition. The absence of an appropriate control groups against which to demonstrate the impact of the treatments limited the validity of the analysis of clinical outcome. For example, only 1.1% to 1.8% of the control group had osteoporosis recorded in their patient records and 0.3% to 1.1% had a previous hip fracture recorded in their records. From the outset this provides indication that the susceptibility of fracture in the control group may have been lower than that in the treatment groups.

Recognising that observational studies inherently are susceptible to bias and confounding factors, which are the major threats to study validity, the design incorporated collection of information on factors that may have influenced the validity of the effect estimates. Age has previously been identified as a covariate of hip fracture. (Dennison and Cooper 2001) For the purpose of arriving at an unbiased estimate of treatment effect, several other variables were tested for their predictive power of hip fracture. Only age and the recording of a previous fracture featured in the final model of time to fracture between the groups. Previous hip fracture increased the relative risk of having a hip fracture recorded by a spectacular amount (Hazard ratio 108.3, 95% CI from 58.4 to 200.8), and age above 75 increased the relative risk of having a hip fracture by nearly 300% (Hazard ratio 3.5, 95% CI from 1.9 to 6.5).

These implausible results indicate that it was impossible to arrive at an unbiased estimate of fracture risk in the cohorts. When the treatments were added to the final model and compared to controls, none of the drug treatments reduced the incidence of fracture significantly. In fact, the analysis estimated that alendronate increased the risk of hip fracture significantly by as much as 95%. This should merely be seen as a result of the non-randomised design of the study and selection of high-risk women into the alendronate group rather than indication that alendronate actually is a risk factor for hip fractures in women. In comparison, randomised clinical trials of alendronate have estimated a hip fracture risk reduction of about 49%. (Black et al 1996) No randomised clinical trials of etidronate, tibolone or the HRTs have been powered to show efficacy on hip fracture as primary endpoint. (Royal college of Physicians 1999)

In summary, the analysis of clinical effectiveness in terms of risk reduction of hip fracture was unfeasible in the Mediplus® database for two main reasons. Firstly, it was not possible to compare fracture risk to a matched control group and therefore

the effect estimates are biased. Secondly, multiple recording of the same fracture was a problem, and analysis of hip fracture was undertaken based on the number of women experiencing a fracture rather than the absolute number of fractures. This did however not control for the artefact that would be created if a fracture was experienced before index date and then re-recorded after index date, a situation that would be picked up by this study design as a new fracture after initiation of drug therapy. This phenomenon may undermine the predictive value of previous fracture in the models by resulting in the prediction of a fracture with a previous record of the same event.

Two papers published in the BMJ in the autumn of 2000 illustrated similar key issues in observational studies undertaken in routinely collected patient record databases. Both studies, Farmer et al (2000) and Jick et al (2000) analysed the incidence of venous thromboembolism before and after the warning from the UK Committee on Safety of Medicines about third generation oral contraceptives using the computerised patient record database General Practitioners Research Database (GPRD). Conducting a time-series analysis Farmer et al (2000) found that the incidence among pill users had not dropped. By means of a cohort analysis with a nested case-control Jick et al (2000) found that, both before and after the warning in October 1995, the risk of venous thromboembolism in women using third generation oral contraceptives was about twice that in users of preparations using levonorgestrel. Moreover fewer cases occurred after the warning than would have been expected if the prescribing of oral contraceptives had not changed.

These two studies arrived at opposite conclusions on the same research question using the same database, and the incidence provides important learning points for database studies. (Skegg 2000) The study undertaken in the Mediplus® database is no exception. It was not possible to review individually all cases. It was not possible to provide any details of validity and specificity of the diagnoses of the hip fracture or any other clinical diagnosis. Computer-recorded diagnosis of hip fracture and other clinical conditions were used, without documentation from clinical records or hospital data. The potential for misclassification bias is substantial. The analysis of clinical endpoints was therefore not pursued in the second study of raloxifene and the bisphosphonates.

Drug therapy pattern

Compliance with therapy is generally higher within the context of controlled clinical trials than in usual care settings, and as such can be a contributing factor to the cost-effectiveness of a drug when used in a clinical practice setting. (Hughes et al. 2001)
Clinical guidelines in osteoporosis recommend long-term treatment in order to experience the full preventive effect of the prescribed drugs. (Royal college of Physicians 1999)

The impact of discontinuation on the cost effectiveness ratio is likely to be greatest where the effect of discontinuation is that the costs increase and the benefits decrease. When treatment is not taken for long enough to experience preventive effect of treatment resources are lost to society without benefits gained. This may be the case with preventive treatments in osteoporosis, where long-term adherence is key to experience preventive effect from treatment. Unfortunately, failure to persist with therapy is particularly prevalent with asymptomatic diseases. For example, Jones et al (1995) estimated that a third of patients discontinue antihypertensive drugs during the first year of treatment, and between 40% and 50% of the patients may change or discontinue therapy after 6 months. It should be noted that in those cases where the patients experienced side effects from treatment discontinuation might in fact be clinically appropriate and cost saving in the long run.

Cessation of therapy is not routinely recorded in the Mediplus® database. It was therefore essential to develop a set of definitions to estimate the date at which a patient stopped taking the medication. The imputed date at which the patient was anticipated to stop therapy was based on the duration of the last prescription of study medication and the date at which a new prescription should have been issued. Two assumptions were used to derive two definitions of stop date in the first study. The first stop date definition assumed that patients who had a new prescription issued within **three** months of the last day of therapy was a continuous user. In other words, those patients for whom a new prescription was not issued within three months of last prescription qualified as having stopped therapy. In order to assess the sensitivity of this assumption, the second stop date definition assumed that patients who had a new prescription issued within **six** months of the last day of therapy was a continuous user.

Both studies evaluated attrition rate as an outcome. The first study revealed that the estimate of attrition rate was very sensitive to the definition of stop date. For example, by using the main indicator of stop date (three months drug free period allowed), 48.1 % of women initiating therapy on etidronate and 10.1 % of the women initiating therapy on alendronate stopped taking the drug within the first 12 months. The risk of stopping therapy with etidronate was 5.31 (3.84 to 7.36) times that of stopping therapy with alendronate. The results of using the 3-month stop date strongly suggested that patients using etidronate seemed to tolerate treatment less than those initiating therapy on alendronate resulting in a very poor real-life adherence with etidronate. In comparison, by allowing a six-month drug free period before a stop date was imputed, the analysis showed that 16.0% of the etidronate patients and 4.6% of the alendronate patients stopped therapy within the first year. The average time to stop was significantly shorter for etidronate patients than that of alendronate patients also with this definition of stop date. The attrition rate for the hormone replacement therapies and tibolone remained similar between the two stop date definitions. The second study provided different results. By using the main indicator of stop date (three month drug-free period) the second study indicated a drop out rate of 29.3% after 6 months for patients initiating therapy with etidronate and 44.0% for patients initiating therapy with alendronate.

A limitation of this analysis is the fact that Mediplus® is only recording issued prescriptions, and provides no insight as to whether the prescriptions are dispensed at the pharmacy or not. Beardon et al (1993) found that 5% of 20,000 prescriptions issued from a large UK general practice were never actually redeemed at the pharmacy. Approximately 80% of the unredeemed prescriptions were for asymptomatic conditions, and the remainder for prophylactic drugs. Furthermore, once the prescription has been redeemed from the pharmacy and the acquisition cost to society has occurred there is no guarantee that the medication is actually taken by the patient.

Resource consumption

The first study attempted to identify resource use indicators for GP visits, drug acquisition costs, costs of concomitant medication and indicators for referrals to secondary care. Analysis of these variables provided mixed results. Medication costs derived from the database seemed plausible and enabled analysis. This aspect was therefore pursued in the second analysis. The analysis of the other

resource indicators, such as GP visits and referral rates, did not encourage faith in the data and they were therefore not included in the second study.

The Mediplus® database does not record directly whether the patient is present in the surgery or not when an activity is undertaken that results in information entry to the patient record, and only a series of dates at which an action was taken is recorded in the patient records. For that reason, one date can be recorded several times, and activity not directly involving the GP, such as recording of laboratory results and repeat prescriptions, receive equal weight. IMS Health has developed a formula in the database to estimate number of GP visits from the number of dates. In order to evaluate this formula, both the number of dates recorded in the records as well as the estimated number of GP visits was analysed. The results revealed high similarity between the dates and estimated GP visits, resulting in implausibly high estimates of GP activity. It was therefore not possible to make meaningful assessment of the number of GP visits by the patients in the observation period, and this resource consumption item was not reiterated in the second study.

The cost of treatment was recorded in both studies to estimate the cost of acquisition drug. The recording of prescriptions for medications is close to 100% and the system provides a calculation of cost of the prescribed amount directly based on MIMS. (IMS Health 1998) During analysis of this outcome it became apparent that the cost estimate not necessarily is informative. The acquisition cost varied hugely within each study group, for the reason that follow-up time and attrition rate varied within the cohorts. Therefore, this analysis was unable to estimate treatment costs relative to each other. This illustrates difficulties in developing appropriate variables for censored cost data rather than revealing a shortcoming of the database.

Concomitant medication costs may be an important consideration when undertaking an economic evaluation for osteoporosis in primary care. The costs of treating gastrointestinal side effects associated with the use of the bisphosphonates were analysed using appropriate methodology for the analysis of highly skewed cost data. The bisphosphonates were grouped in the main analysis of the cost of gastrointestinal (GI) side effects. The average costs of H2 antagonists and proton pump inhibitors (GI drugs) were £44.4 in the bisphosphonate group and £15.2 in the control group. During the first year of bisphosphonate use the average cost of GI drugs increased by £11.3 (95% CI 5.3 to 17.5) when compared to the difference in the control group over the same time period. Similar increases were observed in the

subsequent years. The analysis therefore provides indication that the bisphosphonates are associated with a 20-25% increase in the cost of GI drugs in the first year after initiating therapy, however when analysed separately this increase was not significant.

The absolute number of women taking GI drugs increased between the prior and the first year of treatment with the bisphosphonates, but this increase was not significant. In the subsequent years the increase is not that remarkable in spite of the fact that the average cost of GI drugs increase, also in these years. This may indicate that those women who already taking the GI medication when initiating therapy with the bisphosphonates increased on average their use of the drugs. This may provide indication that caution should be exercised when prescribing bisphosphonates to those women that are already taking GI drugs, and consequently are susceptible to GI problems. These findings confirm those of other observational studies. (Ettinger 1998) The findings of van Staa et al (1997) indicate that etidronate does not induce GI side effects. This is also confirmed in the separate analysis of the bisphosphonates.

The second study provided a very different picture of GI drug costs altogether. Firstly, the baseline costs of GI drugs was substantially higher in the second study, with an average cost of GI drugs in the prior six months of £94.2 and £110.8 and £110.2 for the control group, etidronate takers and alendronate takers, respectively. In comparison, the corresponding estimates were £15.2, £51.9 and £42.0 in the first study. There was no significant difference between GI drug use in the prior period and the first six months after initiating therapy, neither when the bisphosphonates were analysed grouped nor separately.

The same limitations to the use of prescriptions as a measure of treatment discontinuation apply in the analysis of concomitant medication. There is no information about whether the issued prescriptions are collected at the pharmacy and drugs collected at the pharmacy may not be used in the last instance. An additional limitation of the analysis of the first study is the fact that time period prior during which resource consumption was quantified prior to index date was only six months, whereas the costs in subsequent years were accumulated over intervals of 12 months. The costs in the prior six months were multiplied by two in the first analysis in order to escalate costs to a full year. It is not clear whether this could have biased

the analysis in any systematic manner and in that case which direction the results were skewed.

Finally, the analysis revealed some limitations of the meaningful assessment of referral and admissions to secondary care. Multiple recording seemed to be a problem with accident and emergency admissions, and for example as much as 86 patients had 10 or more referrals in their first year. The database therefore failed to provide plausible data for assessment of costs in other sectors than primary care.

In summary, it was not feasible to analyse indicators of GP visits, referral rates or to provide a relevant presentation of acquisition costs for study medication. This prevented inclusion of these variables in the second study. The first study did on the other hand suggest that analysis of resource use associated with treatment of side effects seemed feasible. When comparing the results of the studies however, discrepancies emerged, which questioned the validity of the first evaluation altogether.

5.6.3 Feasibility of Mediplus® for economic evaluation

Specific research objectives in the two pilot studies presented in this Chapter were linked to the study of fracture risk reduction and resource consumption associated to use of osteoporosis drugs in the UK clinical practice setting. Problems with the data became apparent during analysis, and these prevented answering of the specific research questions set at the outset of the study. Did the two studies presented in this chapter enable answering the bigger question of whether the Mediplus® system is a feasible vehicle for the clinical and economic evaluation of pharmaceuticals?

The first study was comprehensive and included variables that may have enabled economic evaluation and resulted in a series of cost-effectiveness estimates for the study drugs. However, the analysis of the individual resource consumption components and the clinical effectiveness estimates provided implausible results. Similar comprehensive data extraction for the second study was not pursued due to the discouraging results of the first study, and the second study was not designed to enable full economic study. Therefore, the confirmation from IMS Health on August 21st 2001 that the first study was flawed provided a dilemma. Would an economic

evaluation have been possible if the second study had been comprehensive? Would an economic evaluation have been possible if the first study had been un-flawed? Three main aspects that discourage the use of Mediplus® for economic evaluation are summarised below.

Firstly, the fact that unreliable data could be released from IMS indicated a shortcoming in the internal quality assurance procedures for researchers at IMS. The problem turned out to be that one analyst did all the data extraction on his own and wrote SAS code that was unreadable to the other IMS researchers. A second researcher did the data extraction for the second study. The result was two datafiles that differed profoundly in spite of the fact that they were meant to measure the same outcomes. The problem detected by comparison of the two studies was not primarily an issue with the data, but rather revealed flaws in the data management process.

Secondly, it was not possible to derive estimates of cost-effectiveness of pharmaceuticals from the database because of the failure to provide valid estimates of clinical effectiveness, and comprehensive estimates of costs. The issue of multiple recording of clinical events and administrative activity such as referrals and admissions disclosed in the first study were not likely to be due to faulty data extraction of these variables. The gathering of observational data for study can be flawed by bias at many stages, including subject selection, data recording, accuracy in data entry, data construct and definition of analytical variables. The findings therefore cast doubt about whether the data recording is complete and sufficiently structured to enable valid economic evaluation. The development of quality audits in routinely collected datasets may provide opportunity to minimise the vulnerability of observational studies to bias resulting from non-random assignment of patients to therapy. (Black 1999)

Finally, the scope for using statistical methods to control for bias and confounding to evaluate costs and clinical outcome was restricted. In clinical research there is a need to demonstrate that changes in costs and outcomes are caused by, rather than merely associated with, the intervention under investigation. The database did not provide sufficient level of data quality to enable the use of modelling techniques methodologies to control for observed and unobserved confounding variables. Statistical methods to control for observed and unobserved biases can only be applied if the data are collected and managed at a certain quality level ensuring valid recording. No amount of sophisticated modelling is able to counteract invalid data.

Methods such as the use of instrumental variables and propensity scores have added value to other analyses of observational data. (Radford et al. 2001, McClelland et al 1994) In this study however, statistical methods could not have minimised the bias introduced by lack of randomisation, simply because the recording of outcomes and covariates was unreliable.

The advantages of using data reflective of clinical practice for the purpose of economic evaluation have been outlined elsewhere in this thesis. In spite of the fact that the potential advantages of study in the Mediplus® database for economic evaluation may not exist in reality, the studies outlined in this chapter made the best possible use of available observational data. The two studies illustrated ways in which an observational patient record database can be used to study aspects of cost effectiveness in clinical practice. The second study provided plausible indicators of attrition rate and cost of concomitant medication. Some aspects of the database may therefore have value in economic evaluation.

5.7 Conclusions and recommendations

This feasibility study in the Mediplus® database was limited to the therapy area of osteoporosis. Therefore, generalisation to other therapy areas may be restricted. Replication of the study is therefore encouraged in other therapy areas that are considered different from osteoporosis on key clinical aspects.

The (first) pilot study did not produce unbiased estimates of clinical effectiveness and resource use and failed to provide a full economic evaluation. Importantly, however, the aim was to assess whether the Mediplus® database is feasible for the economic evaluation of pharmaceuticals in UK clinical practice, and this aim was met.

Conclusions

Health economic researchers need to explore ways in which the needs and requirements of decision-makers can be satisfied. The choice of methodologies is made in a time and money constrained environment. The approach taken requires a series of choices and compromises on the part of the researcher. One challenge is to select a strategy that is capable of generating information that have an acceptable

balance between internal and external validity, between immediate solutions and long-term study and between local information and broader policy setting.

Observational databases provide vast amounts of real-life data that may have a potential value for determining whether the benefits expected from the evidence provided by RCTs are achieved then new pharmaceuticals are adopted in the wider context of clinical practice.

A design in the Mediplus ® database was successfully developed that aimed to demonstrate benefits of pharmaceutical treatments in osteoporosis outside of clinical trials. Findings in the second, reliable, study indicated that concomitant medication use and attrition rates may be evaluated using the data. Provided that research questions and variables are developed with caution, the database may be a valuable source of input for example to economic models. It is important that the question of the potential of routine data should not be dismissed because of the inadequacies of circumstances of this study.

The Mediplus® data are stored in a complex manner, and it is a specialised task to access and extract data. The findings of this study indicated that checks and quality assurance procedures have been inadequate. In this particular case, the individual(s) that was responsible for the data extraction according to design was (were) 'the weakest link', and the fact that the data not reproducible throws doubt on the validity of the studies. This is not to suggest that the database were infeasible for all aspects of economic evaluation. Rather, it has emphasised the importance of having appropriate quality assurance steps in place to prevent unscientific data management processes. It would be premature to recommend wide use of the database for components of economic evaluation analysis before steps are taken by the database providers to ensure valid data extraction.

It would not be relevant to argue that observational data in any sense should – or can – replace randomised controlled trial (RCT) estimates of drug efficacy. RCTs are the appropriate vehicle for assessing outcomes of treatments in all stages in the life cycle of a drug. However, it is time to question the consistent use of randomised data only for the effect estimate within economic evaluation. Rigidly designed studies in good-quality comprehensive observational databases can be a useful supplement for trial-based estimates of cost-effectiveness.

Recommendations

Some barriers would need to be overcome if more studies were to be undertaken using data from the Mediplus® database. The primary issue for further consideration is that of data quality and management. The experience generated through this research may provide constructive guidance on the use of observational data. How could problems of this sort be avoided in the future?

- A carefully written protocol should describe the extraction of data from the database. This could require that two people independently conduct the data extraction according to the protocol. The same subset of patients and variables should result from the two extractions and comparison would provide a quality-check. The introduction of this quality step could be implemented at a marginal cost to the database providers, and it would in the long run contribute to increasing the faith in observational data, a faith that this study has jeopardised. To guide the conduct of clinical trials, the good clinical practice (GCP) regulations apply to all commercially sponsored clinical trials and provide guidance to the data entry process, subject monitoring, and source verification. (ICH 1996) These could aid the development of quality assurance step for data management systems.
- In order that the Mediplus® database could be used on a larger scale for full economic evaluation, procedures for valid recording of clinical events implemented across the whole panel of information suppliers (GPs) should be in place. Multiple recording was a problem in the pilot study of this chapter, as was lack of recording of relevant clinical conditions (for example osteoporosis in patients receiving osteoporosis treatment).
- The recording of clinical outcomes in Mediplus® needs to be validated comprehensively in a rigorously designed validation study. Furthermore, researchers using the data should have the opportunity to go back to the original patient record once the dataset has been extracted to source verify clinical events of major importance to the study results.

At a more general level the findings from the studies presented in this chapter highlight the following:

- None of the current guidelines for cost-effectiveness analysis of pharmaceuticals explicitly state whether observational study designs are an acceptable source for

economic evaluation studies. Such guidelines should take an active standpoint as to when a pharmaceutical evaluation should be based on investigation under ideal conditions (randomised clinical trials) or under conditions of normal practice (observational data).

- The likelihood of poor adherence to therapy in clinical practice should be studied and incorporated in economic models of pharmaceutical interventions.
- Multivariate techniques can be adopted for reducing bias and issues with confounding factors. Such methodologies are making important contribution to epidemiological studies in general and have the potential to add significant value to the use of observational data for economic evaluation. Techniques such as instrumental variables methods and imputing techniques for missing data should be developed to increase the ability to deal with issues of non-random allocation of treatment in observational studies.

CHAPTER 6

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

The aim of this thesis was to assess and further develop research methods to increase the generalisability of economic evaluation studies of pharmaceuticals. Factors that limit the generalisability of economic evaluations across locations include clinical practice patterns, incentive structures for healthcare providers, relative price levels and demographic profile of disease (O'Brien 1997). Empirical studies have provided evidence that these differences may be important. (Willke et al 1998) The sensitivity of cost and clinical estimates to factors that vary by location raises methodological questions on the design and conduct of economic evaluation studies. The objectives of the research presented here were framed in the context of the increasing need to transfer economic evaluation data across settings and the scarcity of empirical and methodological research into the limiting factors and methods to overcome them. Availability of clinical research designs able to reflect context-specific data ranged from controlled pivotal trials through to observational studies, with the methodology of pragmatic trials providing an intermediate degree of clinical practice data through maintaining randomisation but relaxing other design features. The principal focus has been on economic evaluation throughout the thesis.

This chapter considers the extent to which the research objectives were achieved. The first section outlines the empirical findings from the three source projects that constitute the thesis and review the methodological contribution of the source projects in light of the research context outlined in Chapter 2. The second section compares the methods and provides an assessment of the relative merits of each. Finally, some implications of this work are outlined and accompanied by specific recommendations for future research and policy.

6.1 *Contribution to methodological development*

The projects in this thesis explored aspects of three complementary methods for generalising economic evaluations. Firstly, effect sizes from pragmatic trials were examined to assess whether external validity was achieved at the cost of internal validity. Secondly, economic models were assessed according to how well they accommodated contextual differences in resource use and clinical assumptions. Finally the feasibility of an observational database was explored in an attempt to further develop a method for the conduct of economic evaluations with externally valid results.

6.1.1 Pragmatic or explanatory design?

Concerns about the generalisability of randomised controlled trials have been raised in the context of clinical evaluation. (Altman and Bland 1998) In pivotal RCTs, the ‘laboratory’ environment of the experiment provides atypical conditions that may limit external validity of findings, with the result that pharmaceutical interventions are often evaluated under one set of circumstances but implemented in clinical practice under another. Reliance on clinical trial data for economic evaluation has led health economists to investigate the problems of generalisability of economic evaluations undertaken alongside randomised controlled trials. (Drummond and Davies 1991, Coyle et al 1998) These authors have advocated the use of trials adopting a more pragmatic design, through which external validity may be increased by reflecting clinical practice circumstances. It was emphasised in the introduction that a binary concept of ‘pragmatic’ versus ‘explanatory’ trials may not be appropriate. Rather, we may consider a continuum of pragmatism versus explanatory trials, in which investigators adopt design features rendering the trial more, or less, pragmatic (section 2.1.1).

In the first source project, a checklist of ‘pragmatic’ design features was developed and applied to the randomised clinical trials from 318 randomised controlled trials in two therapy areas. Generalised linear models were fitted to assess whether key components of trial design predict the estimate of outcome. The models included a

specified hierarchical factor reflecting the structure of the data, and the final model only included items that predicted outcome at conventional levels of statistical significance.

The study provided empirical evidence that on average, 'pragmatic' design features do not systematically impact effect size estimate. The findings favour the hypothesis that pragmatic trials do not jeopardise internal validity when seeking to maximise external validity to a particular setting. These features may safely be adopted in economic evaluation research, without the fear that the effect size may be biased by design.

In spite of reflecting clinical practice to a larger extent, trials adopting pragmatic features are still specific to the specific healthcare contexts within which it was conducted. The results of pragmatic trials may still be context-specific to the extent that they add limited value to healthcare decisions in other settings.

Because of the way in which economic evaluation incorporates assessment of relative costs and collapses this with an estimate of clinical effectiveness into one outcome measure, the cost-effectiveness ratio, economic evaluations may be particularly sensitive to factors in the study environment that may limit the generalisability of the results. A limitation of this project was that it only examined the clinical effect estimate from pragmatic trials and paid little attention to the numerator (costs) in the cost-effectiveness ratio. Examination of resource consumption collected alongside pragmatic trials would have provided more insight into how resource-use varies between settings.

6.1.2 To model or not to model?

Economic models are simulation experiments evaluating cost-effectiveness based on a series of parameter estimates. These individual parameter estimates are likely to be influenced by the setting in which the input data were derived, and the setting which these parameter estimates are based on do to an extent limit the generalisability of the results of a model. In the second source project, a checklist was devised for assessment of generalisability of economic models that incorporated both clinical and resource consumption estimates. The checklist was intended to facilitate critical

assessment of the generalisability of economic models to judge the study's adaptability to a specific context. The framework may provide guidance to those conducting economic evaluation models on how they can increase the generalisability of their results.

The checklist was applied to economic models published in the area of osteoporosis, with the critical review revealing that very few models attempt to generalise findings to settings outside of one specific context regarding economic assumptions. None of the models published prior to June 1999 assessed a range of unit costs relevant to different settings or countries and most assumed 100% compliance with treatment over substantial periods of time. A few models accommodated for clinical differences such as case-mix in patient population and baseline risks. So in addition to providing a framework for the assessment of models, this project provided empirical evidence that few models actually employ sensitivity analysis to explore the impact of the therapy under evaluation in different sets of circumstances.

The generic framework enables assessment across a range of models in different areas of therapy. Applying the checklist to models in one therapy area facilitated between-study comparison. A limitation of this project was that it was unable to study differences across study results that may have been explained by location, or to study difference in cost-effect estimates over time. The models in this therapy area presented cost-effectiveness ratios by a variety of outcomes, e.g. cost per vertebral fracture avoided, cost per hip fracture avoided and cost per quality adjusted life year. The patient subgroup studied differed across models, as did currency and choice of incremental versus average cost-effectiveness ratio. In order to provide empirical data on how cost-effectiveness ratios travel between settings and across time, the checklist could be applied to models in other areas and regression analysis attempted on a larger and more homogeneous data set. Insight into whether models in other therapy areas have accommodated for variation in setting to a larger extent than models in osteoporosis would test the merits of models as a tool for generalising economic evaluations.

6.1.3 To randomise or to observe?

The need for economic evaluation data of relevance to clinical practice settings led to the investigation of the potential contribution of electronic patient record databases. The aim of the third source project was to evaluate the feasibility of the Mediplus® database to provide data for a full economic evaluation. The therapy area of osteoporosis was used as a model and the study objectives were to combine resource consumption attributable to osteoporosis drugs with real-life fracture data in clinical practice, in an analysis of incremental cost-effectiveness. An observational, retrospective cohort study was designed, including women over 55 using bisphosphonates, HRTs and tibolone. Fractures, hospital referrals, GP visits, concomitant medication and attrition rate over a 3-year long follow-up time were recorded outcomes. A second study was designed, reiterating those aspects of the first design that seemed feasible after analysis, however the second study included patients using bisphosphonates and raloxifene in a different timeframe.

Analysis from the first study revealed limitations with the data in terms of availability of information required for economic evaluation, limitations that negated a full analysis of cost-effectiveness of the study drugs. For example, clinical diagnosis and resource consumption data were difficult to obtain, and a problem with multiple recording became apparent. The first study was, nevertheless, able to generate some data of relevance to economic evaluation, and these were the use of concomitant medication and patient attrition rate in clinical practice. These aspects were therefore reiterated in a second study, extracting a different subset of patients. In spite of similarities in design and inclusion criteria, comparison of the results from the two studies identified major discrepancies, revealing shortcomings with the consultancy's data management procedures.

The observational database was pursued for its prospect of providing two potential practical advantages. Firstly, administration, funding, quality control and considerable infrastructure are necessary for the conduct of RCTs. In economic evaluation, the drugs under scrutiny should be compared to that of alternative treatments representing the current best practice. Over the course of a trial of long duration, the opinion of what is considered 'current best practice' may change. A database study compared to the trial

approach is appealing, as it would enable head-to head comparisons as well as a comparison to 'no intervention', according to what is relevant to a given decision-maker at any time. Secondly, in a clinical trial, observations are recorded on a case report form (CRF). Investigators would seek to reduce the size of the CRF to simplify the process for the physician enrolled, minimise interference and patient burden and increase trial participation. Routinely collected data, in contrast, may contain vast amounts of patient information. The first observational study of this thesis collected information on 180 variables, equating to a clinical trial with a CRF of 180 questions and numerous pages. The richness of the information and the potential volumes of observational studies in longitudinal databases may offer a distinct advantage to clinical trials.

Substantial disadvantages do however discourage use of observational databases for the purpose of economic evaluation. Observational data are susceptible to the influence of known and unknown factors that may systematically bias the effect estimates. The basic function of randomised clinical research designs is to provide a fair and unbiased comparison of outcomes between patient cohorts with and without an exposure that impacts the outcome. (Piantadosi 1997, Pocock 1984) Fundamentally, though, an observational study shifts the focus of the analysis from attributing a causal relationship between intervention and outcome to identifying associations between the intervention and changes in costs and outcomes. It is true for observational studies as well as clinical trials, that estimates from any evaluation study are only as good as the data themselves. The prospective nature of randomised trials enables special care to be taken to maximise valid, complete and reliable recording of information, (ICH 1996) whereas the recording in observational databases is subject to less restrictive monitoring. The studies undertaken and reported in this thesis provide empirical evidence of poor data recording management procedures that flawed the results. Lack of transparency in procedures meant that this was not detected until a second study was undertaken, and data from the first study had already been presented. (Urdahl et al 2000)

Strategies adopted in the statistical analysis of the observational study of this thesis were primarily generalised linear modelling. However, strategies such as instrumental variables have made valuable contribution to the analysis of observational data, a methodology that was developed responding to the fact that important aspects of health

status cannot be observed. (McLelland and Newhouse 1997) Such unobserved patient factors may include underlying severity of disease, patient preferences, compliance, and physician prescribing preferences. These all are potentially strong predictors of the treatment outcome. A possible instrumental variable in the Mediplus® database may have been GP identity, but instrumental variables analysis was not pursued here, principally because the Mediplus® study was exploratory and the issues to be explored ranged from appropriate design of the study to availability and validity of the data. In the light of the inconsistent findings, it is doubtful whether such sophisticated statistical techniques would have provided a more reliable result given the nature of and lack of sophistication of the data collected in the studies.

An additional observation that was made throughout the design, conduct and analysis of the Mediplus® studies was the fact that in spite of the fundamental difference between observational and randomised study designs the processes of undertaking the research have key features in common. Time and cost constraints and issues of data quality needing consideration in a trial are also important to the conduct of an observational study. The initial time investment to acquire the skills and insight to conduct an analysis in an observational database are substantial, and may result in a time-lag between the inception of the idea readiness for dissemination. The cost per patient is lower in an observational study of this kind than in a clinical trial, yet there are still substantial costs involved and in the case of Mediplus®, these are calculated on the basis of patient numbers included in the study.

Well-designed observational studies conducted in valid, reliable and accurate databases, and the concomitant use of statistical methodology developed over the last few years to control for both observed and unobserved bias in observational studies, may be valuable when assessing cost-effectiveness in clinical practice. Despite the rigid design and appropriate methodology, the database that was explored in this thesis failed to meet the data requirements for economic analysis, and problems with data management planted seeds of doubt about the scientific contents of files extracted.

In conclusion, the feasibility of the Mediplus® database in its current format was found to be infeasible for full economic evaluation. Both the lack of relevant variables and the invalid data extraction encountered, provide limitations for the database. The project did not

however reject the hypothesis that observational data in general may be used in economic evaluation. The second study indicated that the database may be of value in studying aspects of drug use such as likelihood estimates of compliance rates, cost of treatment for side-effects as well as a description of patient population in real life.

6.2 *Relative merits of alternative approaches; Conclusions*

Three methods were evaluated individually in the thesis in spite of not being mutually exclusive. In fact, all study design approaches evaluated in this thesis are less than perfect for the purpose of increasing generalisability of economic evaluations and may therefore be combined in one or more ways in order to generalise cost-effectiveness estimates across settings. Estimates of effect derived in a clinical trial are based on a series of assumptions incorporated in the design of the trial and may not be reflective of clinical practice, (Coyle et al 1998) but estimates based on pragmatic trials can be 'contaminated by reality'. (Freemantle and Drummond 1997) Economic models may be 'black boxes' and difficult to quality assess, (Sculpher et al 2000, Hill et al 2000) but evaluations based on epidemiological studies can be flawed by poor data management and invalid estimates (Chapter 5). *Table 6.1* outlines the relative merits of the methods assessed in this thesis to generalise economic evaluation results.

The first source project evaluating pragmatic trials provided evidence that trial features do not, on average, provide systematically different effect estimates from those adopting explanatory features. Pragmatic trials complement late phase III drug development, where aspects of generalisability in terms of cost-effectiveness from clinical trial to clinical practice and from country to country could be assessed. Their use is limited in generalising economic evaluation analyses as they are conducted in one particular setting, and so could be generalised *only* to this setting. For example, pragmatic trials frequently use 'current practice' as comparator to the therapy under evaluation, but 'current practice' may vary between healthcare systems and settings. (O'Brien 1997) The extent to which this is a problem will depend on whether the data is to be used by a specific decision-maker faced by context-specific set of financial and practical constraints, or whether it is to inform a broader debate surrounding resource allocation. (Murray et al 2000)

Table 6.1 *Relative merits of different methodological approaches to generalising economic evaluations*

	Pragmatic trials	Economic modelling	Observational study (patient records)
Advantages	<p>Random allocation ensures good basis for group comparison</p> <p>Data reflect clinical practice without, on average, jeopardising internal validity</p>	<p>Sensitivity analysis enables testing of generalisability using context-specific parameters</p> <p>Enables combination of randomised and non-randomised data</p>	<p>Data reflect clinical practice</p> <p>Large patient numbers increases power and may enable study of final outcomes</p> <p>Time and financial constraints may be overcome</p>
Dis-advantages	<p>Results are context-specific</p> <p>Choice of comparator or unit costs may not translate beyond the setting of the trial</p>	<p>Dependent on availability of context-specific data for parameter estimates</p> <p>Presentation of results complicated and space-consuming</p>	<p>Non-randomised approach jeopardises study validity</p> <p>Inability to demonstrate causal relationships</p> <p>Constraints in real-life limits quality of recording</p> <p>Activity-driven records limits availability of factors relevant to the analysis</p>
Potential uses	<p>Clinical trial to context-specific clinical practice</p> <p>Phase III of drug development</p>	<p>Early phase drug development to explore impact of real-life scenarios on cost-effectiveness ratio</p> <p>Late phase cross-national extrapolations</p> <p>The synthesis of results from trial data with observational data on resource use and baseline risks</p>	<p>Clinical trial to context-specific clinical practice</p> <p>Input to economic models</p>

The choice of an appropriate method or combination of methods to generalise the results of an economic evaluation depends critically on the use of the results. For an evaluation targeted at a specific decision-making body facing a known budget and a given set of options, the pragmatic trials and observational study approaches offer advantages over the pivotal trial-based, and even the model-based, evaluation because

of their context-specific properties. Based on such evaluations, the decision-maker faced with limiting factors such as human resources and healthcare provision infrastructure, may decide to reallocate resources or indeed allocate a budget increase. However, for the provision of general information on the relative costs and benefits of selected healthcare interventions, models may be more appropriate, as they enable testing of relative cost-effectiveness under different set of assumptions. Indeed, relative estimate of clinical effectiveness from a pragmatic trial may be combined with baseline risks and resource consumption data derived from an observational study within the framework of a computerised model.

Patient-level analyses such as pragmatic trials and observational studies do not provide scope for an overall sectoral perspective in which costs and effectiveness of all possible interventions could be compared, enabling identification of the mix that maximises healthcare under given resource constraints. By default, patient-level studies are context-specific, even if that context is more reflective of clinical practice patterns than clinical trial setting through adoption of observational or pragmatic design. An implicit assumption of analyses presenting cost-effectiveness of a single proposed new intervention as compared with an alternative intervention, or with a fixed cut-off point representing an assumed societal 'willingness to pay' for an additional unit of benefit is that, to improve overall efficiency, resources would need to be transferred to the more efficient intervention or from another sector. Broadly, cost-effectiveness analyses based on models may be seen as a tool to allocate a fixed health budget in order to maximise a society's health (Garber et al 1996)

The second source project indicated that models are feasible vehicles for generalising trial to practice estimates, for example through exploring different compliance rates and country to country unit costs estimates. They may represent a method that, in theory, could be used for the broader remit of maximising resource allocation within different fixed healthcare budgets under different sets of circumstances. Economic models seeking to present results in a generalised fashion may risk being criticised of insensitivity and inaccuracy when trying to present cost-effectiveness across a span of healthcare contexts. Also, space-restriction in scientific journals and concerns voiced about the quality of reporting of economic models (Hill et al 2000) limits the degree to which models may make use of the opportunity to explore ranges of scenarios.

However sophisticated a model is, some issues are unlikely to be solved by more modeling, and some may only be solved by collecting new clinical data. Data input to economic models should reflect clinical practice and conventions in order to increase the merits of the model for providing generalisable results.

The various approaches evaluated in this research may be appropriate methodologies at different phases of product development. Modelling may be appropriate in an early phase of clinical development, to moderate explanatory trial features to reflect clinical practice. Pragmatic trials may be appropriate for an intermediate phase. Valid observational studies may be relevant in a late phase where the intervention in question has been in routine use for some time. Then a return to modelling at a more advanced phase may facilitate transferability of cost-effectiveness results between countries. Importantly, models may at any stage in the drug lifecycle provide a valuable tool to combine trial data with observational studies or population-based data.

There is little value in searching for a universal methodology to increase generalisability of economic evaluation. The research team would need to decide which approach to apply in a given circumstance, as it would depend on the nature of the intervention, drug development phase, target patient population and the decision to be informed. The methods reviewed in this thesis may all be feasible approaches across different stages in the development of a drug and across different settings either individually or in combination. Experimentation and observation can be complementary tools to improve external validity of economic evaluations.

The emphasis here is that, as yet, there is no perfect approach to generalisability. Only through insight into the strengths and weaknesses of each approach, can researchers establish the one that satisfies the information need. The methods reviewed in this thesis are not mutually exclusive, but complementary tools for researchers and decision-makers wishing to throw light on these issues. When used by themselves or in combination they may increase understanding of how technologies change across settings and provide deeper insight into cost-effectiveness for clinical decision-makers in particular healthcare environments.

6.3 *Recommendations*

There is a prevailing need for methodologies to evaluate generalisability of economic evaluation and empirical research to assess how economic and clinical data travel across settings. Research presented in this thesis has added to the current knowledge base of methods to generalise economic evaluation results and provided empirical data demonstrating relative merits of three alternative methods. The findings may have implications for research and policy agenda:

1. The examination of internal versus external validity of pragmatic trials did not evaluate resource aspects of cost-effectiveness analyses alongside pragmatic trials. The project concluded that effect estimates from pragmatic trials do not jeopardise internal validity at the expense of externally valid estimates. However, a key component of economic evaluation is the identification of unit costs. Further research should explore the resource-side of this and provide insight into transferability of context-specific resource information.
2. Empirical data demonstrating the use of modelling to enhance the generalisability of context-specific evaluations is scarce. The way that models can be used to explore cost-effectiveness under different scenarios should be applied to increase understanding of cost-effectiveness ratios across settings.
3. Those undertaking economic modelling should, to a larger extent, explore alternative assumptions relating to context. Trial-based probabilities should be subject to sensitivity analysis reflecting assumptions about practice patterns in clinical practice. Similarly, those that commission studies must be prepared to explore a range of scenarios relevant to the transfer of a therapy between settings. They must be prepared to justify the values chosen with respect to generalisability.
4. The feasibility study of the Mediplus® database revealed limitations relating to data quality and data management. Firstly, formal validation studies should be undertaken to assess whether current recording is reliable. Secondly, quality control mechanisms should be introduced to ensure validity of recording in the Mediplus® database. Finally, scientific method should be adopted in future extractions of data

from Mediplus® for similar studies, for example through dual data extraction and introduction of procedures to document the data extraction steps.

5. Further research should seek to evaluate the influence of context specific factors on cost-effectiveness estimates, and assess which characteristics an economic evaluation should encompass, in terms of overall design, data collection, analysis and presentation, in order to optimise the generalisability of the results.
6. The research presented in this thesis has focused on methods that can be adopted to increase and explore generalisability across settings and locations. Limitations to generalisability can occur over time, for example through changes in demography, treatment patterns and available interventions, illustrating the need for empirical research to explore how cost-effectiveness profiles of interventions change over time. This has specific implications in those jurisdictions where economic evaluation systematically contributes to decision-making such as in Australia, Canada and the UK, where reimbursement decisions may be reassessed.

REFERENCES

Altman DG. Practical statistics for medical research. Chapman and Hall, London 1997.

Altman DG, Bland JM. Generalisation and extrapolation. *British Medical Journal* 1998;317:409-410

Altman DG, Schulz KF. Concealing treatment allocation in randomised trials. *British Medical Journal* 2001;323:446-447.

Beardon PHG, McGilchrist MM, McKendrick AD, McDevitt DG, MacDonald TM. Primary non-compliance with prescribed medication in primary care. *British Medical Journal* 1993;307:846-848.

Begg CB, Cho MK, Eastwood S, Horton R, Moher D, Olkin I et al. Improving the quality of reporting randomised controlled trials: the CONSORT statement. *JAMA* 1996;276:637-9

Benson K, Harz A. A comparison of observational studies and randomised controlled trials. *The New England Journal of Medicine* 2000;342:1878-1886

Black N. Why we need observational studies to evaluate the effectiveness of health care. *British Medical Journal*. 1996;312:1215-1218.

Black DM, Cummings SR, Karpf DB, Cauley JA et al. Randomised trial of effect of alendronate and risk of fracture in women with existing vertebral fractures. *Lancet* 1996;348:1535-1541

Bloor K, Freemantle N. Lessons from international experience in controlling pharmaceutical expenditure II: influencing doctors. *British Medical Journal* 1996;312:1525-27

Bloor K, Maynard A, Freemantle N. Lessons from international experience in controlling pharmaceutical expenditure II: influencing doctors. *British Medical Journal* 1996;313:33-35

British National Formula. British Medical Association and Royal Pharmaceutical Association. The Pharmaceutical Press, Oxon, England 2000

Bombardier C, Maetzel A. Pharmacoeconomic evaluation of new treatments: efficacy versus effectiveness studies? *Ann Rheumatic Dis* 1999;58:182-185

Brennan A, Akehurst R. Modelling in Health Economic Evaluation: What is its place? What is its value? *Pharmacoeconomics* 2000;17(5)445-459.

Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000;17(5):479-500

Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technol Assess* 1999;3(2)

Briggs A, Gray A. Using cost-effectiveness information. *British Medical Journal* 2000;320:246

Briggs A, Schulpher MJ. An introduction to Markov modelling for economic evaluation. *Pharmacoeconomics* 1998;13:397-409.

Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C. Choosing between randomised and non-randomised studies: a systematic review. *Health Technology Assessment* 1998;2(13).

Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, Vray M. Modelling in economic evaluation: an unavoidable fact of life. *Health Economics* 1997;6:217-227

Byford S, Palmer S. Common errors and controversies in pharmacoeconomic analyses. *Pharmacoeconomics* 1998;13(6):659-666

Canadian Co-ordinating Office for Health Technology Assessment. Guidelines for Economic Evaluation of Pharmaceuticals: Canada. 2nd edition 1997.

Cappelleri JC, Ioannidis JPA, Schmid CH, de Ferranti SD, Aubert M, Chalmers TC et al. Large trials vs meta-analysis of smaller trials. How do their results compare? *JAMA* 1996; 276: 1332-8.

Claxton K, Posnett J. AN economic approach to clinical trial design and research priority setting. *Health Economics* 1996;5:513-524

Chalmers TC, Smith H, Blackburn B, Silverman B, Schroeder B, Reitman D, Ambroz A. A method for assessing the quality of a randomised controlled trial. *Controlled Clinical Trials* 1981;2:31-49

Chisholm J. The Read clinical classification. *British Medical Journal* 1990;300:1092

Cochrane AL. Effectiveness and efficiency. Random reflections on health services. Cambridge University Press. 1971.

Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I:Medical. *Statistics in Medicine*. 1989;8:441-454.

Collett D. Modelling survival data in medical research. Chapman and Hall, London 1994.

Commonwealth Department of Human Services and Health. Guidelines for the Pharmaceutical Industry on Preparation of submissions to the pharmaceutical Benefits Advisory Committee Including Major Submissions Involving Economic Analyses. Canberra: Australian Government Publishing Service; 1995. Updated 2001: <http://www.health.gov.au/pbs/pubs/pharmpac/gusubpac.htm>

Concato J, Shah N, Horwitz RI. Randomised controlled trials, observational studies and the hierarchy of research designs. *The New England Journal of Medicine* 2000;342:1887-1892

Cooper C. Epidemiology and definition of osteoporosis. In: Compston JE, ed. *Osteoporosis: New perspectives on causes, prevention and treatment*. London: Royal college of Physicians of London: 1-10.

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(1):135-144.

CRD Report number 4 (2nd edition) Undertaking systematic reviews of research on effectiveness. CRD's guidance for those carrying out or commissioning reviews. March 2001.

Cummings SR, Black DM, Thompson DE, Applegate WB, Barrett-Connor E, Musliner TA, Palermo L, Prineas R, Rubin S, et al. for the Fracture Intervention Trial Research Group. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures. *JAMA* 1998;280:2077-2082

Davies HTO, Crombie IK. Outcomes from observational studies: understanding causal ambiguity. *Drug Information Journal* 1999;33:153-158

Day SJ, Altman DG. Blinding in clinical trials and other studies. *British Medical Journal* 2000;321:504

Delva D. Hormone replacement therapy. Risks, benefits and costs. *Canadian Family Physician* 1993;39:2149-2154.

Dennison E, Cooper C. Epidemiology of osteoporotic fractures. In Barlow DH, Francis RM, Miles A. (eds) *The effective management of Osteoporosis*. Aesculapius Medical Press, London 2001

Dolan P, Torgerson DJ. The cost of treating osteoporotic fractures in the United Kingdom female population. *Osteoporosis International* 1998;8:611-617

Donaldson C, Hundley V, McIntosh E. Using economics alongside clinical trials: why we cannot choose the evaluation technique in advance. *Health Economics* 1996;5:267-269

Drummond MF. Economic evaluation of pharmaceuticals. Science or marketing? *Pharmacoeconomics* 1992^a;1(1):8-13

Drummond M. Cost-effectiveness guidelines for reimbursement of pharmaceuticals: is economic evaluation ready for its enhanced status? *Health economics* 1992^b;1:85-92

Drummond MF. Experimental versus observational data in economic evaluation of pharmaceuticals. *Medical decision making*. 1998;18(2):S12-S18

Drummond M, Cooke J, Walley T. Economic evaluation under managed care: evidence from the UK. *Social Science and Medicine* 1997;45(4):538-595.

Drummond MF, Davies L. Economic analysis alongside clinical trials. Revisiting the methodological issues. *International Journal of Technology Assessment in Health Care*. 1991;7(4):561-573

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

Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *British Medical Journal* 1996;313:275-283

Drummond MF, Stoddard GL. Economic analysis and clinical trials. *Controlled Clinical Trials* 1984;5:115-128.

Drummond M, Torrance G, Mason J. Cost-effectiveness league tables: more harm than good? *Social science and medicine* 1993;37(1):33-40.

Eccles M, Freemantle N, Mason J. North of England evidence based guideline development project: guideline on the use of aspirin as secondary prophylaxis for vascular disease in primary care. North of England Aspirin Guideline Development Group. *BMJ* 1998 Apr 25;316(7140):1303-9

Eddy DM. Technology assessment: the role of mathematical modelling. In: Mosteller F, editor. *Assessing medical technologies*. Washington DC. National Academy Press 1985:144-53

Effron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1986;1(1):54-77.

Egger M, Smith D. Meta analysis: potential and promise. *British Medical Journal* 1997;315:1371-1374

Ettinger B, Pressman A, Silver P. Effect of age on reasons for initiation and discontinuation of hormone replacement therapy. *Menopause* 1999;6:282-289.

^a Ettinger B, Pressman A, Schein J, Chan J, Silver J Connolly N. Alendronate use among 812 women: prevalence of gastrointestinal complaints, noncompliance with patient instructions and discontinuation. *Academy of Managed Care Pharmacy* 1998;4:488-492

^b Ettinger B, Pressman A, Schein J. Clinical visits and hospital admissions for care of acid-related upper gastrointestinal disorders in women using alendronate for osteoporosis. *American Journal of Managed Care* 1998;4:1377-1382

Fayers PM, Hand DJ. Generalisation from phase III clinical trials: survival, quality of life and health economics. *The Lancet* 1997;350:1025-1027

Farmer RDT, Williams TJ, Simpson EL, Nightingale AL. Effect of 1995 pill scare among women taking combined oral contraceptives: analysis of General Practice Research Database. *British Medical Journal* 2000;321:477-9

Farmer RDT, Lawrenson RA. Oral contraceptives and venous thromboembolic disease: The findings from database studies in the United Kingdom and Germany. *American Journal of Obstetrics and Gynaecology* 1998;179(no 3 part 2):S78-S86.

Farmer RDT, Lawrenson RA, Thompson CR, Kennedy JG, Hambleton IR. Population-based study of risk of venous thromboembolism associated with various oral contraceptives. *Lancet* 1997;349:83-88

Freemantle N. Does the UK National Health Service need a fourth hurdle for pharmaceutical reimbursement to encourage the more efficient prescribing of pharmaceuticals? *Health Policy* 1999;46:255-265

Freemantle, N. Interpreting the results of secondary endpoints and subgroup analyses in clinical trials: should we lock the crazy aunt in the attic? *British Medical Journal*, 2001; 322: 989-91.

Freemantle N, Bloor K. Lessons from international experience in controlling pharmaceutical expenditure. I: Influencing patients. *British Medical Journal* 1996;312:1469-71

Freemantle N, Urdahl H, Eastaugh J, Hobbs FDR. What is the place of beta blockade in patients who have experienced a myocardial infarction with preserved left ventricular function? Evidence and (mis)interpretation. *Prog Cardiovascular Disease* in press

French F, Torgerson DJ, Porter RW. Overestimation of cost savings in prevention of osteoporosis (Letter) *British Medical Journal* 1994;308:138

Garber AM, Weinstein MC, Torrance GW, Kamlet MS. Theoretical foundations of cost-effectiveness analysis. In Gold MR, Siegel JE, Russell LB et al. *Cost-Effectiveness Analysis in Health and Medicine*. New York: Oxford University Press, 1996.

Gardner MJ, Altman DG. *Statistics with confidence. Confidence intervals and statistical guidelines*. The British Medical Journal. The University Press (Belfast) Ltd. 1989

Geddes, J, Freemantle N, Harrison P, Bebbington P. Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. *British Medical Journal* 2000 Dec 2;321(7273):1371-6

Glick H, Willke R, Polsky D, Llana T, Alves WM, Kassell N, Schulman K. Economic analysis of trilazad mesylate for aneurysmal subarachnoid hemorrhage. *International Journal of Technology assessment in Health Care* 1998;14(1):145-160.

Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness analysis in health and medicine. New York (NY): Oxford University Press, 1996.

Gray AM, 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

Hailey D, Sampietro-Colom L, Marshall D, Rico R, Granados A, Asua J. The effectiveness of bone density measurement and associated treatments for prevention of fractures. An international collaborative review. *International Journal of Technology Assessment in Health Care*. 1998;14(2):237-254

Hennekens CH, Buring JE. *Epidemiology in medicine*. Little, Brown and Company. Boston/Toronto 1987

Hill SR, Mitchell AS, Henry DA. Problems With the Interpretation of Pharmacoeconomic Analyses: A Review of Submissions to the Australian Pharmaceutical Benefits Scheme. *JAMA* 2000;283: 2116-

Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *British Medical Journal* 1999;319:670-4.

Hornberger J, Wrono E. When to base clinical policies on observational versus randomised trial data. *Annals of Internal Medicine* 1997;127(8): 697-703

HTA project 98/22/05. Sculpher M. An investigation of methods to increase the generalisability and transferability of economic evaluation studies.
<http://hta.nhsweb.nhs.uk/>

Hughes DA, Bagust A, Haycox A, Walley T. The impact of non-compliance in the cost-effectiveness of pharmaceuticals: a review of the literature. *Health Economics* 2001;10(7):601-615.

Hurley J. An overview of the normative economics of the health care sector. In *Handbook of Health Economics volume 1A*. Eds. Culyer AJ, Newhouse JP. Elsevier Science BV. 2000

Hutton J, Maynard A. A nice challenge for health economics. *Health Economics* 2000;9:89-93

International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use. ICH Harmonised Tripartite Guideline: Guideline for Good Clinical Practice 1996

Ioannidis JPA, Cappelleri JC, Lau J. Issues in comparisons between meta-analyses and large trials. *JAMA* 1998; 279: 1089-93.

IMS Health. UK Primary Care Database (Mediplus®). Quality, Validity, Representativeness and Completeness of data. IMS Health 1998.

Jick H, Rodriguez LAG, Perez-Gutthann S. Principles of epidemiological research on adverse and beneficial drug effects. *The Lancet* 1998;352:1767-1770.

Jick H, Jick S, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. *British Medical Journal* 1991;302:766-768.

Jick H, Kaye JA, Vasilakis-Scaramozza C, Jick S. Risk of venous thromboembolism among users of third generation oral contraceptives compared with users of oral contraceptives with levonorgestrel before and after 1995: cohort and case-control analysis. *British Medical Journal* 2000;321:1190-5

Jick H, Terris B, Derby LE, Jick S. Further validation of information recorded on a general practitioner based computerized data resource in the United Kingdom. *Pharmacoepidemiology and drug safety* 1992;1:347-349.

Johnston K, Buxton MJ, Jones DR, Fitzpatrick R. Assessing the costs of healthcare technologies in clinical trials. *Health Technol Assess* 1999;3(6)

Jones JK, Gorkin J, Lian JF, Staffa JA, Fletcher AP. Discontinuation of, and changes in treatment after start of new courses of antihypertensive drugs: a study of United Kingdom population. *British Medical Journal* 1995;311:293-295

Jonsson B, Christiansen C, Johnell O, Hedbrandt J. The cost effectiveness of fracture prevention in established osteoporosis. *Scandinavian Journal of Rheumatology* 1996;25(suppl 103):30-38

Juni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *British Medical Journal* 2001;323:42-46.

Kunz R, Oxman A. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *British Medical Journal* 1998;317:1185-1190.

Lancet. From research to practice. Editorial *The Lancet* 1994;344:417-418

Lewsey JD, Leyland AH, Murray GD, Boddy FA. Using routine data to complement and enhance the results of randomised controlled trials. *Health Technol Assess* 2000;4(22)

Lis Y, Carter M, Lumley C, Walker S. Can the UK Automated databases provide relevant information for health economics research? Centre for Medicines Research International report (CMR97-97R) 1997.

MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AMS. A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies. *Health Technol Assess* 2000;4(34).

MacDonald TM, Collins D, McGilchrist MM, Stevent J, McKendrick AD, McDevitt DG, Davey PG. The utilisation and economic evaluation of antibiotics prescribed in primary care. *Journal of Antimicrobial Chemotherapy* 1995;35:191-204

- Mauskopf J, Schulman K, Bell I, Glick H. A strategy for collecting pharmaco-economic data during phase II/III clinical trials. *Pharmacoeconomics*, 1996;9:264-277
- Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost-effectiveness analysis. In: Gold MR, Siegel JE, Russell LB et al. editors. *Cost effectiveness in Health and Medicine*. New York (NY): Oxford University Press, 1996:247-75
- Mason J. The generalisability of pharmaco-economic studies. *Pharmacoeconomics*. 1997;11(6):503-514
- Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *British Medical Journal* 1996;312:1254-1259
- Maynard A. Developing the health care market. *The Economic Journal* 1991;101:1277-1286
- Maynard A, Bloor K. Regulating the pharmaceutical industry. *British Medical Journal* 1997;315:200-201
- Maynard A, Bloor K. *Our certain fate: rationing in health care*. Office of Health Economics, London 1998.
- McCabe C, Dixon S. Testing the validity of cost-effectiveness models. *Pharmacoeconomics* 2000;17(5):501-513
- McClelland M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA* 1994;272:859-866
- McCullagh P, Nelder JA. *Generalized linear models*. New York: Chapman & Hall, 1989
- McGuire A, Henderson J, Mooney G. *The economics of health care*. Routledge 1988
- Moher D, Ba'Pham, Jones A, Cook D, Jadad AR, Moher M, Tugwell P, Klassen PT. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *The Lancet* 1998; **352**:609-13.
- Moher M, Schulz KF, Altman DG, for the CONSORT group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomised trials. *JAMA* 2001;285:1987-91. Simultaneously in *Ann Intern Med* 2001;134:657-62 and *Lancet* 2001;357:1191-4
- Moher D, Jadad AR, Tugwell P. Assessing the quality of randomised controlled trials. Current issues and future directions. *International Journal of Technology Assessment in Health Care*. 1996;12(2):195-208

Mosteller F. Assessing medical technologies. National Academy Press, Washington 1985.

Murray CJL, Evans DB, Acharya A, Baltussen R. Development of WHO guidelines on generalised cost-effectiveness analysis. *Health Economics* 2000;9:235-251

National Institute for Clinical Excellence. Technical Guidance for Manufacturers and Sponsors on making a Submission to a Technology Appraisal. London 2001.

Nuijten MJC. The selection of data sources for use in modelling studies. *Pharmacoeconomics* 1998;13(3):305-316.

O'Brien BJ. A tale of two (or more) cities: geographic transferability of pharmaco-economic data. *The American Journal of Managed Care* 1997;3(Suppl):S33-S40.

O'Brien BJ. Economic Evaluation of Pharmaceuticals: Frankenstein's monster or vampire of trials? *Medical Care* 1996;34(12):DS99-108

O'Hagan A, Stevens JW. A framework for cost-effectiveness analysis from clinical trial data. *Health Economics* 2001;10:303-315

Oster G, Borok GM, Menzin J, Heyse JF, Epstein RS, Quinn V, Benson V, Dudl J, Epstein A. A randomised trial to assess effectiveness and cost in clinical practice: rationale and design of the Cholesterol Reduction Intervention Study (CRIS). *Controlled Clinical Trials* 1995;16:3-16

Pablos-Mendez A, Barr G, Shea S. Run-in periods in randomised trials. Implications for the application of results in clinical practice. *JAMA* 1998;279:222-225

Pang F, Drummond MF, Song F. The use of meta-analysis in economic evaluation. Discussion Paper 173, Centre for Health Economics, University of York 1999.

Pocock SJ. *Clinical Trials. A practical approach.* John Wiley & Sons, New York 1983

Pocock SJ, Elbourne DR. Randomised trials or observational tribulations? *The New England Journal of Medicine.* 2000;342:1907-1909

Piantadosi S. *Clinical Trials. A methodological perspective.* John Wiley and Sons, New York 1997

Radford MJ, Foody JM. How do observational studies expand the evidence base for therapy? *JAMA* 2001;286:1228-1230

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;9:191-198

Revicki DA, Frank L. Pharmacoeconomic evaluation in the real world. Effectiveness versus efficacy studies. *Pharmacoeconomics* 1999;15(5):423-434

Rodriguez LAG, Perez-Gutthann S. Use of the UK General Practice Research Database for pharmacoepidemiology. *British Journal of Clinical Pharmacology* 1998;45:419-425.

Rothwell PM. Can overall results of clinical trials be applied to all patients? *The Lancet* 1995;345:1616-19

Royal College of Physicians of London. Osteoporosis. Clinical guidelines for prevention and treatment. 1999.

Roland M, Torgerson DJ. What are pragmatic trials? *British Medical Journal* 1998;316:285

Rothwell PM. Can overall results of clinical trials be applied to all patients? *The Lancet* 1995;345:1616-1619

Russell LB, Siegel JE, Daniels N, Gold MR, Luce BR, Mandelblatt JS. Cost-effectiveness analysis as a guide to resource allocation in health: Roles and limitations. In Gold MR, Siegel JE, Russell LB et al. *Cost-Effectiveness Analysis in Health and Medicine*. New York: Oxford University Press, 1996.

Sackett DL, Haynes RB, Tugwell P. *Clinical epidemiology: A basic science for clinical medicine*. Toronto: Little, Brown and Co., 1985

Sacristan JA, Soto J. Drug utilisation studies as tools in health economics. *Pharmacoeconomics* 1994;5(4):299-312.

Schulman KA, Buxton M, Glick H, Schulpher M, Guzman G, Kong J, Backhouse M, Mauskopf J, Bell L, Eisenberg J. Results of the economic evaluation of the FIRST study. A multinational Prospective Economic Evaluation. *International Journal of Technology assessment in Health Care*. 1996;12(4):698-713.

Schulpher M, Torgerson D, Goeree R, O'Brien B. A critical structured review of economic evaluations of interventions for the prevention and treatment of osteoporosis. Discussion paper 169. Centre for Health Economics, University of York 1999.

Schulpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models. A suggested framework and example of application. *Pharmacoeconomics* 2000;17(5):461-477.

Schulz K, Chalmers I, Hayes R, Altman D. Empirical evidence of bias. (1996) Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 273:408-412.

Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Chronic Diseases* 1967;20:637-648.

Sheldon TA. Problems of using modelling in the economic evaluation of health care. *Health Economics* 1996;5:1-11

Simon G, VonKorff M, Heiligenstein JH, Revicki D, Grothaus L, Katon W, Wagner EH. Initial antidepressant choice in primary care. Effectiveness and cost of fluoxetine vs. tricyclic antidepressants. *JAMA* 1996;275(24):1897-1902

Simon G, Wagner E, Vonkorff M. Cost-effectiveness comparisons using "real world" randomized trials: the case of new antidepressant drugs. *Journal of Clinical Epidemiology* 1995;48(3):363-373

Skegg DCG. Pitfalls of pharmacoepidemiology. Oral contraceptive studies show a need for caution with databases. *British Medical Journal* 2000;321:1171-2

Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *British Medical Journal* 2000;320:1197-1200

Torgerson DJ, Bell-Syer SEM. Hormone replacement therapy and prevention of non-vertebral fractures. A meta-analysis of randomised trials. *JAMA* 2001;285:2891-2897

Torgerson DJ, Reid DM. The pharmacoeconomics of hormone replacement therapy. *Pharmacoeconomics* 1999;16(1):9-16.

Torgerson DJ, Reid DM. The economics of Osteoporosis and its prevention. A review. *Pharmacoeconomics* 1997;11(2):126-138

Tosteson ANA, Weinstein MC. Cost-effectiveness of hormone replacement therapy after the menopause. *Baillieres clinical obstetrics and gynaecology* 1991;5:943-959

Torrance GW, Siegel JE, Luce BR. Framing and designing the cost effectiveness analysis. In Gold MR, Siegel JE, Russell LB et al. *Cost-Effectiveness Analysis in Health and Medicine*. New York: Oxford University Press, 1996.

Udvarhelyi S, Colditz G, Rai A, Epstein A. Cost-effectiveness and cost-benefit analyses in the medical literature. Are the methods being used correctly? *Annals of Internal Medicine* 1992;116:238-244

Urdahl HF, Young P, McKendrick J, Freemantle N. Use of bisphosphonates and the cost of treating gastrointestinal side effects in a UK clinical practice setting. *Osteoporosis International* 2000;(11)S5:S10-S11.

Van Staa T, Abenhaim L, Cooper C. Upper gastrointestinal adverse events and cyclical etidronate. *The American Journal of Medicine* 1997;103:462-467

Van Staa T-P, Abenhaim L, Cooper C. Use of cyclical etidronate and prevention of non-vertebral fractures. *British Journal of Rheumatology* 1998;37:87-94

Van Staa T-P and Abenhaim L. The quality of information recorded on a UK Database of primary care records: a study of hospitalisations due to hypoglycemia and other conditions. *Pharmacoepidemiology and drug safety* 1994;3:15-21.

Weinstein MC, Fineberg HV. Clinical decision analysis. Philadelphia: WB Saunders Co., 1980.

Weinstein MC. Estrogen use in postmenopausal women – costs, risks and benefits. *New England Journal of Medicine* 1980;303:308-16

Wenger NK. Exclusion of the elderly and women from coronary trials. Is their quality of care compromised? *JAMA* 1992;268:1461-2

Williams A. Cost-effectiveness analysis: is it ethical? *Journal of Medical Ethics*. 1992;18:7-11.

Williams A. Cost-benefit analysis: Applied welfare economics or general decision aid? In Williams A, Giardina E. Efficiency in the public sector. The theory and practice of cost-benefit analysis. University Press, Cambridge, Great Britain 1993

Willke RJ, Glick H, Polsky D et al. Estimating country-specific cost-effectiveness from multinational clinical trials. *Health Economics* 1998;7:481-493.

Zaethraeus N, Johannesson M, Jönsson B. A computer model to analyse the cost-effectiveness of hormone replacement therapy. *International Journal of Technology Assessment in Health Care* 1998;15(2):352-365

APPENDIX 1

TRIALS INCLUDED IN THE ANALYSIS OF PRAGMATIC FEATURES

TRIALS IN POST-MI

4S Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian simvastatin survival study (4S). *Lancet* 1994; 344: 1383-9.

Ahlmark G, Saetre H, Korsgren M. Reduction of sudden deaths after MI. *Lancet* 1974; i: 1563

Ahlner J, Bergdahl B, Dahlström U, Ohlsson J. Once daily dosing of enalapril in congestive heart failure. *Acta Med Scand* 1988; 223: 313-20

Ambrosioni E, Borghi C, Magnani B for the Survival of MI Long Term Evaluation (SMILE) Investigators. The effect of angiotensin-converting-enzyme inhibitor zofenopril on mortality and morbidity after anterior MI. *N Engl J Med* 1995;332:80-5.

AMIS Aspirin MI Study (AMIS) Research Group. AMIS: the aspirin MI study: final results. *Circulation* 1980;62(Suppl V)79-84.

Andersen MP, Bechsgaard P, Frederiksen J, Hansen DA, Jurgensen HJ, Nielsen B, et al. Effect of alprenolol on mortality among patients with definite or suspected acute MI. Preliminary results. *Lancet* 1979;ii:865-7.

Anderson JL, Lutz JR, Gilbert EM et al. A randomised trial of low-dose beta-blockade therapy for idiopathic dilated cardiomyopathy. *Am J Cardiol* 1985; 55: 471-5

Anon. Rationale, design, and organization of the Metoprolol CR/XL randomised intervention trial in heart failure (MERIT-HF). The international steering committee. *American Journal of Cardiology* 1997; 80: 54J –58J.

APSI] Boissel JP, Leizorovicz A, Picolet H, Ducruet T, and the APSI Investigators. Efficacy of acebutolol after acute MI (the APSI trial). *Am J Cardiol* 1990;66:24c-31c.

ARIS] Anturane Reinfarction Italian Study (ARIS) Research Group. Sulphinpyrazone in post-MI. *Lancet* 1982;i:237-42.

Aronow WS, Ahn C, Kronzon AI. Effect of Propanolol versus no propanolol on total mortality plus nonfatal MI in older patients with prior MI, congestive heart failure and left ventricular ejection fraction \leq 40% treated with diuretics plus angiotensin converting enzyme inhibitors. *American Journal of Cardiology* 1997; 80: 207-9.

Australian and Swedish Pindolol Study Group. The effect of pindolol on the two years mortality after complicated MI. *European Heart Journal* 1983; 4: 367-75

Aronow WS, Ahn C, Kronzon I. Effect of propranolol versus no propranolol on total mortality plus nonfatal MI in older patients with prior MI, congestive heart failure, and left ventricular ejection fraction $\geq 40\%$ treated with diuretics plus angiotensin converting enzyme inhibitors. *Am J Cardiol* 1997; 80: 207-9

Aronow WS, Kronzon AI. Effect of propranolol on mortality in patients with prior MI, congestive heart failure and left ventricular ejection fraction $\leq 40\%$. *JACC* February 1997 464A-465A.

ART] Anturane Reinfarction Trial (ART) Research Group. The Anturane reinfarction trial: re-evaluation of outcome. *N Engl J Med* 1982;306:1005-8.

Australia - New Zealand Heart Failure Research Collaborative Group. Effects of carvedilol, a vasodilator- β -blocker in patients with congestive heart failure due to ischemic heart disease. *Circulation* 1995; 92: 212-8.

Australia -New Zealand Heart Failure Research Collaborative Group. Effects of carvedilol in patients with congestive heart failure due to ischaemic heart disease: final results from the Australia-New Zealand Heart Failure Research Collaborative Group Trial *Lancet* 1997; 349: 375-80

Azancot I, Lorente P, Georgiopoulos G, Beaufils P, Masquet C, Baudouy Y et al. Effects of acebutolol on myocardial infarct extension: a randomized electrocardiographic enzymatic and angiographic study. *Circulation* 1982; 66: 986-94.

Baber NS, Wainwright Evans D, Howitt G, Thomas M, Wilson C, Lewis JA, et al. Multicentre post-infarction trial of propranolol in 49 hospitals in the United Kingdom, Italy and Yugoslavia. *Br Heart J* 1980;44:96-100.

Baber NS, Wainwright Evans D, Howitt G, Thomas M, Wilson C, Lewis JA, et al. Multicentre post-infarction trial of propranolol in 49 hospitals in the United Kingdom, Italy and Yugoslavia. *Br Heart J* 1980;44:96-100.

Barabino A, Galbariggi G, Pizzorni C, Lotti G. Comparative effects of long-term therapy with captopril and ibopamine in chronic congestive heart failure in old patients. *Cardiology*. 1991;78:289-296.

Barber JM, Boyle McC, Chaturvedi NBC, Singh N, Walsh MJ. Propranolol in acute MI. *Acta Med Scand (Suppl)* 1976; 587: 213-9

Barber JM, Murphy FM, Merrett JD. Clinical trial of propranolol in acute MI. *Ulster Med J* 1967; 36: 127-30

Basu S, Senior R, Raval U, van der Does R, Bruckner T, Lahiri A. Beneficial effects of intravenous and oral carvedilol treatment in acute MI: A placebo controlled, randomized trial. *Circulation* 1997; 96: 183-91.

Baur LHB, Schipperhyen JJ, Wall EE van der; Velde EA van der; Schaliij MJ; Eck-Smit BLF van et al. Beneficial effect of enalapril on left ventricular stenosis after acute anterior wall infarction. *Eur Heart J* 1997; 18: 1313-21

Beijing] Ning W, Shongjie F, for the Beijing Collaborative Study Group. Secondary prevention of cardiac events following MI: effects of atenolol and enalapril. *Chinese Medical Journal* 1997; 110: 602-6

Bengtsson K. Rehabilitation after myocardial infarction. *Scand J Med* 1983; 15: 1-9.

Bethell HJN, Mullee MA. A controlled trial of community based coronary rehabilitation. *Br Heart J* 1990; 64: 370-5.

B-HAT] β -Blocker Heart Attach Trial Research Group. A randomized trial of propranolol in patients with acute MI. 1. Mortality Results. *JAMA* 1982;247:1707-14.

Branagan JP, Walsh K, Kelly P, Collins WC, McCafferty D, Walsh MJ. Effect of early treatment with nifedipine in suspected acute MI. *European Heart Journal* 1986; 7: 859-65.

Brilla CG, Krämer B, Risler T, Seipel L. Niedrigdosertes enalapril bei schwerer chronischer. *Z Kardiologie* 1991; 80 (Suppl 2): 44-9.

Bristow MR, O'Connell JB, Gilbert EM et al. For the Bucindolol Investigators. Dose-response of chronic beta blocker treatment in heart failure from either idiopathic dilated or ischaemic cardiomyopathy. *Circulation* 1994; 89: 1632-42

Brune S, Schmidt T, Tebbe U, Kreuzer H. Hemodynamic effects of nebivolol at rest and on exertion in patients with heart failure. *The Journal of Vascular Diseases* 1990; September: 696-701

Bussmann WD, Storger H, Hadler D, et al. Long-term treatment of severe chronic heart failure with captopril: a double-blind randomized, placebo-controlled long-term study. *J Cardiovasc Pharmacol.* 1987;9 (suppl 2): S50-S60

Campbell 1984 - Reported in: Yusuf S, Peto R, Lewis J, Collins R, Sleight P. β -blockade during and after MI. An overview of the randomized trials. *Progress in Cardiovascular Diseases* 1985;27:225-71

Cannavà G, Buttafarro A, Mangano A, Silipigni C, Cordiani A, Casella F et al. Efficacia del trattamento a lungo termine con enalapril in pazienti con scompenso cardiaco congestizio. *Cardiologica* 1990; 35: 485-7.

CAPRIE Steering Committee. A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). *Lancet* 1996; 348: 1329-39.

Cardiff 1] Elwood PC. Trial of acetylsalicylic acid in the secondary prevention of mortality from MI. *BMJ* 1981;282:481.

Cardiff 2] Elwood PC, Sweetnam PM. Aspirin and secondary mortality after MI. *Lancet* 1979;ii:1313-15.

CARE] Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford JD, Cole TG et al. The effect of pravastatin on coronary events after MI in patients with average cholesterol levels. *N Engl J Med* 1996; 335: 1001-9.

Carson P, Phillips R, Lloyd M, Tucker H, Neophytou M, Buch NJ et al. Exercise after myocardial infarction: a controlled trial. *J Royal College Physicians* 1982; 16: 147-51

CASSIS] Widimsky J, Kremer HJ, Jerie P, Uhlir O, for the CASIS investigators. Czech and Slovak spirapril intervention study (CASSIS): A randomized, placebo and active-controlled, double-blind multicentre trial in patients with congestive heart failure. *Eur J Clin Pharmacol* 1995;49:95-102.

CCAIT] Waters D, Higginson L, Gladstone P, Kimball B, Le May M, Boccuzzi SJ et al. Effects of monotherapy with an HMG-CoQ reductase inhibitor on the progression of coronary atherosclerosis as assessed by serial quantitative arteriography. The Canadian Coronary Atherosclerosis Intervention Trial. *Circulation* 1994; 89: 959-68

CCMG] The Cilazapril-Captopril Multicentre Group. Comparison of the effects of cilazapril and captopril versus placebo on exercise testing in chronic heart failure patients: A double blind, randomized multicentre trial. *Cardiology* 1995;86(Suppl 1):34-40.

CDMRG] Captopril-Digoxin Multicentre Research Group. Comparative effects of therapy with captopril and digoxin in patients with mild to moderate heart failure. *JAMA*. 1988;259:539-544.

CDP-A] Coronary Drug Project (CDP) Research Group. Aspirin in coronary heart disease. *J Chronic Dis* 1976;29:625-42.

Chinese Cardiac Study Collaborative Group. Oral captopril versus placebo among 13 634 patients with suspected acute MI: interim report from the Chinese cardiac Study (CCS-1). *Lancet* 1995;345:686-87.

CIBIS Investigators. A randomised trial of beta-blockade in heart failure. The Cardiac Insufficiency Bisoprolol Study. *Circulation* 1994; 90: 1765-73.

CIBIS-II investigators and committees. The cardiac insufficiency bisoprolol study II (CIBIS-II): a randomised trial. *Lancet* 1999; 353: 9-13

Circio A, Platania F, Mangiameli S, Putignano E. Multicenter, randomized placebo-controlled, double-blind study of the safety and efficacy of oral delapril in patients with congestive heart failure. *Am J Cardiol* 1995; 75: 18F to 24F

Clausen J, Felsby M, Jorgensen FS, Nielsen BL, Roin J, Strange B. Absence of prophylactic effect of propranolol in MI. *Lancet* 1966; ii: 920-4.

Cleland JGF, Dargie HJ, Ball SG, et al. Effects of enalapril in heart failure: a double blind study of effects on exercise performance, renal function, hormones and metabolic state. *Br Heart J* 1985;54:305-11.

Cleland JGF, Freemantle N, McGowan J, Clark A. The Evidence for Beta-Blockers equals that for ACE Inhibitors in Heart Failure. *British Medical Journal* 1999;318:824-825

CMRG] Captopril Multicenter Research Group: A placebo-controlled trial of captopril in refractory chronic congestive heart failure. *J Am Coll Cardiol* 1983;2(4):755-63.

Cohn JN, Ziesche S, Smith R, Anand I, Dunkman B, LoebH, et al. Effect of the calcium antagonist felodipine as supplementary vasodilator therapy in patients with chronic heart failure treated with enalapril. V-HeFT III. *Circulation* 1997; 96: 856-63.

Colfer HT, Ribner HS, Gradman A, Hughes V, Kapoor A, Laideau JC, for the Benazepril Heart Failure Study Group. Effects of once-daily benazepril therapy on exercise tolerance and manifestations of chronic congestive heart failure. *Am J Cardiol.* 1992;70:354-8.

Colucci WS, Packer M, Bristow MR et al for the US carvedilol heart failure study group. Carvedilol inhibits clinical progression in patients with mild symptoms of heart failure. *Circulation* 1996; 94: 2800-6

CONSENSUS Trial Study Group. Effects of enalapril on mortality in severe congestive heart failure. *N Engl J Med.* 1987;316:1429-35.

Cook TD, Campbell DT. Quasi-experimentation: design and analysis issues for field settings. Chicago: Rand McNally, 1979

CPRG] Coronary Prevention Research Group. An early intervention secondary prevention study with oxprenolol following MI. *European Heart Journal* 1981;2:389-93.

Crea F, Deanfield J, Crean P, Sharom M, Davies G, Maseri A. Effects of verapamil in preventing early postinfarction angina and reinfarction. *Am J Cardiol* 1985; 55: 900-4

Currie PJ, Kelly MJ, McKenzie A et al. Oral beta-adrenergic blockade with metoprolol in chronic severe dilated cardiomyopathy. *J Am Coll Cardiol* 1984; 3: 203-9

Curtis JL, Houghton JL, Patterson JH, Koch G, Bradley DA, Adams KF. Propranolol therapy alters estimation of potential cardiovascular risk derived from submaximal postinfarction exercise testing. *Am Heart J* 1991; 121: 1655

Danish Study Group on Verapamil in MI. The effect of verapamil on mortality and major events after acute MI (The Danish Verapamil Infarction Trial II – DAVIT II). *American Journal of Cardiology* 1990; 66: 779-85

Danish Study Group on Verapamil in MI. Verapamil in acute myocardial infarction. *European Heart Journal* 1984; 5: 516-528

DART 1989] Burr ML, Fehily AM, Gilbert JF, Rogers S, Holliday RM, Sweetnam PM et al. Effects of changes in fat, fish and fibre intakes on death and myocardial reinfarction : Diet and reinfarction trial (dart). *Lancet* 1989; ii: 758-61

De Bock V, Mets T, Romagnoli M, Derde MP. Captopril treatment of chronic heart failure in the very old. *Journal of Gerontology* 1994;49:M148-52.

Dickstein K, Barvik S, Aarsland T. Effect of long-term enalapril therapy on cardiopulmonary exercise performance in men with mild heart failure and previous MI. *J Am Coll Cardiol* 1991;18:596-602.

Dössegger L, Aldor E, Baird MG et al. Influence of angiotensin converting enzyme inhibition on exercise performance and clinical symptoms in chronic heart failure: a multicentre, double-blind, placebo-controlled trial. *Eur Heart J* 1993; 14(Suppl C): 18-23.

Dotremont G, De Geest H, The Use of Propranolol (Inderal) in Acute MI. *Acta Clinica Belgica*. 1968;23(3):163-76

Drexler H, Banhardt BS, Meinertz T, Wollschlager H, Lehmann M, Just H. Contrasting peripheral short-term and long-term effects of converting enzyme inhibition in patients with congestive heart failure: a double blind, placebo controlled trial. *Circulation* 1989;79:491-502.

Dubach P, Myers J, Dziekan G, Goebbels U, Reinhart W, Vogt P et al. Effect of exercise training on myocardial remodeling in patients with reduced left ventricular function after myocardial infarction. Application of magnetic resonance imaging. *Circulation* 1997; 95: 2060-7.

Dunselman PHJM, Kuntze CEE, van Bruggen A, Hamer JPM, Scaf AHJ, Wesseling H et al. Efficacy of felodipine in congestive heart failure. *European Heart Journal* 1989; 10: 354-64

Eichhorn EJ, Heejch CM, Barnet JH et al. Effect of metoprolol on myocardial function and energetics in patients with non-ischaemic dilated cardiomyopathy: a randomised double blind, placebo controlled study. *J Am Coll Cardiol* 1994; 24: 1310-20

EIS] European Infarction Study Group. European Infarction Study (EIS). A secondary prevention study with slow release oxprenolol after MI: morbidity and mortality. *European Heart Journal* 1984; 5: 189-202.

Enalapril CHF Investigators. Long-term effects of enalapril in patients with congestive heart failure: a multicentre, placebo-controlled trial. *Heart Failure* 1987;3:102-107.

Engelmeier RS, O'Connell JB, Walsh R, Rad N, Scainlon PJ, Gunnar RM. Improvement in symptoms and exercise tolerance by metoprolol in patients with

dilated cardiomyopathy: a double blind randomised placebo controlled trial. *Circulation* 1985; 72: 536-46

Erbel R, Pop T, Meinertz T, Olshausen KV, Treese N, Henrichs KJ et al. Combination of calcium channel blocker and thrombolytic therapy in acute MI. *Am Heart J* 1988; 115: 529-38.

Evemy KL, Pentecost BL. Intravenous and oral practolol in the acute stages of MI. *European Journal of Cardiology*. 1978; 7: 391-8.

Farchtein I, Roque AR, Cardoso C, Toledo F. Enalapril no tratamento da insuficiencia cardiaca congestiva severa. Estudo duplo-cego, controlado com placebo. *Arq Bras Cardiol* 1986; 47: 119-22.

Federman 1984] Reported in Yusuf S, Peto R, Lewis J, Collins R, Sleight P. β -blockade during and after MI. An overview of the randomized trials. *Progress in Cardiovascular Diseases* 1985;27:225-71

FHFSG] Brown EJ, Chew PH, MacLean A, Gelperin K, Ilgenfritz JP, Blumethal M, for the FOSINOPRIL HEART FAILURE STUDY GROUP. Effects of fosinopril on exercise tolerance and clinical deterioration in patients with chronic congestive heart failure not taking digitalis. *Am J Cardiol* 1995;75:596-600.

Figulla HR, Gietzen F, Zeymer U, Raiber M, Hegselmann J, Soballa R et al. Diltiazem improves cardiac function and exercise capacity in patients with idiopathic dilated cardiomyopathy. *Circulation* 1996; 94: 346-52.

Fisher ML, Gottlieb SS, Plotrick GD et al. Beneficial effects of metoprolol in heart failure associated with coronary artery disease: a randomised trial. *J Am Coll Cardiol* 1994; 23: 943-50

Foy SG, Crozier IG, Turner JG, Richards AM, Frampton CM, Nicholls G, Ikram H. Comparisons of enalapril versus captopril on left ventricular function and survival three months after acute MI (the "PRACTICAL" study). *Am J Cardiol* 1994;73:1180-6.

Fridlund B, Högstedt B, Lidell E, Larsson PA. Recovery after myocardial infarction: effects of a caring rehabilitation programme. *Scand J Caring Sci* 1991; 5: 23-32.
Fuccella LM: report on the double blind trial with compound CIBA 39089 (trasicor) in MI. In Sowton E: *Prog Cardiovasc Dis* 1968; 10: 561-74

Galcera-Tomas J, Antonio Nuno De La Roas J, Torres-Martinez G, Rodriguez-Garcia P, Castillo-Soria FJ, Canton-Martinez A, Campos-Peris JV et al. Effects of early use of captopril on haemodynamics and short-term ventricular remodelling in acute anterior MI. *Eur Heart J* 1993;14:259-66.

German-Austrian Study Group] Breddin K, Loew D, Lechner K, Oberla K, Walter E on behalf of the German-Austrian Study Group. The German-Austrian Aspirin Trial: a comparison of acetylsalicylic acid, placebo and phenprocoumon in secondary prevention of MI. *Circulation* 1980; 62(Suppl V): V63-V72

Giannuzzi P, Tavazzi L, Temporelli PL, Corrà U, Impartato A et al. Long-term physical training and left ventricular remodeling after anterior myocardial infarction: results of the exercise in anterior myocardial infarction (EAMI) trial. *J Am Coll Cardiol* 1993; 22: 1821-9.

Gibson RS, Boden WE, Theroux P, Strauss HD, Pratt CM, Gheorghide M et al. Diltiazem and reinfarction in patients with non-Q-wave MI: Results of a double-blind, randomised, multicenter trial. *N Engl J Med* 1986; 315: 423-9.

Gilbert EM, Anderson JL, Deitchman D et al. Long-term beta-blocker vasodilator therapy improves cardiac function in idiopathic dilated cardiomyopathy: a double-blind randomised study of bucindolol versus placebo. *Am J Med* 1990; 88: 223-9.

Gilbert EM, Sandoval A, Larabee P, Renlund D, O'Connell JB, Bristow MR. Lisinopril lowers cardiac adrenergic drive and increases β -receptor density in the failing human heart. *Circulation*. 1993;126:798-826.

Giles TD, for Lisinopril Chronic Heart Failure Group. Lisinopril treatment of congestive heart-failure results of a placebo controlled trial. *Circulation*. 1990;82(suppl 4):III-323.

Goldbourt U, Behar S, Reicher-Reiss H, Zion M, Mandelzweig L, Kaplinsky E et al. Early administration of nifedipine in suspected acute MI. *Arch Intern Med* 1993; 153: 345-53

Gordon M. Evaluation Of The Efficacy And Safety Of Ramipril (HOE 498) In Patients With Congestive Heart Failure In A Placebo-Controlled Trial. Somerville, NJ: Hoechst-Roussel Pharmaceuticals; October 1991. Unpublished report.

Gottlieb SO, Becker LC, Weiss JL, Shapiro EP, Chandra NC, Flaherty JT et al. Nifedipine in acute MI: an assessment of left ventricular function, infarct size and infarct expansion. *Br Heart J* 1988; 59: 411-8

Gøtzsche CO, Søgaard P, Ravkilde J, Thygesen K. Effects of captopril on left ventricular systolic and diastolic function after acute MI. *Am J Cardiol* 1992; 70: 156-60.

Gruppo Italiano per lo Studio della Sopravvivenza nell'infarto Miocardico (GISSI-3). GISSI-3: effects of lisinopril and transdermal glyceryl trinitrate singly and together on 6-week mortality and ventricular function after acute MI. *Lancet* 1994;343:1115-22.

Gupta RC, Butaney B, Narang NK. Effect of IV propranolol on the extent of myocardial ischemic injury in patients of acute anterior MI (AAMI). *JAPI* 1984; 32: 67-8.

Gupta RC, Sharma SK, Arora YK, Tripath K. Reduction of infarct size by early use of oral propranolol and verapamil in acute MI. *JAPI* 1985; 33: 577-80

Gupta RC, Sharma SK, Mittal SR. Effect of oral propranolol on the extent of acute anterior MI. *Abstracts of 9th World Congress on Cardiology, Moscow* 1982; 2: 616

Gutiérrez JV, Navarro-Robles J, Tello Osorio RT, Borges-Moreno J, Santos HLDL. Prevencion de la dilatacion del ventriculo izquierdo postinfarto del miocardio por captopril. Arch Inst Cardiol Méx 1992; 62: 533-9

Hansteen V, Møinichen E, Lorentsen E, Andersen A, Strøm O, Søliland K et al. One year's treatment with propranolol after MI: preliminary report of Norwegian multicentre trial. BMJ 1982; 284: 155-60.

Hattori Y, Atsushi S, Hiroaki F, Toyama J. Effects of cilazapril on ventricular arrhythmia in patients with congestive heart failure. Clinical Therapeutics 1997; 19: 481-6

Heber ME, Rosenthal E, Thomas N, Haskett VL, Burwood RD, Lutkin J, et al. Effect of labetalol on indices of myocardial necrosis in patients with suspected acute infarction. European Heart Journal 1987; 8: 11-18.

Hjalmarson Å, Abelardo N, Caidahl K, Reyes C, Waagstein F, Wallentin I et al. Effects of prenalterol administered orally in patients with congestive heart failure. Acta Medica Scandinavica - Supplementum 1982;659:201-20

Hjalmarson A, Elmfeldt D, Herlitz J, Holmberg S, Malek I, Nyberg G et al. Effect on mortality of metoprolol in acute MI: A double blind randomised trial. Lancet 1981; ii: 823-7.

HOPE] The Heart Outcomes Prevention Evaluation Study Investigators. Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. N Engl J Med 2000;342:145-53.

Hutton I, Vallance BD, Beattie JM et al. A prospective randomized trial of propranolol in acute MI. Excerta Med Inst Congr Ser 1979; 2: 824-6.

ICSG] International Collaborative Study Group. Reduction of infarct size with the early use of timolol in acute MI. N Engl J Med 1984; 310: 9-15.

Ikram H, Fitzpatrick D. Double-blind trial of chronic oral beta-blockade in congestive cardiomyopathy. Lancet 1981; 2: 490-3.

International Steering Committee on behalf of the MERIT-HF Study Group. Rationale, design and organisation of the Metoprolol CR/XL randomized intervention trial in Heart Failure (MERIT-HF). Am J Cardiol 1997; 80(Suppl 9B): 54j-58j.

ISIS-1 (First International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous atenolol among 16,027 cases of suspected acute MI: ISIS1. Lancet, 1986;ii:57-67

ISIS-4 (fourth International Study of Infarct Survival) Collaborative Group. ISIS-4: A randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58050 patients with suspected acute MI. Lancet 1995;345:669-85.

ISIS-4 (Fourth International Study of Infarct Survival) Pilot Study Investigators. Randomized control trial of oral captopril, of oral isosorbide mononitrate and of intravenous magnesium sulphate started early in acute MI: safety and haemodynamic effects. *European Heart Journal*. 1994;15:608-19.

Jakob H, Sigmund M, Eschenhagen T, Mende U, Patten M, Schmitz W et al. Effect of captopril on myocardial α adrenoceptor density and $G_{i\alpha}$ proteins in patients with mild to moderate heart failure due to dilated cardiomyopathy. *Eur J Clin Pharmacol* 1995; 47: 389-94.

Jansson JH, Boman K, Nilsson TK. Enalapril related changes in the fibrinolytic system in survivors of MI. *Eur J Clin Pharmacol* 1993; 44: 485-8

Johansson BW. A comparative study of cardioselective β -blockade and diazepam in patients with acute MI and tachycardia. *Acta Med Scand* 1980; 207: 47-53.

Johnson DB, Foster RE, Barilla F, Blackwell GG, Roney M, Stanley AWH et al. Angiotensin-converting enzyme inhibitor therapy affects left ventricular mass in patients with ejection fraction > 40% after acute MI. *J Am Coll Cardiol* 1997; 29: 49-54.

Julian DG, Prescott RJ, Jackson FS, Szekely P. Controlled trial of sotalol for one year after MI. *Lancet* 1982;i:1142-7.

Kafatos A, Diacatou A, Voukiklaris G, Nikolakakis N, Vlachonikolis J, Kounali D et al. Heart disease risk-factor status and dietary changes in the Cretan population over the past 30 y: the Seven Countries Study. *American Journal of Clinical Nutrition* 1997; 65:1 882-6

Kahler RL, Brill SJ, Perkins WE. The role of propranolol in the management of acute MI. In Kattus AA, Ross G, Hall YE (eds). *Cardiovascular β adrenergic responses*. Los Angeles: California Press, 1968, 213-222.

Kaul AU, Verma R, Garg KC. Early intervention with propranolol after acute MI: serial left ventricular function determined by M-Mode and cross sectional echocardiography. *International Journal of Cardiology* 1988; 21: 301-10.

Keren G, Pardes A, Eschar Y, Hansch E, Scherez J, Laniado S. Left ventricular filling dynamics by doppler echocardiography in dilated cardiomyopathy: one-year follow up in patients treated with captopril compared to placebo. *Cardiology* 1992; 81: 196-206

Kingma JH, van Gilst WH, Peels CH, Cambrink H-H E, Verheugts FWA, Wielenga RP for the CATS investigators. *Eur Heart J* 1994;15:898-907.

Kleber FX, Niemoller L, Doering W. Impact of converting enzyme inhibition on progression of chronic heart failure: results of the Munich mild heart failure trial. *Br Heart J*. 1992;67:289-96.

Kleber FX, Sabin GV, Winter UJ, Reindl I, Beil S, Wenzel M et al. Angiotensin-converting enzyme inhibitors in preventing remodeling and development of heart failure after acute MI: results of the German multicenter study of the effects of captopril on cardiopulmonary exercise parameters. *Am J Cardiol* 1997; 80(Suppl 3A):162A-167A.

KleberFX, Niemöller L. Long-term survival in the Munich Mild Heart Failure Trial (NHFT). *Am J Cardiol* 1993; 71: 1237-9

Kobayashi Y, Goto Y, Daikoku S, Itoh A, Miyazaki S, Ohshima S et al. Cardioprotective effect of intravenous nicorandil in patients with successful reperfusion for acute MI. *Jpn Circ J* 1998; 62: 183-9

Kober L, Torp-Pedersen C, Carlesen JE, Bagger H, Eliassen P, Lyngborg K et al for the Trandolapril Cardiac Evaluation (TRACE) Study Group. *N Engl J Med* 1995;333:1670-6.

Krum H, Sackner-Bernstein J, Goldsmith RL et al. Double blind, placebo controlled study of the long term efficacy of carvedilol in patients with severe chronic heart failure. *Circulation* 1995; 92: 1499-506.

Latini R, Avanzini R, De Nicolao A, Rocchetti M, and the GISSI-3 investigators. Effects of lisinopril and nitroglycerin on blood pressure early after MI: The GISSI-3 pilot study. *Clin Pharmacol Ther* 1994;56:680-92.

Launbjerg J, Fruergaard P, Madsen JK, Fischer Hansen J & The Danish Study Group on Verapamil in MI. Three-year mortality in patients suspected of acute MI with and without confirmed diagnosis. *Am Heart J*,1991;122:1270-3

Lechat P, Garnham J, Desche P, Bounhoure JP. Efficacy and acceptability of perindopril in mild to moderate chronic congestive heart failure. *Am Heart J*. 1993;126:798-826.

Ledwich JR. A trial of propranolol in MI. *Canad Med Ass J* 1968; 98: 988-94

Lemarie JC. Multicentre Double Blind Placebo Controlled Study Of The Efficacy And Safety Of Ramipril Administered Orally For 24 Weeks In The Treatment Of Stable Chronic Congestive Cardiac Failure. Paris, France: Laboratories Hoechst; December 1992. Unpublished report.

Leung W, Lau C, Wong C, chang C, Tai Y, Lim S. A double-blind cross-over trial of labetalol in patients with dilated cardiomyopathy. *Am Heart J* 1990; 119: 884-90.

LIPID] Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. *N Engl J Med* 1998; 339: 1349-57

LIT] Lopressor Intervention Trial Research Group. The lopressor intervention trial: multicentre study of metoprolol in survivors of acute MI. *European Heart Journal* 1987; 8: 1056-64.

Littler WA, Sheridan DJ, on behalf of the UK Study Group. Placebo controlled trial of felodipine in patients with mild to moderate heart failure. *Brit Heart J* 1995; 73: 428-33

Lloyd EA, Charles RG, Gordon GD, Adams CM, Mabin TA, Commerford PJ, Opie LH. β -blockade by sotalol in early MI decreases ventricular arrhythmias without increasing left ventricular volume. *SAMJ* 1988; 74: 5-10.

Lombardo M, Selvini A, Belli C, Motolese M, Pedroni P. β -blocking therapy in 440 cases of acute MI: a double-blind trial with oxprenolol. *Proceedings of Florence International Meeting on MI* 1979; 2: 803-7.

Lu Cai. The treatment of acute MI with oral captopril: a randomized double blind and placebo controlled pilot study. *Chinese Medical Journal* 1993; 106: 717

Lubsen J, Chadha DR, Yotof YT, Swedberg K. Meta analysis of morbidity and mortality in five exercise capacity trials evaluating ramipril in chronic congestive cardiac failure. *Am J Cardiol* 1996; 77: 1191-6

Lyon 1994] Lorgeril M, Renaud S, Mamelle N, Salen P, Martin JL et al. Mediterranean alpha-linolenic acid-rich diet in secondary prevention of coronary heart disease. *Lancet* 1994; 343: 1454-9.

MAAS Investigators. Effect of simvastatin on coronary atheroma: the multicentre anti-atheroma study (MAAS). *Lancet* 1994; 344: 633-8.

Maass L. Double Blind Comparative Trial With Ramipril And Placebo In Patients With Heart Failure (NYHA Class III-IV) Stabilised On Digitalis And Ferusemides. Frankfurt, Germany: Hoechst Aktiengesellschaft; October 1991a. Unpublished report.

MacFadyen RJ, Barr CS, Sturrock NDC, Fenwick M, Struthers AD. Further evidence that chronic perindopril treatment maintains neurohormonal suppression but does not lower blood pressure in chronic cardiac failure. *Br J Clin Pharmacol* 1997; 44: 69-76

Macleod A, Fananapazir L, Kitchin AH, Murray A, Neilson J.M. Prophylactic selective β -blockade in acute MI. *Abstracts of VIII European Congress on Cardiology Paris*. 1980, (abstr).

Magnani B, Magelli C, for the Multicentre Research Groups on Mild Heart Failure. Captopril in mild heart failure: preliminary observations of a long-term double-blind, placebo-controlled multicentre trial. *Postgrad Med J*. 1986;62(suppl I):153-158.

Maisel AS. Beneficial effects of metoprolol treatment in congestive heart failure: Reversal of sympathetic induced alterations of immunologic function. *Circulation* 1994; 90: 1774-80.

Malmberg K. Prospective randomised study of intensive insulin treatment on long term survival after acute myocardial infarction in patients with diabetes mellitus.

DIGAMI (Diabetes Mellitus, Insulin Glucose Infusion in Acute Myocardial Infarction) Study Group. *BMJ* 1997 ;314 :1512-5

Manger Cats V, van Capelle FJL, Lie KI, Durrer D. Effect of treatment with 2x100 mg metoprolol on mortality in a single-center study with low placebo mortality rate after infarction. *Circulation* 1983; 68(Suppl 3): 181

Marra S, Paolillo V, Spadaccini F, Angelino PF. Long-term follow up after a controlled randomized post-myocardial infarction rehabilitation programme: effects on morbidity and mortality. *European Heart Journal* 1985; 6: 656-63

MARS] Blankenhorn DH, Azen SP, Krams DM, Mack WJ, Cashin-Hemphill L, Hodis HN et al. Coronary angiographic changes with lovastatin therapy. The monitored atherosclerosis regression study (MARS). *Ann Intern Med* 1993; 119: 969-76

Mazur NA, Kulginskaya IV, Ivanova LA, Ostrovskaya TP, Smirnova TM, Svet EA et al. Results of long term propranolol treatment in MI survivors with advanced grades of ventricular extasystoles: Randomized study. *Cor Vasa* 1984; 26: 241-7.

McGarry R. Randomized, double blind, multicentre study comparing benazepril to digoxin and to placebo as add on therapy to diuretic in patients with CHF, NYHA Class II-III during a 12-week treatment period, GHBA-194. Summit, NJ: Ciba-Geigy Pharmaceuticals; September 1991. Unpublished Report.

McMurray J, Lang CC, MacLean D, Struthers AD, McDevitt DG. Effects of xamaterol in acute MI: blood pressure, heart rate, arrhythmias and early clinical course. *International Journal of Cardiology* 1991;31:295-304.

Metra M, Nardi M, Giubbini R, Dei Cas L. Effects of short and long-term carvedilol administration on rest and exercise haemodynamic variable, exercise capacity and clinical conditions in patients with idiopathic dilated cardiomyopathy. *J Am Coll Cardiol* 1994; 24: 1678-87

Miscristin] Vogel G, Fischer C, Huyke R. Prevention of reinfarction with acetylsalicylic acid. In: Breddin HK, Loew D, Uberla K, Dorndoff W, Marx R, eds. *Prophylaxis of venous peripheral cardiac and cerebral vascular diseases with acetylsalicylic acid.* Stuttgart: Shattauer, 1981:123-8.

Mueller HS, Ayres SM. Propranolol decreases sympathetic nervous activity reflected by plasma catecholamines during evolution of MI in man. *J Clin Invest* 1980; 65: 338-46.

Muller JE, Morrison J, Stone PH, Rude RE, Rosner B, Roberts R et al. Nifedipine therapy for patients with threatened and acute MI: a randomized, double blind, placebo-controlled comparison. *Circulation* 1984; 69: 740-7

MULTICENTRE 1966] Bath JCJL, Blakc S, Bloom RA, Brown R, Fleming HA, Franklin AJ et al. Propranolol in acute MI: A Multicentre Trial. *Lancet* 1966; ii: 1435-8.

MULTICENTRE 1966] Bath JCJL, Blakc S, Bloom RA, Brown R, Fleming HA, Franklin AJ et al. Propranolol in acute MI: A Multicentre Trial. *Lancet* 1966; ii: 1435-8.

Nabel EG, Topo EJ, Galeana A, Ellis SG, Bates ER, Werns SW et al. A randomized placebo-controlled trial of combined early intravenous captopril and recombinant tissue-type plasminogen activator therapy in acute MI. *JACC* 1991;17:467-73.

Nguyen KN, Aursnes I, Kjekshus J. Interaction between enalapril and aspirin on mortality after acute MI: subgroup analysis of the cooperative new scandinavian enalapril survival study II (CONSENSUSII). *Am J Cardiol* 1997; 79: 115-9.

Nicholls MG, Ikram H, Espiner EA, Webster MWI, Fitzpatrick MA. Enalapril in heart failure. *Br J Clin Pharmacol* 1984;18:163s-7s.

Nicolau JC, Ramires JAF, Maggioni AP, Garzon SAC, Pinto MAFV, Silva DG et al. Diltiazem improves left ventricular systolic function following acute MI treated with streptokinase. *Am J Cardiol* 196; 78: 1049- 52.

Nigam PD, Popli R, Charan S. Cardioprotection by β -blockade in acute myocardial infarction. Abstracts of the 9th World Congress of Cardiology, Moscow, 2:616, 1983 (abstr)

Norris RM, Barnaby PF, Brown MA, Geary GG, Clarke ED, Logan RL et al. Prevention of ventricular fibrillation during acute MI by intravenous propranolol. *Lancet* 1984; ii: 883-6.

Norris RM, Caughey DE, Scott PJ. Trial of Propranolol in Acute MI. *BMJ* 1968; 2:398-400.

Norris RM, Clarke ED, Sammel NL, Smith WM, Williams B. Protective effect of propranolol in threatened MI. *Lancet* 1978;ii:907-9.

North of England Evidence-based Guideline Project. The Primary Care Management of Stable Angina. Centre For Health Services Research, Newcastle Upon Tyne 1999. Report No 98.

Northridge DB, Rose E, Raftery ED, Lahiri A, Elder AT, Shaw TRD et al. A multicentre, double-blind, placebo-controlled trial of quinapril in mild chronic heart failure. *Eur Heart J*. 1993;14: 403-9.

Norwegian Multicentre Study Group. Timolol induced reduction in mortality and reinfarction in patients surviving acute MI. *N Engl J Med* 1981;304:801-7.
Office for National Statistics. 1997 Mortality statistics: cause. Series DH2 no. 24. London: The Stationery Office, 1998.

Olsen SL, Gilbert EM, Renlund DG et al. Carvedilol improves left ventricular function and symptoms in chronic heart failure: a double blind randomised study. *J Am Coll Cardiol* 1995; 25: 1225-31.

Owensby DA, O'Rourke MF. Failure of pindolol to alter determinants of myocardial oxygen requirements, enzyme release or clinical course in acute MI. *Circulation* 1984; 70(Suppl II): 156.

P.RE.COR Group. Comparison of a rehabilitation programme, counselling programme and usual care after an acute myocardial infarction: results of a long-term randomized trial. *European Heart Journal* 1991; 12: 612-6.

Pacher R, Hurlsmann M, Stanek B, Berger R, Wutte M, Globits S, et al. Atenolol: additional short-term effects of a cardioselective beta blocker to tailored therapy in congestive heart failure. A randomized placebo-controlled pilot study. *European Heart Journal*, 1995; 16(Suppl): 157.

Packer M, Bristow M, Cohn J et al. For the US Carvedilol Heart Failure Study Group. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. *N Engl J Med* 1996; 334: 1349-55

Packer M, Colucci WS, Sackner BJ et al. Double blind placebo controlled study of the effects of carvedilol in patients with moderate to severe heart failure: the prospective randomized evaluation of carvedilol in patients with symptoms and exercise. *Circulation* 1996; 94: 2793-9.

Packer M, O'Connor CM, Ghali JK, Pressler ML, Carson PE, Belkin RN et al. Effect of amlodipine on morbidity and mortality in severe chronic heart failure. *N Engl J Med* 1996; 335: 1107-14

Paolisso G, Gambardella A, Marrazzo G, Vera M, Teasuro P, Varricchio M, D'Onofrio F. Metabolic and cardiovascular benefits deriving from beta-adrenergic blockade in chronic congestive heart failure. *Am Heart J* 1992; 123: 1397-404.

PARIS 1] Persantine-Aspirin Reinfarction Study (PARIS) Research Group. Persantine and aspirin in coronary heart disease. *Circulation* 1980a;62:449-61.

PARIS 2] Klimt CR, Knatterud GL, Stamler J, Meier P. Persantine-aspirin reinfarction study. Part II. Secondary coronary prevention with Persantine and aspirin. *J Am Coll Cardio* 1986;7:251-69.

Peter T, Norris RM, Clarke ED, Heng MK, Singh BN, Williams B et al. Reduction of enzyme levels by propranolol after acute MI. *Circulation* 1978; 57: 1091-5.

Pfeffer MA, Branunwald E, Moy LA, Basta L, Brown EJ, Cuddy TE et al for the SAVE Investigators. Effect of captopril on mortality and morbidity in patients with left ventricular dysfunction after MI. Results of the Survival and Ventricular Enlargement Trial. *N Engl J Med* 1992;327:669-77.

Pfeffer MA, Greaves SC, Arnold MO, Glynn RJ, LaMotte FS, Lee RT et al. Early versus delayed angiotensin-converting enzyme inhibition therapy in acute MI: The Health and Early Afterload Reducing Therapy Trial. *Circulation* 1997; 95: 2643-51.

Pflugfelder PW, Baird MG, Tonkon MJ, DiBianco R, Pitt B for the Quinapril Heart Failure Trial Investigators. *J Am Coll Cardiol* 1993; 22: 1557-63

Pitt B, Zannad F, Remme WJ, Cody R, Castaigne A, Perez A, et al. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. *N Engl J Med* 1999; 341; 709-17.

Pitt B and the RALES investigators. the Randomised Aldactone Evaluation Study (RALES) Parallel dose finding trial (abstr). *J Am Coll Cardiol* 1995;25:45A

Pitt B, Weiss JL, Schulze RA, Taylor DR, Kennedy HL, Caralis D. Reduction of myocardial infarct extension in man by propranolol. *Circulation* 1976; 29(Suppl2): 53-4.

PLAC-II] Crouse JR, Byington RP, Bond MG, Espeland MA, Craven TE, Sprinkel JW et al. Pravastatin, lipids, and atherosclerosis in carotid arteries (PLAC-II). *Am J Cardiol* 1995; 75: 455-9.

Platsi I. Feasibility of physical training after myocardial infarction and its effect on return to work, morbidity and mortality. *Acta Med Scan* 1976; (Suppl 599): 4 – 84.

Pollock SG, Lytash J, Tedesco C, Craddock G, Smucker ML. Usefulness of bucindolol in congestive heart failure. *Am J Cardiol* 1990; 66: 603-7.

PROVED] Darasz KH, Bayliss J, Underwood SR, Keegan J, Poole Wilson PA, Sutton GC. *International Journal of Cardiology* 1995;51:137-42.

Ranganathan N, Rautaharju PM, Jablonsky GC, Laroche P, Lopez JF, Matangi MF. Prophylaxis of post MI dysrhythmias by long term timolol therapy. *Am Heart J* 1988; 115: 340-50.

Ray S, Pye M, Oldroyd KG, Christie J, Connelly DT, Northridge DB, et al. Early treatment with captopril after acute MI. *Br Heart J* 1993;69:215-22.

Rehmqvist N, Ahnve S, Erhardt L, Lindvall K, Lundman T, Olsson G, Sjogren A. Effect of metoprolol after acute MI. *Proc Eur Cong Cardiol* 16, 1980 (abstr)

Rehmqvist N, Olsson G. Influence on ventricular arrhythmias by chronic post infarction treatment with metoprolol. *Circulation* 1983; 68 (Suppl 3): 369.

Rengo F, Carbonin P, Pahor M, De Caprio L, Bernabei R, Ferrara N et al. A controlled trial of verapamil in patients after acute MI: results of the calcium antagonist reinfarction Italian study (CRIS). *Am J Cardiol* 1996; 77: 365-9.

Reynolds JL, Whitlock RML. Effects of a β -adrenergic receptor-blocker in MI treated for one year from onset. *Br Heart J* 1972; 34: 252-9.

Riegger GAJ. The effects of ACE inhibitors on exercise capacity in the treatment of congestive heart failure. *J Cardiovasc Pharmacol.* 1990;15(suppl 2):41-6.

Roberts R, Croft C, Gold HK, Hartwell TD, Jaffe AS, Muller JE et al. Effect of propranolol on myocardial infarct size in a randomized blinded multicenter trial. *N Engl J Med* 1984;311:218-25.

Rodríguez JA, Godoy I, Castrol PG, Quintana JCF; Chávez ES, Yovanovich JS et al. Ramipril vs espirolactona en el remodelamiento ventricular izquierdo post-infarto: estudio randomizado y doble ciego. *Rev Méd Chile* 1997; 125: 643-52

Román O, Gutierrez M, Luksick I, Chavez E, Camuzzi AL, Villalón E, et al. Cardiac rehabilitation after acute myocardial infarction: 9 year controlled follow up study. *Cardiology* 1983; 70: 223-31.

Rome] Gentile R, Laganà B, Calcagni S, Borgia MC, Baratta L. Efficacy of platelet inhibiting agents in the prevention of reinfarction in smoker patients. In: *Proceedings of X world congress on cardiology, Washington, 1986:302. (Abstract 1724).*

Rossi P, Giordano A, Tamiz A, Schiavo B, Minuco G. Effects of training and of verapamil on exercise capacity in patients recovering from MI: a double blind study. *Cardiology* 1979; 64: 372-85.

Rousseau MF, Chapelle F, Van Eyll C, Stoleru L, Hager D, Van Nueten L et al. Medium term effects of beta blockade on left ventricular mechanics; a double-blind, placebo-controlled comparison of nebivolol and atenolol in patients with ischemic left ventricular dysfunction. *Journal of Cardiac Failure* 1996; 2: 15-23.

Rucinska EJ. A double blind placebo controlled study to evaluate the effects of enalapril in patients with chronic heart failure. West Point, Pa: Merck Sharpe & Dohme Research Laboratories; September 1991a. Unpublished report.

Rucinska EJ. Enalapril vs placebo in previously untreated patients with CHF. West Point, Pa: Merck Sharpe & Dohme Research Laboratories; September 1991b. Unpublished report.

Ruckinska EJ. Lisinopril first line therapy in CHF. West Point, Pa: Merck Sharpe & Dohme Research Laboratories; September 1991c. Unpublished report.

Russo R, Rao MAE, Mele AF, Cangianiello S, Giunta A, Cardei S, et al. Long-term effects of felodipine in patients with mild heart failure treated chronically with enalapril: a randomized placebo controlled trial. *Current Therapeutic Research* 1998; 59: 288-306.

Salathia KS, Barber JM, McIlmoyle EL, Nicholas J, Evans AE, Elwood JH, et al. Very early intervention with metoprolol in suspected acute MI. *Eur Heart J* 1985; 6: 190-8.

Sanderson JE, Chan SKW, Yu CM, Yeung LYC, Chan WM, Raymond K, et al. β blockers in heart failure: comparison of a vasodilating β blocker with metoprolol. *Heart* 1998; 79: 86-92.

Sano H, Kawabata N, Yonezawa K, Hirayana H, Sakuma I, Yasuda H. Metoprolol was more effective than captopril for dilated cardiomyopathy in Japanese patients (Abstr.). *Circulation* 1989; 80(Suppl II): II-118

Santiparpluacha C, Yipintsoi T, Jintapakorn W, Jirathamopas W, Kunjara-N-Ayudhya R, Akaravinake A. Double-blind evaluation of enalapril in patients with systolic heart failure. *J Med Assoc Thai* 1992; 75: 146-52

Schulman SP, Wiess JL, Becker LC, Guerci AD, Shapiro EP, Chandra NC et al. Effect of early enalapril therapy on left ventricular function and structure in acute MI. *Am J Cardiol* 1995; 76: 764-70

Sharpe N, Smith H, Murphy J, Greaves S, Har H, Gamble G. Early prevention of left ventricular dysfunction after MI with angiotensin-converting-enzyme inhibition. *Lancet* 1991; 337: 872-76.

Shaw LW. Effects of a prescribed supervised exercise program on mortality and cardiovascular morbidity in patients after a myocardial infarction. *Am J Cardiol* 1981; 48: 39-46

Sigmund M, Jakob H, Becker H, Hanrath P, Schumacher C, Eschenhagen T et al. Effects of metoprolol on myocardial α -adrenoceptors and $G_{i\alpha}$ -proteins in patients with congestive heart failure. *Eur J Clin Pharmacol* 1996; 51: 127-32

Singh S, Saxena PC, Shukla SK. Intravenous propranolol in acute MI. *The Clinician* 1985; 49 :384-91.

Sirnes PA, Overskeid K, Pedersen TR, Bathen J, Drivenes A, Frøland GS et al. Evolution of infarct size during early use of nifedipine in patients with acute MI: the Norwegian Nifedipine Multicenter Trial. *Circulation* 1984; 70: 638-44.

Sivarajan ES, Bruce RA, Lindskog BD, Almes MJ, Belanger L, Green B. Treadmill test responses to an early exercise program after myocardial infarction: a randomised study. *Circulation* 1982; 65: 1420-8

Sloman G, Stannard M. β -adrenergic-blockade and cardiac arrhythmias. *BMJ* 1967; four :508-12.

Snow 1980, Reported in Yusuf S, Peto R, Lewis J, Collins R, Sleight P. β -blockade during and after MI: An overview of the randomized trials. *Prog Cardio Dis* 1985; 27: 335-71.

SOLVD Investigators. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N Engl J Med* 1991;325:293-302

SOLVD Investigators. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N Engl J Med* 1991;325:293-302.

SSSD] López FN, Cosin J, Marrugat J, Guindo J, Lunda ABD for the SSSD investigators. Comparison of the effects of amiodarone versus metoprolol on the frequency of ventricular arrhythmias and on mortality after acute MI. *Am J Cardiol* 1993; 72: 1234-8.

Swedberg G, Held P, Kjekshus J, Rasmussen K, Rydon L, Wedel H for the CONSENSUS II Study Group. Effects of the early administration of enalapril on mortality in patients with acute MI. Results of the co-operative New Scandinavian Enalapril Survival Study II (CONSENSUS II). *N Engl J Med* 1992;327:678-84.

Swedberg K, Amtorp O, Gundersen T, Remes J, Nilsson B. Is maximal exercise testing a useful method to evaluate treatment of moderate heart failure? *Circulation*. 1991;57:226. Abstract.

Taylor SH, Silke B, Ebbutt A, Sutton GC, Prout BJ, Burley DM. A long term prevention study with oxprenolol in coronary heart disease. *N Engl J Med* 1982; 307: 1293-1301.

The Acute Infarction Ramipril Efficacy (AIRE) study investigators. Effect of ramipril on mortality and morbidity of survivors of acute MI with clinical evidence of heart failure. *Lancet* 1993;342:821-8.

The DEFIANT Research Group (Doppler Flow and Echocardiography in Functional cardiac Insufficiency: assessment of Nisoldipine Therapy). Improved diastolic function with the calcium antagonist nisoldipine (coat-core) in patients post MI: results of the DEFIANT study. *European Heart Journal* 1992; 13: 1496-505.

The DEFIANT-II Research Group. Doppler flow and echocardiography in functional cardiac insufficiency: assessment of nisoldipine therapy. Results of the DEFIANT-II Study. *Eur Heart J* 1997; 18: 31-40

The Israeli Sprint Study Group. Secondary prevention reinfarction Israeli nifedipine trial (SPRINT). A randomised intervention trial of nifedipine in patients with acute MI. *European Heart Journal* 1988; 9: 354-64

The MIAMI Trial Research Group. Metoprolol in acute MI (MIAMI). A randomised placebo-controlled international trial. *European Heart Journal* 1985;6:199-226.
The Multicenter Diltiazem Postinfarction Trial Research Group. The effect of diltiazem on mortality and reinfarction after MI. *N Engl J Med* 1988; 319: 385-92

The TIMI Study Group. Comparison of invasive and conservative strategies after treatment with intravenous tissue plasminogen activator in acute MI: Results of the thrombolysis in MI (TIMI) phase II trial. *N Engl J Med* 1989; 320: 618-27.

Thompson PL, Jones AS, Noon D, Katavatis V. A randomised trial of oral β -blockade during MI - lack of effect on enzymatic indices of myocardial necrosis. *Aust NZ J Med* 1979; 9: 757.

Tonkin AM. β -blockade in acute MI: inability of relatively late administration to influence infarct size and arrhythmias. *Med J Aust* 1981; 2: 145-6.

Tsuyuki RT, Yusuf S, Rouleau JL, Maggioni AP, McKelvie RS, Wiecek EM et al. Combination neurohormonal blockade with ACE inhibitors, angiotensin II antagonists and beta-blockers in patients with congestive heart failure: Design of the randomized evaluation of strategies for left ventricular dysfunction (RESOLVD) Pilot Study. *Can J Cardiol* 1997; 13: 1166-74.

Uhlir O, Dvorak I, Gregor F, Malie I, Spinarova L, Vojacke I et al. Nebivolol in the treatment of cardiac failure: a double-blind controlled clinical trial. *Journal of Cardiac Failure* 1997; 3: 272-6

UKCSG] UK Collaborative Study Group: Mortality results from the timolol trials in various British hospitals. Reported in: Yusuf S, Peto R, Lewis J, Collins R, Sleight P. β -blockade during and after MI: An overview of the randomized trials. *Prog Cardio Dis* 1985; 27: 335-71.

Uprichard A. A 12-Week Double Blind, Placebo Controlled Study To Determine The Efficacy And Safety Of Orally Administered Quinapril Hydrochloride In Patients With Congestive Heart Failure (906-63). Ann Arbor, Mich: Parke-Davis Pharmaceutical Research; April 1994c. Unpublished report.

Uprichard A. An 18-Week Double Blind, Optional Titration, Multicentre Study To Compare The Efficacy And Safety Of Orally Administered Quinapril Hydrochloride With Captopril And Placebo In Patients With Congestive Heart Failure (906-226). Ann Arbor, Mich: Parke-Davis Pharmaceutical Research; April 1994b. Unpublished report.

Van de Werf F, Janssens L, Brzostek T, Mortelmans L, Wackers F, Willems GM et al. Short term effects of early intravenous treatment with a β -adrenergic-blocking agent or a specific bradycardiac agent in patients with acute MI receiving thrombolytic therapy. *J Am Coll Cardiol* 1993; 22: 407-16

Vermeulen A, Lie KI, Durrer D. Effects of cardiac rehabilitation after myocardial infarction: changes in coronary risk factors and long term prognosis. *Am Heart J* 1983; 105: 798-801

VON ESSEN] Essen Rv, Merx W, Neis W, Ritz R. Wirkung von metoprolol auf die infarkttröbe beim akuten myokardinfarkt: doppelbindstudie. *Deutsches Medizin Wochenschr* 1982; 107: 1267-73.

Vries RJM d, Dunselman PHJM. The potential role of calcium antagonists in the management of congestive heart failure: initial experience with lacidipine. *Journal of Cardiovascular Pharmacology* 1995; 25(Suppl 3):S33 to 39

Waagstein F, Bristow MR, Swedberg K et al for the Metoprolol in Dilated Cardiomyopathy (MDC) Trial Study Group. Beneficial effects of metoprolol in idiopathic dilated cardiomyopathy. *Lancet* 1993; 342: 1441-6.

Waagstein F, Hjalmarson ÅC. Double blind study of the effect of cardioselective β -blockade on chest pain in acute MI. *Acta Med Scand* 1975; 587(Suppl): 201-11

Walker LJE, MacKenzie G, Adgey AA. Effect of nifedipine on enzymatically estimated infarct size in the early phase of acute MI. *Br Heart J* 1988; 59: 403-10

Walsh JT, Andrews R, Curtis S, Evans A, Cowley AJ. Effects of amlodipine in patients with chronic heart failure. *Am Heart J* 1997; 134: 872-8

WHO] Lamm G, Denolin H, Dorossiev D, Pisa Z. *Adv Cardiol* 1982; 31: 107-11

WHO] O'Connor GT, Buring JE, Yusuf S, Goldhaber SZ, Olmstead EM, Paffenbarger RS, et al. An overview of randomized trials of rehabilitation with exercise after myocardial infarction. *Circulation* 1989; 80: 234-44.

Wilcox RG, Hampton JR, Banks DC, Birkhead JS, Brooksby IAB, Burns-Cox CJ et al. Trial of early nifedipine in acute MI: the Trent study. *BMJ* 1986; 293: 1204-8

Wilcox RG, Rowley JM, Hampton JR, Mitchell JRA, Rolan JM, Banks DC. Randomised placebo-controlled trial comparing oxprenolol with diopyramide phosphate in immediate treatment of suspected MI. *Lancet* 1980;ii:765-9.

Wilhelmsen L, Sanne H, Elmfeldt D, Grimby G, Tibblin G, Wedel H. A controlled trial of physical training after myocardial infarction. *Preventive Medicine* 1975; 4: 491-508

Wisnibaugh T, Katz I, Davi J et al. Long term (3 month) effects of a new beta blocker (nebivolol) on cardiac performance in dilated cardiomyopathy. *J Am Coll Cardiol* 1993; 21: 1094-100

Woodley SL, Gilbert EM, Anderson JL et al. Beta-blockade with bucindolol in heart failure caused by ischaemic versus idiopathic dilated cardiomyopathy. *Circulation* 1991; 84: 2426-41.

Yang XS, Coupez R, Ector H, Kesteloot H, Geest HD. Effects of betaxolol on heart rate in patients with a recent transmural MI. *Acta Cardiologica* 1987;XLII:273-86

Yodfat Y. Functional status in treatment of heart failure by captopril: a multicentre, controlled, double blind study in family practice. *Family Practice* 1991; 8: 409-11.

Yoshitomi Y, Kojima S, Yano M, Sugi T, Matsumoto Y, Saotome M et al. Early administration of the angiotensin-converting enzyme inhibitor imidapril attenuates ventricular remodeling after acute MI. *Jpn Pharmacol Ther* 1998; 26: 1545-52.

Yusuf S, Pepine CJ, Garces C, Pouleur H, Salem D, Kostis J et al. Effect of enalapril on MI and unstable angina in patients with low ejection fractions. *Lancet* 1992; 340: 1173-8.

Yusuf S, Ramsdale D, Peto R, Furse L, Bennet D, Bray C et al. Early intravenous atenolol treatment in suspected acute MI. *Lancet* 1980; 2: 273-6

Zannad F, Amor M, Karcher G, Maurin P, Ethevenot G, Sebag C et al. Effect of diltiazem on myocardial infarct size estimated by enzyme release, serial thallium-201 single-photon emission computed tomography and radionuclide angiography. *Am J Cardiol* 1988; 61: 1172-7.

Zwehl W, Rucinska E, for Lisinopril Chronic Heart Failure Investigators. Long-term effects of lisinopril in patients with chronic heart failure: a multicentre, placebo-controlled trial. In: Nicholls MG, ed. *A Focus on the Clinical Effects of a Long Acting ACE inhibitors/Heart Failure*. New York, NY: Raven Press; 1990:31-40.

TRIALS IN SCHIZOPHRENIA

Beasley CM, Hamilton SH, Crawford AM, Dellva MA, Tollefson GD, Tran PV et al. Olanzapine versus haloperidol: acute phase results of the international double-blind trial. *European Neuropsychopharmacology* 1997; 7: 125-37

Beasley CM, Tollefson G, Tran P, Satterlee W, Sanger T, Hamilton S, and the Olanzapine HGAD Study Group. Olanzapine versus placebo and haloperidol: acute phase results of the North American double-blind trial. *Neuropsychopharmacology* 1996; 14: 111-23.

Blin O, Azorin MJ, Bouhours P. Antipsychotic and anxiolytic properties of risperidone, haloperidol and methotrimeprazine in schizophrenic patients. *Journal of Clinical Psychopharmacology* 1996; 16: 38-44

Bondolfi G, Dufour H, Patris M, May JP, Billeter U, Eap CB et al. Risperidone versus clozapine in treatment resistant chronic schizophrenia: a randomised double blind study. *Am J Psychiatry* 1998; 155: 499-504

Borison RL, Pathiraja AP, Diamond BI, Meibach RC. Risperidone: clinical safety and efficacy in schizophrenia. *Psychopharmacology Bulletin* 1992; 28: 213-8.

Bouchard RH et al. ABSTRACT from APA 1998

Bowling A (1991). *Measuring health*. Buckingham: Open University Press

Breier A, Buchanan RW, Kirkpatrick B, Davis OR, Irish D, Summerfelt A et al. Effects of clozapine on positive and negative symptoms in outpatients with schizophrenia. *Am J Psychiatry* 1994; 151: 20-6.

Cešková E, Švestka J. Double-blind comparison of risperidone and haloperidol in schizophrenic and schizoaffective psychoses. *Pharmacopsychiatr* 1993; 26: 121-4

Chiu E, Burrows G, Stevenson J. *Australian & New Zealand Journal of Psychiatry* 1976; 10: 343-7.

Chouinard G, Jones B, Reminton G, Bloom D, Addington D, MacEwan GW et al. A Canadian multicenter placebo-controlled study of fixed doses of risperidone and haloperidol in the treatment of chronic schizophrenic patients. 1993; 13: 25-40

Ciurezu T, Ionescu R, Udangiu SN, Niturad D, Oprioiu L, Todorache D et al. Étude clinique en «double blind» DU HF 1854 (=LX 100-129 = clozapine=leponex) comparé À L'Halopéridol. *Neurol Psychiat* 1976; 14: 29-34.

Claus A, Bollen J, De Cuyper H, Eneman M, Malfrid M, Peuskens J et al. Risperidone versus haloperidol in the treatment of chronic schizophrenic inpatients: a multicentre double-blind comparative study. *Acta Psychiatr Scand* 1992; 85: 292-305

Clinical Trial GCP No. 93302. A dose ranging study comparing the efficacy, tolerability and safety of four doses of sertindole and one dose of haloperidol in schizophrenic patients: A multinational, prospective, randomised, double blind, controlled, parallel group study. Denmark: H Lundbeck A/S, 1997.

Conley RR, Tamminga CA, Bartko JJ, Richardson C, Peszke M, Lingle J, et al. Olanzapine compared with chlorpromazine in treatment-resistant schizophrenia. *Am J Psychiatry* 1998a; 155: 914-20

Conley RR, Brecher M and the Risperidone/Olanzapine Study Group. Paris, ECNP, 1998b.

Daniel DG, Wozniak P, Mack RJ, McCarthy BG and the Sertindole Study Group. Long-term and safety comparison of sertindole and haloperidol in the treatment of schizophrenia. *Psychopharmacology Bulletin* 1998; 34: 61-9

Delcker A, Schoon ML, Oczkowski B, Gaertner HJ. Amisulpride versus haloperidol in treatment of schizophrenic patients – results of a double-blind study. *Pharmacopsychiatry* 1990; 23: 125-30

Engelhart L, Mahmoud R and the Risperidone Outcome Study of Effectiveness (ROSE) Group. After schizophrenia relapse: findings from a prospective 684 patient cohort. Toronto, Annual Meeting of the American Psychiatric Association, May 30- June 4, 1998

Erlandsen C. Utprøving av et nytt nevroleptikum, Leponex (clozapin) hos schizofrene med lang sykehistorie. *Nord Psykiatr Tidsskr* 1981; 35: 248-53.

Essock S, Hargreaves WA, Covell NH, Goethe J. Clozapine's effectiveness for patients in state hospitals: results from a randomised trial. *Psychopharmacology Bulletin* 1996a; 32: 683-97.

Essock SM, Hargreaves WA, Dohm FA, Goethe J, Carver L, Hipshman L. Clozapine eligibility among state hospital patients. *Schizophrenia Bulletin* 1996b; 22: 15-25

Farde L, Nordstrom AL, Wiesel FA, Pauli S, Halldin C, and Sedvall G. Positron emission tomographic analysis of central D1 and D2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine. Relation to extrapyramidal side effects. *Arch Gen Psychiatry* 49 (7):538-544, 1992.

Gelenberg AJ, Doller JC. Clozapine versus chlorpromazine for the treatment of schizophrenia: preliminary results from a double-blind trial. *J Clin Psychiatry* 1979; 40: 238-40

Gerlach J, Koppelhus P, Helweg E, Monrad A. Clozapine and haloperidol in a single-blind cross over trial: therapeutic and biochemical aspects in the treatment of schizophrenia. *Acta Psychiatr Scand* 1974b; 50: 410-24

Gerlach J, Thorsen K, Fog R. Extrapyrarnidal reactions and amine metabolites in cerebrospinal fluid during haloperidol and clozapine treatment of schizophrenic patients. *Psychopharmacologia (Berl)* 1974a; 40: 341-50

Guirguis E, Voineskos G, Gray J, Schlieman E. Clozapine (leponex) versus chlorpromazine (largactil) in acute schizophrenia (a double-blind controlled study). *Current Therapeutic Research* 1977; 21: 707-19.

Hillert A, Philipp M, Gattaz WF, Sauer H, Adler G, Wetzel H, Benkert O and the Amisulpride Study Group. Amisulpride vs flupentixol in the treatment of schizophrenia with predominant positive symptomatology: a controlled double-blind study. *Neuropsychopharmacology* 1994; 10 (Suppl 3s): 31s.

Hong CJ, Chen JY, Chiu HJ, Sim CB. A double-blind comparative study of clozapine versus chlorpromazine on Chinese patients with treatment-refractory schizophrenia. *International Clinical Psychopharmacology* 1997; 12: 123-30

Honigfeld G, Patin J, Singer J. Clozapine: antipsychotic activity in treatment resistant schizophrenics. *Advances in Therapy* 1984; 1: 77-97

Høyberg OJ, Fensbo C, Remvig J, Lingjærde O, Sloth-Nielsen M, Salvesen I. Risperidone versus perphenazine in the treatment of chronic schizophrenic patients with acute exacerbations. *Acta Psychiatry Scand* 1993; 88: 395-402.

Huttunen MO, Piepponen T, Rantanen H, Larmo I, Hyholm R, Raitasuo V. Risperidone versus zuclopentixol in the treatment of acute schizophrenic episodes: a double-blind parallel-group trial. *Acta Pschiatrica Scandinavica* 1995; 91: 271-7

Itoh H, Miura S, Yagi G, Sakurai S, Ohtsuka N. Some methodological considerations for the clinical evaluation of neuroleptics – comparative effects of clozapine and haloperidol on schizophrenics. *Folia Psychiatrica et Neurologica Japonica* 1977; 31: 17-24

Kane JM, Honigfeld G, Singer J, Meltzer H and the Clozaril Collaborative Study Group. Clozapine for the treatment-resistant schizophrenic: results of a US multicenter trial. *Psychopharmacology* 1989; 99(Suppl): S60-63

Kane JM, Honigfeld G, Singer J, Meltzer H. Clozapine in treatment-resistant schizophrenics. *Psychopharmacology Bulletin* 1988; 24: 62-7.

Klieser E, Lehmann E, Kinzler E, Wurthmann C, Heinrich K. Randomized, double-blind, controlled trial of risperidone versus clozapine in patients with chronic schizophrenia. *Journal of Clinical Psychopharmacology* 1995; 15: (Suppl 1):45S-51S.

Kumra S, Frazier JA, Jacobsen LK, McKenna K, Gordon CT, Lenane MC et al. Childhood-onset schizophrenia: a double-blind clozapine-haloperidol comparison. *Arch Gen Psychiatry* 1996; 53: 1090-7

Léon CA. A 4-year follow-up of a controlled clinical trial. *Acta Psych Scand* 1979; 59: 471-80

M93-098. A double-blind, placebo-controlled, haldol-referenced study of the safety and efficacy of two doses of sertindole in schizophrenic patients. Illinois: Abbott Laboratories, 1995

M93-113. A double-blind, placebo-controlled, dose response comparison of the safety and efficacy of three doses of sertindole and three doses of haldol in schizophrenic patients. Illinois: Abbott Laboratories, 1995.

M93-132. A double blind comparison of sertindole and haldol: an assesment of the chronic safety, efficacy, quality of life and relapse in stable schizophrenic patients. Illinois: Abbott Laboratories, 1997

Mahmoud R, Engelhart L, Oster G, Ollendorf D and the Risperidone Outcome Study of Effectiveness (ROSE) Group. Psychiatric resource use under usual care conditions: what is the cost of risperidone's benefits? Toronto, Annual Meeting of the American Psychiatric Association, May 30- June 4, 1998b.

Mahmoud RA, Engelhart LM and the Risperidone Outcome Study of Effectiveness (ROSE) Group. Risperidone vs conventional antipsychotics in usual care: a prospective randomized effectiveness trial of outcomes for patients with schizophrenia and schizoaffective disorder. Glasgow: XXI CINP Congress July 12-16, Glasgow, Scotland, 1998a.

Mahmoud RA, Engelhart LM, Oster G, Stevens MC, Meredith C, Lee DM and the Risperidone Outcome Study of Effectiveness (ROSE) Group. Hawaii, Annual Meeting of the American College of Neuropsychopharmacology, 1997.

Marder SR, Meibach RC. Risperidone in the treatment of schizophrenia. *Am J Psychiatry* 1994; 151: 825-35

Mason P, Harrison G, Glazebrook C, Medley I, Dalkin T, and Croudace T. Characteristics of outcome in schizophrenia at 13 years. *British Journal of Psychiatry* 167 (5):596-603, 1995.

Meredith C, Mahmoud RA, Englehart LM, Ramirez L and the Risperidone Outcome Study of Effectiveness (ROSE) Group. Clinical and quality of life superiority of risperidone over conventional antipsychotics under usual care conditions: a prospective randomized trial in schizophrenia and schizoaffective disorder. Toronto, Annual Meeting of the American Psychiatric Association, May 30- June 4, 1998

Mesotten F. Clinical Research Report RIS-BEL-7; August 1991 (N 85163). Unpublished report from Janssen.

Min SK, Rhee CS, Kim CE, Kang DY. Risperidone versus haloperidol in the treatment of chronic schizophrenic patients: a parallel group double blind trial. *Yonsei Medical Journal* 1993; 34: 179-90

Möller HJ, Boyer P, Fleurot O, Rein W. Improvement of acute exacerbations of schizophrenia with amisulpride: a comparison with haloperidol. PROD-ASLP Study Group. *Psychopharmacology (Berl)* 1997; 132: 396-401

Peuch A, Fleurot O, Rein W and the Amisulpride Study Group. Amisulpride, an atypical antipsychotic in the treatment of acute episodes of schizophrenia: a dose-ranging study vs. haloperidol. *Acta Psychiatr Scand* 1998; 65-72.

Peuskens J on behalf of the Risperidone Study Group. Risperidone in the treatment of patients with chronic schizophrenia: a multi-national, multi-centre, double-blind, parallel-group study versus haloperidol. *British Journal of Psychiatry* 1995; 166: 712-26.

Peuskens J, Link CGG. A comparison of quetiapine and chlorpromazine in the treatment of schizophrenia. *Acta Psychiatr Scand* 1997; 96: 265-73

Rosenheck R, Cramer J, Zu W, Thomas J, Henderson W, Frisman L et al. A comparison of clozapine and haloperidol in hospitalized patients with refractory schizophrenia. *N Engl J Med* 1997; 337: 809-15.

Schönell H, Klieser E. Comparison of the effects of standardized doses of haloperidol and clozapine on systematic delusions of chronic schizophrenics. (Abstract) *Psychopharmacology* 1988;96(Suppl.):186.

Singer K, Law SK. A double-blind comparison of clozapine (leponex) and chlorpromazine in schizophrenia of acute symptomatology. *J Int Med Res* 1974; 2: 433-5

Tollefson GD, Beasley CM, Tran PV, Street JS, Krueger JA, Tamura RN et al. Olanzapine versus haloperidol in the treatment of schizophrenia and schizoaffective and schizophreniform disorders: Results from an international collaborative trial. *Am J Psychiatry* 1997a; 154: 457-65.

Tollefson GD, Beasley CM, Tamura RN, Tran PV Potvin JH. Blind, controlled, long-term study of the comparative incidence of treatment emergent tardive dyskinesia with olanzapine or haloperidol. *Am J Psychiatry* 1997b; 154: 1248-54.

Tran PV, Hamilton SH, Kuntz AJ, Potvin JH, Andersen SW, Beasley C, et al. Double blind comparison of olanzapine versus risperidone in the treatment of schizophrenia and other psychotic disorders. *J Clin Psychopharmacol* 1997; 17: 407-18.

Zimbroff DL, Kane JM, Tamminga CA, Daniel DG, Mack RJ, Wozniak PJ et al. Controlled, dose-response study of sertindole and haloperidol in the treatment of schizophrenia. *Am J Psychiatry* 1997; 154: 782-91

APPENDIX 2 DATA EXTRACTION FORM FOR ECONOMIC MODELS

General information and verification of study eligibility

Component of study	Study descriptive
Study (year)	
Interventions	
Design	

Study-specific information

Component of study	Study descriptive
Perspective of study	
Currency (year)	
Time horizon	
Clinical outcomes considered	
Outcome measure	
Target population of model	
Target setting of model	

Information relating to generalisability

Factual information of data sources

Data underlying the study	Study descriptive
What was the source of effectiveness data on the main clinical outcome?	
What was the population of trial underlying effect estimate (inclusion and exclusion criteria)?	
What was the source of the costs of the study?	
What was the source of resource use and treatment pattern?	

Checklist for generalisability: Costs

Q	Question for appraisal	Study score	Comment
1	Did the model attempt to reflect variation in different resource use patterns between health care environments nationally (e.g. varying hospitalisation rates and length of stay)?		
2	Did the model attempt to reflect variation in different resource use patterns between health care environments internationally		
3	Did the model attempt to reflect a range of costs relevant to a different health care environment nationally?		
4	Did the model attempt to reflect a range of costs relevant to different health care environments internationally?		

Checklist for generalisability: Effects

Q	Question for appraisal	Study score	Comment
5	Were the effect estimates based on epidemiological studies, pragmatic trials or meta-analyses?		
6	Were the effect estimates from the underlying studies moderated to reflect the target population of the model?		
7*	Did the model accommodate for difference in compliance rates between those observed in clinical trials and those likely in usual care?		
8	Were head to head comparisons synthesised?		
9	Were intermediate outputs of the model compared to external sources?		

* If compliance explored	Result	Comment
Was the outcome of the evaluation sensitive to variation in compliance rate?		

Checklist for generalisability: Results.

Q	Question for appraisal	Study score	Comment
10	Were the results compared to other relevant studies?		

APPENDIX 3

STUDIES IN THE REVIEW OF ECONOMIC MODELS

Ankjaer-Jensen A, Johnell O. Prevention of Osteoporosis: Cost effectiveness of different pharmaceutical treatments. *Osteoporosis International* 1996;6:265-275

Cheung AP, Wren BG. A cost-effectiveness analysis of hormone replacement therapy in the menopause. *The Medical Journal of Australia* 1992;156:312-316

Clark AP, Schuttinga JA. Targeted Estrogen/Progestogen replacement therapy for Osteoporosis: Calculation of health care cost savings. *Osteoporosis International* 1992;2:195-200

Daly E, Roche M, Barlow D, Gray A, McPherson K, Vessey M. HRT: AN analysis of benefits, risks and costs. *British Medical Bulletin* 1992;48(2)368-400

Daly E, Vessey MP, Barlow A, Gray K, McPherson K, Roche M. Hormone replacement therapy in a risk-benefit perspective. *Maturitas* 1996;23:247-259

Francis RM, Anderson FH, Torgerson DJ. A comparison of the effectiveness and cost of treatment for vertebral fractures in women. *British Journal of Rheumatology* 1995;34:1167-1171

Garton MJ, Cooper C, Reid D. Perimenopausal bone density screening – will it help prevent osteoporosis? *Maturitas* 1997;26:35-43

Geelhoed E, Harris A, Prince R. Cost-effectiveness analysis of hormone replacement therapy and lifestyle intervention for hip fracture. *Australian Journal of Public Health* 1994;18(2)153-60

Goddard M. The cost effectiveness of hormone replacement therapy: a review. Discussion paper nr 73, Centre for Health Economics, university of York 1990.

Jonsson B, Christiansen C, Johnell; O, Hedbrandt J. Cost effectiveness of fracture prevention in established osteoporosis. *Osteoporosis International* 1995;5:136-142

Office of Technology assessment (OTA) US Congress. Effectiveness and costs of osteoporosis screening and hormone replacement therapy, volume I: Cost effectiveness analysis. OTA-BP-H-160 (Washington, DC: US Government Printing Office, August 1995).

Rosner AJ, Grinna DT, Torrance GW, Bradley C, Adachi JD, Sebaldt RJ, Willison DJ. Cost effectiveness of multi-therapy treatment strategies in the prevention of vertebral fractures in postmenopausal women with osteoporosis. *Pharmacoeconomics* 1998;14(5)559-573

Torgerson DJ, Reid DM. Osteoporosis prevention through screening: will it be cost effective? *Baillieres Clinical Rheumatology* 1993;7(3)603-622.

Torgerson DJ, Kanis JA. Cost effectiveness of preventing hip fractures in the elderly population using vitamin D and calcium. *Quarterly Journal of Medicine* 1995;88:135-139

Torgerson D, Donaldson C, Reid D. Using economics to prioritise research: a case study of randomised trials for the prevention of hip fractures due to osteoporosis. *Journal of Health Services Research and Policy* 1996;1(3):141-146

Tosteson ANA, Rosenthal DI, Melton LJ, Weinstein MC. Cost effectiveness of screening perimenopausal white women for osteoporosis: bone densitometry and hormone replacement therapy. *Annals of Internal Medicine*. 1990;113(8):598-603.

Tosteson ANA, Weinstein MC. Cost-effectiveness of hormone replacement therapy after the menopause. *Baillieres clinical obstetrics and gynaecology* 1991;5:943-959

Visentin A, Ciravegna R, Fabris F. The cost per avoided hip fracture by osteoporosis treatment in Italy. *Maturitas* 1997;26:185-192

Weinstein MC. Estrogen use in postmenopausal women – costs, risks and benefits. *New England Journal of Medicine* 1980;303:308-16

Weinstein MC, Schiff I. Cost-effectiveness of hormone replacement therapy in the menopause. *Obstretical and gynaecological survey* 1983;38:445-454

Weinstein MC, Tosteson ANA, Cost-effectiveness of hormone replacement. *Annals of the New York Academy Science* 1990;592:162-173

APPENDIX 4

MEDIPLUS BIBLIOGRAPHY OF PUBLICATIONS

Received from IMS, May 2001

Title: The effects of age, body mass index, smoking and general health on the venous thromboembolism in users of combined oral contraceptives.

Type of article: Scientific study

Author: Nightingale AL, Lawrenson RA, Simpson EL, Williams TJ, MacRae KD, Farmer RD

Publication: Eur J Contracept Reprod Health Care. 2000 Dec;5(4):265-74

IMS database:UK MediPlus®

Title: Helicobacter pylori eradication therapy

Type of article: Abstract and poster

Author: Schröder-Bernhardi D, Dietlein G

Publication: 8. Jahrestagung Deutschen Arbeitsgemeinschaft für Epidemiologie (Hamburg), September 2000

IMS database:German MediPlus®

Title: Epidemiology and economic considerations of Helicobacter pylori infection

Type of article: Abstract and poster

Author: Schröder-Bernhardi D, Perez E, Dietlein G

Publication: Medical Infobahn for Europe. Proceedings of MIE2000 and GMDS2000. Technology and Informatics 77, 2000

IMS database:German MediPlus®

Title: Venous thromboembolism and combined oral contraceptives: does the type of progestogen make a difference?

Type of article: Scientific study

Author: Lawrenson R, Farmer R

Publication: Contraception. 2000 Aug 1;62(2 Suppl 1):S21-S28

IMS database:UK MediPlus®

Title: Fluoroquinolones and the risk of Achilles tendon disorders

Type of article: Poster

Author: van der Linden PD, Sturkenboom MC, Herings RM, Leufkens HG, Stricker BH

Publication: Number 333, 16th Intl Conf Pharmacoepidemiology (Barcelona), August 2000

IMS database:UK MediPlus®

Title: Selective serotonin reuptake inhibitor treatment in the UK: risk of relapse or recurrence of depression

Type of article: Scientific study

Author: Claxton AJ, Li Z, McKendrick J

Publication: Br J Psychiatry. 2000 Aug;177:163-8
IMS database: UK MediPlus®

Title: Drug utilisation research on the basis of a longitudinal patient database
Type of article: Abstract
Author: Schröder-Bernhardi D, Dietlein G
Publication: Intl J Clin Pharmacol Therapeut. 2000 Jul;38:364-372
IMS database: German MediPlus®

Title: Prescribing behavior of primary care physicians in diabetes therapy: effect of drug budgeting
Type of article: Scientific study
Author: Junger C, Rathmann W, Giani G
Publication: Dtsch Med Wochenschr. 2000 Feb 4;125(5):103-9
IMS database: German MediPlus

Title: Patterns of diagnosis and referral in women consulting for chronic pelvic pain in UK primary care
Type of article: Scientific study results
Author: Zondervan KT, Yudkin PL, Vessey MP, Dawes MG, Barlow DH, Kennedy SH
Publication: Br J Obstet Gynaecol. 1999 Nov;106(11):1156-61
IMS Database: UK MediPlus®

Title: Prevalence and incidence of chronic pelvic pain in primary care: evidence from a national general practice database
Type of article: Scientific study results
Author: Zondervan KT, Yudkin PL, Vessey MP, Dawes MG, Barlow DH, Kennedy SH
Publication: Br J Obstet Gynaecol. 1999 Nov;106(11):1149-55
IMS Database: UK MediPlus®

Title: Cardiovascular drug prescriptions and risk of depression in diabetic patients
Type of article: Scientific study
Author: Rathmann W, Haastert B, Roseman J, Giani G
Publication: J Clin Epidemiol. 1999 Nov;52(11):1103-9
IMS database: German MediPlus®

Title: Migraine consultation patterns in primary care. Results from the PCAOM study 1994-96.
Type of article: Scientific study results
Author: Krobot KJ, Schroder-Bernhardi D, Pfaffenrath V
Publication: Cephalalgia. 1999 Nov;19(9):831-40

IMS database: German MediPlus®

Title: Coronary heart disease in women with diabetes: positive association with past hysterectomy and possible benefits of hormone replacement therapy

Type of article: Poster

Author: Lawrenson R, Leydon GM, Newson RB, Feher MD

Publication: Number 146, 15th Intl Conf Pharmacoepidemiology (Boston) August 1999

IMS database: UK MediPlus®

Title: Influence of hospitals on the prescription behaviour of doctors in office

Type of article: Abstract

Author: Schröder-Bernhardi D, Dietlein G

Publication: Eur J Clin Pharmacol. 1999 Aug;55:A54, No.37

IMS database: German MediPlus

Title: Oral contraceptives and venous thromboembolic disease. Analyses of the UK General Practice Research Database and the UK Mediplus database.

Type of article: Review

Author: Farmer RD, Lawrenson RA, Todd JC, Williams TJ, MacRae K

Publication: Hum Reprod Update. 1999 Nov-Dec;5(6):688-706

IMS database: UK MediPlus®

Title: The treatment of depression in the UK general practice; selective serotonin reuptake inhibitors and tricyclic antidepressants compared

Type of article: Poster

Author: Lawrenson R, Tyer F, Farmer R

Publication: Number 144, 15th Intl Conf Pharmacoepidemiology (Boston), August 1999

IMS database: UK MediPlus®

Title: Comparison of safety profiles of acetaminophen dosing regimens (3g and 4g daily) in patients with painful chronic rheumatoid diseases

Type of article: Poster

Author: Vesque D, Dardennes J, Ganry H, Pelc A, Pruvot F, Schmidely N

Publication: Number 126, 15th Intl Conf Pharmacoepidemiology (Boston), August 1999

IMS database: UK MediPlus®

Title: Patterns of contraception in UK women Type 1 diabetes: a GP database study

Type of article: Poster

Author: Lawrenson R, Williams T, Leydon G, Newson R

Publication: Number 143, 15th Intl Conf Pharmacoepidemiology (Boston), August 1999

IMS database:UK MediPlus®

Title: Venous thromboembolic disease and combined oral contraceptives: A re-analysis of the MediPlus database

Type of article: Scientific study results

Author: Todd J, Lawrenson R, Farmer RD, Williams TJ, Leydon GM

Publication: Hum Reprod. 1999 Jun;14(6):1500-5

IMS database:UK MediPlus®

Title: High prevalence of comorbid conditions in type 2 diabetes (DM2): UK MediPlus

Type of article: Poster

Author: Lydick E, Gaskin M, Howard T

Publication: Intl Conf Diabetes Cardiovasc Disease, Winnipeg, June 1999

IMS database: UK MediPlus®

Title: Coronary heart disease in women with diabetes. Positive association with past hysterectomy and possible benefits of hormone replacement therapy

Type of article: Letter

Author: Lawrenson R, Leydon GM, Newson RB, Feher MD

Publication: Diabetes Care. 1999 May;22(5):856-7

IMS database:UK MediPlus®

Title: Patterns of contraception in UK women with Type 1 diabetes mellitus: a GP database study

Type of article: Scientific study

Author: Lawrenson R, Williams T, Leydon G, Newson R

Publication: Diabet Med. 1999 May;16(5):395-9

IMS database:UK MediPlus®

Title: A UK primary care database (UKPCD)-UK *MediPlus* is a readily available research tool for determining the prevalence of patients with a history of obesity or overweight, their treatment, and associated co-morbidities

Type of article: Poster

Author: Starling A, Gaskin M, Thorington N, Harrison P

Publication: 4th Intl Soc Pharmacoeconomics Outcomes Res (Washington), May 1999

IMS database:UK MediPlus®

Title: Migraine prescription density and recommendations. Results of the PCAOM study

Type of article: Scientific study results

Author: Krobot KJ, Steinberg HW, Pfaffenrath V

Publication: Cephalalgia. 1999 Jun;19(5):511-9

IMS database: German MediPlus

Title: Clinical information for research; the use of general practice databases

Type of article: Review

Author: Lawrenson R, Williams T, Farmer R

Publication: J Public Health Med. 1999 Sep;21(3):299-304

IMS Database: UK MediPlus®

Title: Patterns of antidepressant prescription in general practice: lessons from the MediPlus database

Type of article: Scientific study results

Author: Lawrenson R

Publication: Primary Care Psychiatry. 1999 Jan;5:3-7

IMS database: UK MediPlus®

Title: Drug prescriptions and costs in the treatment of diabetic polyneuropathy

Type of article: Scientific study results

Author: Rathmann W, Haastert B, Giani G

Publication: Dtsch med Wschr. 1999 Jun;124:681-6

IMS database: German MediPlus®

Title: A UK primary care database (UKPCD-UK *MediPlus*) as a readily available research tool for determining the incidence and prevalence of multiple sclerosis, its treatment, prescription costs and referrals in UK practice

Type of article: Poster

Author: Carmichael R, Grigas E, Harrison P, McDermott J

Publication: Intl Soc Pharmacoeconomics Outcomes Res (Cologne) December 1998

IMS database: UK MediPlus®

Title: Selective serotonin reuptake inhibitor treatment in the UK: and the risk of relapse or recurrence of depression

Type of article: Poster

Author: Li Z, Claxton AJ, McKendrick J.

Publication: ECNP, Paris, November 1998

IMS database: UK MediPlus®

Title: High prevalence of comorbid conditions in type 2 diabetes (DM2): UK MediPlus

Type of article: Poster

Author: Lydick E, Gaskin M, Howard T

Publication: British Diabetic Association, Harrogate, October 1998

IMS database: UK MediPlus®

Title: Prevalence and incidence of chronic pelvic pain in UK primary care
Type of article: Scientific study results
Author: Zondervan KT, Yudkin PL, Vessey MP, Dawes MG, Barlow DH, Kennedy SH
Publication: 42nd Ann Sci Mtg Society of Soc Social Medicine (Cardiff), September 1998
IMS database: UK MediPlus®

Title: The direct costs to the NHS of discontinuing and switching prescriptions for hypertension
Type of article: Scientific study results
Author: Hughes D, McGuire A
Publication: J Hum Hypertens. 1998 Aug; 12(8):533-7
IMS database: UK MediPlus®

Title: A UK Primary Care Database (UKPCD) - UK MediPlus is a readily available research tool for determining the incidence and prevalence of epilepsy (all forms) and their first-line treatment in UK practice
Type of article: Abstract
Author: Harrison PJ, McDermott J, Thompson M
Publication: Poster Abstract Number 74. 14th Intl Conf Pharmacoepidemiology (Berlin) August 1998
IMS Database: UK MediPlus®

Title: Anti-diabetic drugs and the risk of Stevens-Johnson syndrome and toxic epidermal necrolysis
Type of article: Abstract
Author: Martinez C, Schlingmann J, Suissa S, Schopf E, Mockenhaupt M
Publication: Poster Abstract Number 78. 14th Intl Conf Pharmacoepidemiology (Berlin) August 1998
IMS Database: German MediPlus

Title: Chronic pelvic pain in UK primary care: patterns of diagnosis
Type of article: Scientific study results
Author: Zondervan KT, Yudkin PL, Vessey MP, Dawes MG, Barlow DH, Kennedy SH
Publication: Proc World Congr Endometriosis (Canada) July 1998
IMS Database: UK MediPlus®

Title: A study of cardiovascular disease and antidepressants on a computerised general practice database
Type of article: Poster
Author: Tyrer F, Lawrenson RA, Newson R, Farmer R
Publication: Conf Intl de Neuropsychopharm (Glasgow) July 1998

IMS Database: UK MediPlus®

Title: Tiaprofenic acid and severe cystitis
Type of article: Scientific study results
Author: Brown EG, Waller PC, Sallie BA
Publication: Postgrad Med J. 1998 Jul;74:443-6
IMS Database: UK MediPlus®

Title: Cost analysis of palliative care for terminally ill cancer patients in the UK after switching from weak to strong opioids. Palliative Care Advisory Committee
Type of article: Scientific study results
Author: Guest JF, Hart WM, Cookson RF
Publication: Pharmacoeconomics. 1998 Sep; 14(3):258-97
IMS database: UK MediPlus®

Title: Oral contraceptives and venous thromboembolic disease: the findings from database studies in the United Kingdom and Germany
Type of article: Scientific study results
Author: Farmer RD, Lawrenson RA
Publication: Am J Obstet Gynecol. 1998 Mar; 179:S78-86
IMS database: UK MediPlus®

Title: Cost effectiveness of treatment for gastro-oesophageal reflux disease in clinical practice: a clinical database analysis
Type of article: Scientific study results
Author: Eggleston A, Wigerinck A, Huijghebaert S, Dubois D, Haycox A
Publication: Gut. 1998 Jan; 42(1):13-6
IMS database: UK MediPlus®

Title: Modelling different approaches to the management of upper gastrointestinal disease
Type of article: Scientific study results
Author: Duggan AK
Publication: Pharmacoeconomics. 1998; 14 (Suppl 2):25-37
IMS database: UK MediPlus®

Title: The path of least resistance
Type of article: Data citation
Author: Standing Medical Advisory Committee (SMAC) Guidelines
Publication: Department of Health
IMS Database: UK MediPlus®

Title: Prescribing of antidepressant in cardiovascular disease: a study using a computerised general practice data base

Type of article: Scientific study results

Author: Tyrer F, Lawrenson RA, MacRae K, Farmer RD

Publication: Int J Gen Pract and Prim Care. June 1998

IMS database: UK MediPlus®

Title: Coronary heart disease in women with diabetes: the influence of diabetic treatment, hysterectomy and use of hormone replacement therapy

Type of article: Abstract

Author: Lawrenson RA, Leydon GM, Newson RB, Feher MD

Publication: Rept Second Conf Prim Care Diabetes UK, Brit Diabetic Assoc. 1998;17

IMS Database: UK MediPlus®

Title: Patterns of contraception in young UK women with insulin dependent diabetes mellitus (IDDM)

Type of article: Abstract

Author: Leydon GM, Lawrenson RA, Williams T, Newson RB, Feher MD

Publication: Rept Second Conf Prim Care Diabetes UK, Brit Diabetic Assoc. 1998;19

IMS Database: UK MediPlus®

Title: What is the evidence on venous disease

Type of article: Abstract

Author: Lawrenson RA

Publication: Eur J Contracept Reprod Health Care. 1998;3(Suppl 1):73

IMS Database: UK MediPlus®

Title: Age specific prevalences do not suggest association with in utero exposure

Type of article: Letter

Author: Lawrenson RA, Farmer R

Publication: BMJ. 1998;316(7146):1746

IMS Database: UK MediPlus®

Title: Postmarketing surveillance of adverse drug reactions: a correlational study approach using multiple data sources

Type of article: Scientific study results

Author: Rathmann W, Haastert B, Delling B, Gries FA, Giani G

Publication: Pharmacoepidemiol Drug Safety. 1998;7:51-7

IMS database: German MediPlus

Title: Assessment of effectiveness of therapy using retrospective population data

Type of article: Poster

Author: Sverdlov LS, Jones JK, Dixon RB
Publication: Presented at ASCPT conference (New Orleans) March 1998
IMS database: UK MediPlus®

Title: Impact of potassium chloride on renal function in hypertensive patients
Type of article: Poster
Author: Jones JK, Sverdlov LS, Carlile OL, Dixon RB, Shannon MW, Hou EW
Publication: Presented at ASCPT conference (New Orleans) March 1998
IMS database: UK MediPlus®

Title: Neuroleptic use and indications in UK primary care database
Type of article: Poster
Author: Jones JK, Miwa LJ, Staffa JA, Pathiyal A, Schwamlein C
Publication: Presented at ASCPT conference (New Orleans) March 1998
IMS database: UK MediPlus®

Title: The risks of venous thromboembolic disease among German women using oral contraceptives: a database study
Type of article: Scientific study results
Author: Farmer RD, Todd JC, Lewis MA, MacRae KD, Williams TJ
Publication: Contraception. 1998 Feb; 57(2):67-70
IMS database: German MediPlus

Title: Oral contraception was not associated with venous thromboembolic disease in recent study
Type of article: Letter
Author: Farmer RDT, Todd J-C, MacRae KD, Williams TJ, Lewis MA
Publication: BMJ. 1998;316:1090-1
IMS database: German MediPlus

Title: Do women with diabetes receive hormone replacement therapy?
Type of article: Scientific study results
Author: Lawrenson RA, Newson RB, Feher MD
Publication: Pract Diabetes Int. 1998 Mar; 15:71-2
IMS database: UK MediPlus®

Title: Characteristics of practices contributing to the MediPlus database and the implications for its use in epidemiological research
Type of article: Scientific study results
Author: Lawrenson RA, Coles G, Walton K, Farmer RDT
Publication: J Informatics in Prim Care 1998 May;14-8
IMS database: UK MediPlus®

Title: The pivotal role of post-licensing activities
Type of article: Data citation
Author: Wood S, Harman RJ
Publication: The Regulatory Review, July. 1998 May;1:3-8
IMS database: UK MediPlus®

Title: New drug introductions—The issues raised by Viagra
Type of article: Data citation
Author: Stephens P
Publication: Medi-Pharm Opportunities. June 1998;8
IMS database: UK MediPlus®

Title: Prescription drug use and costs among diabetic patients in primary health care practices in Germany
Type of article: Scientific study results
Author: Rathmann W, Haastert B, Roseman J, Gries FA, Giani G
Publication: Diabetes Care. 1998 Mar;21:389-97
IMS database: German MediPlus®

Title: Population-based study of risk of venous thromboembolism associated with various oral contraceptives
Type of article: Scientific study results
Author: Farmer RDT, Lawrenson RA, Thompson CR, Kennedy JG, Hambleton IR
Publication: Lancet. 1997 Jan 11; 349(9045):83-8
IMS database: UK MediPlus®

Title: Oral contraceptives and venous thromboembolism
Type of article: Letter
Author: Farmer RDT, Lawrenson RA
Publication: Lancet. 1997;349:733
IMS database: UK MediPlus®

Title: Oral contraceptives and venous thromboembolism
Type of article: Letter
Author: Lawrenson R, Farmer R
Publication: N Z Med J. 1997;110:63
IMS database: UK MediPlus®

Title: Examining the management of depression in primary care using the MediPlus® database in the Netherlands
Type of article: Abstract
Author: Pathiyal A, Hylan T, Quick R, Jones JK
Publication: Pharmacoepidemiol Drug Safety. 1997;6(Suppl 2):S88
IMS Database: Netherlands MediPlus

Title: Examining the management of depression in primary care using the MediPlus database in the Netherlands

Type of article: Scientific study results

Author: Pathiyal A, Hylan TR, Quick R, Jones JK

Publication: Clin Ther. 1997 Jul-Aug;19(4):798-810

IMS Database: Netherlands MediPlus

Title: Antidepressant use patterns in the naturalistic GP setting in the Netherlands

Type of article: Poster

Author: Pathiyal A, Hylan TR, Quick R, Jones JK

Publication: Presented at ECNP Conference (Vienna) 1997

IMS Database: Netherlands MediPlus

Title: Pregnancy following oral contraception

Type of article: Poster

Author: Farmer RDT, Todd J-C

Publication: Poster abstract Number 317. 13th Intl Conf Pharmacoepidemiology, (Orlando) August 1997

IMS database: UK MediPlus®

Title: The validity of general practice computer data

Type of article: Poster

Author: Farmer RDT, Lawrenson RA, Coles GLA

Publication: Poster abstract Number 318. 13th Intl Conf Pharmacoepidemiology, (Orlando) August 1997

IMS database: UK MediPlus®

Title: Hormone replacement therapy prescriptions for women with diabetes

Type of article: Abstract

Author: Lawrenson RA, Feher M

Publication: Diabet Med. 1997;14(Suppl):S40

IMS database: UK MediPlus®

Title: Antibacterial prescribing for respiratory symptoms in health care. Time to put the brakes on

Type of article: Abstract

Author: Malik F, Gaskin M, Davey PG

Publication: Pharmacoepidemiol Drug Safety. 1996;5(Suppl 1):S34

IMS database: UK MediPlus®

Title: The cost of treatment of uncomplicated cystitis: transfer of data from a clinical trial in the USA based on an epidemiological study in the UK

Type of article: Abstract
Author: Malik F, Gaskin M, Davey PG
Publication: Pharmacoepidemiol Drug Safety. 1996;5(Suppl 1):S34
IMS Database: UK MediPlus®

Title: The scale of repeat prescribing
Type of article: Scientific study results
Author: Harris CM, Dajda R
Publication: Br J Gen Pract. 1996 Nov;46(412):649-53
IMS database: UK MediPlus®

Title: Cough due to ACE inhibitors: a case-control study using automated general practice data
Type of article: Scientific study results
Author: Visser LE, Vlug AE, van der Lei J, Stricker BH Ch
Publication: Eur J Clin Pharmacol. 1996 Jun; 49:439-44
IMS database: Netherlands MediPlus

Title: Specific therapeutic group age-sex related prescribing units (STAR-PUs): weightings for analysing general practices' prescribing in England
Type of article: Scientific study results
Author: Lloyd DC, Harris CM, Roberts DJ
Publication: BMJ. 1995 Oct 14;311(7011):991-4
IMS database: UK MediPlus®

Title: Discontinuation of, and changes in treatment after start of new courses of antihypertensive drugs: a study of United Kingdom population
Type of article: Scientific study results
Author: Jones JK, Gorkin J, Lian JF, Staffa JA, Fletcher AP
Publication: BMJ. 1995 Jul 29;311(7000):293-5
IMS database: UK MediPlus®

Title: Treating menorrhagia in primary care. An overview of drug trials and a survey of prescribing practice
Type of article: Data citation
Author: Coulter A, Kelland J, Peto V, Rees M
Publication: Int J Technol Assess Health Care. 1995 Summer;11(3):456-71
IMS database: UK MediPlus®

Title: Approaching a century after: morbidity and potentially drug induced disease
Type of article: Review
Author: Fletcher AP
Publication: Adverse Drug React Toxicol Rev. 1995 Spring;14(1):45-64
IMS database: UK MediPlus®

APPENDIX 5

ANALYTICAL VARIABLES TO THE OBSERVATIONAL STUDY

APPENDIX 5.1

DEFINITION OF STATISTICAL VARIABLES

The data extracted according to the study design document will be exported into five data files. The first file will contain information on demographic details, the second on therapy pattern of the intervention drugs, the third on resources consumed through concomitant medication, the fourth on administrative activities undertaken and the fifth file will contain information on the clinical events of the patients. Each datafile will contain 1 record per patient.

General and demographic variables (FILE 1, "PDEMOG")

A datafile containing patient age, marital status, previous co-morbidity etc. This will include clinical Problem and Notes (ICD 10 level 1 and 2 and Read codes) and date on which a problem was entered. Previous clinical events, such as occurrence of fractures, cancers and cardiovascular events and the various forms of fractures will be included in this file.

Var. name	Label	Definition and creation
PATNR	Patient ID	Unique identifier for a study subject as given in the source data. Will be used to link source data files.
PATCO	Study cohort	1996, 1997 or 1998
INXDRG	Index drug	The study drug prescribed to a study subject on the index date. The variable categories will be named as following: EVISTA, FOSAMAX, KLIOFEM, LIVIAL, DIDRONEL, PREMIQUE, BISCNTR, HRTCNTR
INXDIAG	Diagnosis for index drug	The diagnosis under which the index drugs was prescribed. Osteoporosis: OSTEO Other: OTHER If no diagnosis recorded, then missing value.
PATAGE	Patient age in index year	The age of a patient as of the year of the index date, computed from patient's age as of December 1999
PATRES	Patient residential area	17 areas in the source data set according to the former indexing of Regional Health Authorities.

Var. name	Label	Definition and creation
PRXID	Practice ID number	Unique identifier for a practice given in the source data.
PRXPAT	Number of patients in the practice	Number of patients enrolled in the practice
PRXGP	Practice size	Size of the practice in terms of number of GPs
GPNUM	GP's ID number	The sequential number in MediPlus database, unique for each GP*
GPSEX	GP's gender	MALE / FEMALE The gender of the GP who issued the index prescription
GPAGE	GP's age	Doctor's age in the index year
PRXFUND	Practice previous fundholding	FUND = The practice was previously a fundholding practice, NOFUND = The practice was not previously a fundholding practice.

* In those cases where two GPs independently issued two prescriptions of the same drug at the same day, and one of the prescriptions was linked to a problem of osteoporosis, then the prescription linked to osteoporosis will have priority. If neither of the two prescriptions were issued in osteoporosis then the patients GP will be selected randomly from the two.

Data on the following clinical status entered in the patient records within five years prior to index prescription will be recorded. In those cases where these details are recorded repeatedly, the value entered closer to the index date will be included in the file. If BMI is recorded in the database then use this direct recording for the variable PATBMI. If BMD is not given directly but weight and height is given, then MBD will be derived from these two.

Var. name	Label	Definition and creation
PATHE	Patient's height	Patient's height (if recorded) in centimetres
PATWE	Patient's weight	Patient's weight (if recorded) in kilograms
PATBMI	BMI	Patients BMI Patient's body mass index (when recorded)
PATSMOK	Smoking status	Patient registered as smoker = 1 Patient registered as non-smoker = 0
PATBMD	Bone Mass Density	Patients Bone Mass Density

The following clinical events may affect outcome variables in this study and therefore this information will be collected on the patients. The recording of these will allow controlling for such confounders in the analysis. Information prevalent in the Problem Headings or in the Notes in the patient records **5 years** before or at the index date will be recorded.

The four variables concerning fractures are a special case in that GPs may record a fracture on more than one occasion. Therefore the following rule was developed: If a new fracture is recorded in Read within a 3 month period following another recording, then it is likely that the second fracture is the same as the first and this will consequently be ignored. Counting for new fractures will start after month 3. This rule does not apply to hip, wrist or vertebral fractures, as the recording of these is likely to be site-specific. Lists of the detailed Read codes characterising each clinical state can be found in appendix 1B.

Var. name	Label	Definition and creation
PRHIP	Previous hip fracture	Note or problem of hip fracture entered in patient record before index date. 1=yes, 0=no
PRWRIST	Previous wrist fracture	Note or problem of wrist fracture entered in patient record before index date. 1=yes, 0=no
PRVERT	Previous vertebral fracture	Note or problem of vertebral fracture entered in patient record before index date. 1=yes, 0=no
PROTH	Previous other fracture	Note or problem of any other fracture entered in patient record before index date. 1=yes, 0=no
PRANY	Any fracture	Any of the above fractures. 1=yes, 0=no
PRHIPN	Number of previous hip fractures	Number of hip fractures
PRWRISTN	Number of previous wrist fractures	Number of wrist fractures
PRVERTN	Number of previous vertebral fractures	Number of vertebral fractures
PROTHN	Number of any other previous fractures	Total number of fractures
PRANYN	Number of total fractures	The total of the above fractures, i.e. all fractures recorded in the patient record
HOST	History of osteoporosis	There is mentioning of an osteoporosis-related diagnosis in the patient record prior to index date (Including problem or note of family history) as defined in appendix 1D
PRHYST	Previous hysterectomy	Note or problem of hysterectomy entered in patient record before index date. 1 = yes, 0 = no
HBRCAN	History of breast cancer	There is mentioning of a breast cancer related diagnosis in the patient record prior to index date (Including problem or note of family history). 1 = yes, 0 = no
HOVCAN	History of ovarian cancer	There is mentioning of an ovarian cancer-related diagnosis in the patient record prior to index date (Including problem or note of family history). 1 = yes, 0 = no
HCHD	History of coronary heart disease	There is mentioning of coronary heart disease in the patient record prior to index date (Including problem or note of family history). 1 = yes, 0 = no
HSTROKE	History of stroke	There is mentioning of stroke in the patient record prior to index date (Including problem or note of family history). 1 = yes, 0 = no
HSTROKEN	Number of strokes	There is mentioning of stroke in the patient record prior to index date (Including problem or note of family history). Entered in the file as number of stroke diagnoses
HMI	History of myocardial infarction	There is mentioning of a myocardial infarction-related diagnosis in the patient record prior to index date (Including problem or note of family history). 1 = yes, 0 = no
HYP	Prior Hypertension	There is mentioning of a Hypertensive disease related diagnosis in the patient record prior to index date (Including problem or note of family history). 1 = yes, 0 = no
HCEREBR	Prior Cerebrovascular disease	There is mentioning of a Cerebrovascular disease related diagnosis in the patient record prior to index date. 1 = yes, 0 = no

Information on the following conditions prevalent in the Problem Headings or in the Notes in the patient records **9 months** before or at the index date will be recorded.

Var. name	Label	Definition and creation
HVCS	Prior vertebral column syndrome	There is mentioning of vertebral column syndrome in the patient record prior to index date
HRESP	Prior Respiratory system disease	There is mentioning of any Respiratory system disease related diagnosis in the patient record prior to index date
HPULM	Prior Pulmonary disease	There is mentioning of Chronic obstructive pulmonary disease in the patient record prior to index date.
HARTH	Prior Arthritis	There is mentioning of Rheumatoid Arthritis / arthrosis related diagnosis in the patient record prior to index date.
HRHEUM	Prior Rheumatism	There is mentioning of Nonarticular rheumatism related diagnosis in the patient record prior to index date.
HASTHMA	Prior Asthma	There is mentioning of Asthma or Asthma-related diagnosis in the patient record prior to index date.
HOES	Prior GI disease	There is mentioning of Oesophag / stomach / duoden disease diagnosis in the patient record prior to index date.

Therapy progress (FILE 2, “THPROGR”)

This file contains information on the progress of osteoporosis drug therapy. A decision-tree that shows the flow of patients initiating therapy on study drugs and then moving on to stopping therapy, swapping therapy or continuing therapy will be developed. The variables will enable the analysis of treatment patterns such as attrition rates, rate of swapping in the different cohorts and therapy cost incurred over the course of the treatment. The data are recorded on patient level and will enable the development of attrition curves and potential migration towards other any drugs in particular. Definitions were:

Therapy Stop: No further study drug issued in the 6 months following the end date of the duration of the last prescription of drug.

Switch: New study drug (or other drug in osteoporosis) issued within the time from the period covered by the last prescription of the study drug to the end of a 6 month period following the coverage of the last prescription.

Failure: Switch or stop

Therapy Break: A new prescription issued within 6 months after the duration of the last prescription of the study drug. Will be computed from the variables in the dataset.

Var. name	Label	Definition and creation
INDEXDAT	Index date	Date of the initial prescription of index drug
STOPDAT	Last day of index therapy	Date at which the coverage of the last prescription of the first treatment episode with index drug ended
INXCOST	Cost of index therapy	Total cost of the index therapy until failure date
FOSDATE	Date, Fosamax second drug	Date at which the patients initiates therapy with Fosamax as second drug
FOSCOST	Cost of Fosamax second drug	Total cost of Fosamax as second drug
DIDDATE	Date, Didronel second drug	Date at which the patients initiates therapy with Didronel as second drug
DIDCOST	Cost of Didronel second drug	Total cost of Didronel as second drug
LIVDATE	Date, Livial second drug	Date at which the patients initiates therapy with Livial as second drug
LIVCOST	Cost of Livial second drug	Total cost of Livial as second drug
KLIDATE	Date, Kliofem second drug	Date at which the patients initiates therapy with Kliofem as second drug
KLICOST	Cost of Kliofem second drug	Total cost of Kliofem as second drug
PREDATE	Date, Premique second drug	Date at which the patients initiates therapy with Premique as second drug
PRECOST	Cost of Premique second drug	Total cost of Premique as second drug
HRTDATE	Date, other HRT second drug	Date at which the patients initiates therapy with other HRT as second drug (see appendix 1D)
HRTCOST	Cost of other HRT second drug	Total cost of other HRT as second drug
BISDATE	Date, other Bisphosphonate second drug	Date at which the patients initiates therapy with other bisphosphonate therapy as second drug (see appendix 1D)
BISCOST	Cost of Bisphosphonate second drug	Total cost of other bisphosphonate as second drug
EVIDATE	Date, Evista second drug	Date at which the patients initiate therapy with Evista
EVICOST	Cost of Evista second drug	Total cost of Evista as second drug
OSTDATE	Date, other osteoporosis drug as second drug	Date at which the patients initiate therapy with other drug with an impact in osteoporosis
OSTCOST	Cost of other osteoporosis drug as second drug	Total cost of other drug in osteoporosis as second drug

Concomitant medication (FILE 3, “CONCOM”)

This file includes concomitant prescribing over the 1st, 2nd and 3rd year, respectively, following index prescription as well as patient-level data on prescriptions issued to the patients in the 6 months period before the index date. The file will include variables indicating whether each prescription was linked to an osteoporosis - related problem or not (appendix 1D), and the cost of the concomitant prescriptions.

Variables have a _0_5, 1, 2, or 3 added to identify the period of 6 months prior, 1, 2, or 3 years after the index date.

Var. name	Definition and creation
A02A_0_5	Cost of prescriptions of antacids issued in the six months prior to inclusion in the study.
A02A1	Cost of prescriptions of antacids issued in the first year after inclusion in the study.
A02A2	Cost of prescriptions of antacids issued in the second year after inclusion in the study.
A02A3	Cost of prescriptions of antacids issued in the third year after inclusion in the study.
A02B_0_5	Cost of prescriptions of H2 antagonists and proton pump inhibitors issued in the six months prior to inclusion in the study
A02B1	Cost of prescriptions of H2 antagonists and proton pump inhibitors issued in the first year after inclusion in the study
A02B2	Cost of prescriptions of H2 antagonists and proton pump inhibitors issued in the second year after inclusion in the study
A02B3	Cost of prescriptions of H2 antagonists and proton pump inhibitors issued in the third year after inclusion in the study
A2B1_0_5	Cost of prescriptions of H2 antagonists issued in the six months prior to inclusion in the study.
A2B11	Cost of prescriptions of H2 antagonists issued in the first year after inclusion in the study.
A2B12	Cost of prescriptions of H2 antagonists issued in the second year after inclusion in the study.
A2B13	Cost of prescriptions of H2 antagonists issued in the third year after inclusion in the study.
A2B2_0_5	Cost of prescriptions of Proton pump inhibitors issued in the six months prior to inclusion in the study.
A2B21	Cost of prescriptions of Proton pump inhibitors issued in the first year after inclusion in the study.
A2B22	Cost of prescriptions of Proton pump inhibitors issued in the second year after inclusion in the study.
A2B23	Cost of prescriptions of Proton pump inhibitors issued in the third year after inclusion in the study.
A11_0_5	Cost of prescriptions of Multivitamins and minerals (ATC A11A and A11B only) issued in the six months prior to inclusion in the study.
A111	Cost of prescriptions of Multivitamins and minerals issued in the first year after inclusion in the study.
A112	Cost of prescriptions of Multivitamins and minerals issued in the second year after inclusion in the study.
A113	Cost of prescriptions of Multivitamins and minerals issued in the third year after inclusion in the study.
A12A_0_5	Cost of prescriptions of Mineral supplements (Calcium, vitamin C and D) issued in the six months prior to inclusion in the study.
A12A1	Cost of prescriptions of Mineral supplements issued in the first year after inclusion in the study.
A12A2	Cost of prescriptions of Mineral supplements issued in the second year after inclusion in the study.
A12A3	Cost of prescriptions of Mineral supplements issued in the third year after inclusion in the study.
A06A_0_5	Cost of prescriptions of Laxatives issued in the six months prior to inclusion in the study.
A06A1	Cost of prescriptions of Laxatives issued in the first year after inclusion in the study.
A06A2	Cost of prescriptions of Laxatives issued in the second year after inclusion in the study.
A06A3	Cost of prescriptions of Laxatives issued in the third year after inclusion in the study.
G03_0_5	Cost of prescriptions of Sex hormones issued in the six months prior to inclusion in the study.
G031	Cost of prescriptions of Sex hormones issued in the first year after inclusion in the study.
G032	Cost of prescriptions of Sex hormones issued in the second year after inclusion in the study.
G033	Cost of prescriptions of Sex hormones issued in the third year after inclusion in the study.
H02A_0_5	Cost of prescriptions of Corticosteroids issued in the six months prior to inclusion in the study.

H02A1	Cost of prescriptions of Corticosteroids issued in the first year after inclusion in the study.
H02A2	Cost of prescriptions of Corticosteroids issued in the second year after inclusion in the study.
H02A3	Cost of prescriptions of Corticosteroids issued in the third year after inclusion in the study.
M01_0_5	Cost of prescriptions of Antirheumatic agents issued in the six months prior to inclusion in the study.
M011	Cost of prescriptions of Antirheumatic agents issued in the first year after inclusion in the study.
M012	Cost of prescriptions of Antirheumatic agents issued in the second year after inclusion in the study.
M013	Cost of prescriptions of Antirheumatic agents issued in the third year after inclusion in the study.
N02_0_5	Cost of prescriptions of Non-narcotic analgesics (N02A and N02B only) issued in the six months prior to inclusion in the study.
N021	Cost of prescriptions of Non-narcotic analgesics issued in the first year after inclusion in the study.
N022	Cost of prescriptions of Non-narcotic analgesics issued in the second year after inclusion in the study.
N023	Cost of prescriptions of Non-narcotic analgesics issued in the third year after inclusion in the study.
R03_0_5	Cost of prescriptions of Bronchodilators issued in the six months prior to inclusion in the study.
R031	Cost of prescriptions of Bronchodilators issued in the first year after inclusion in the study.
R032	Cost of prescriptions of Bronchodilators issued in the second year after inclusion in the study.
R033	Cost of prescriptions of Bronchodilators issued in the third year after inclusion in the study.

Administrative activity (FILE 4, "ADMIN")

Administrative Read codes indicating GP activity (dates entered in patient record) and referrals to secondary health care.

Variables have a _0_5, 1, 2, or 3 added to identify the period of 6 months prior, 1, 2, or 3 years after the index date.

Var. name	Definition and creation
GPVS_0_5	GP visits as estimated at the MediPlus terminal in the six months prior to inclusion in the study.
GPVS1	GP visits as estimated at the MediPlus terminal in the first year after inclusion in the study.
GPVS2	GP visits as estimated at the MediPlus terminal in the second year after inclusion in the study.
GPVS3	GP visits as estimated at the MediPlus terminal in the third year after inclusion in the study.
ENTN_0_5	Number of dates of entries in the patient record in the six months prior to inclusion in the study.
ENTN1	Number of dates of entries in the patient record in the first year after inclusion in the study.
ENTN2	Number of dates of entries in the patient record in the second year after inclusion in the study.
ENTN3	Number of dates of entries in the patient record in the third year after inclusion in the study.
NURS_0_5	Number of recordings of patients seen by practice nurse in the six months prior to inclusion in the study.
NURS1	Number of recordings of patients seen by practice nurse in the first year after inclusion in the study.
NURS2	Number of recordings of patients seen by practice nurse in the second year after inclusion in the study.
NURS3	Number of recordings of patients seen by practice nurse in the third year after inclusion in the study.

Var. name	Definition and creation
HREF_0_5	Number of referrals to hospital in the six months prior to inclusion in the study.
HREF1	in the first year after inclusion in the study.
HREF2	in the second year after inclusion in the study.
HREF3	in the third year after inclusion in the study.
HADM_0_5	Number of accident and emergency admissions to hospital in the six months prior to inclusion in the study.
HADM1	Number of accident and emergency admissions to hospital in the first year after inclusion in the study.
HADM2	Number of accident and emergency admissions to hospital in the second year after inclusion in the study.
HADM3	Number of accident and emergency admissions to hospital in the third year after inclusion in the study.
GYNE_0_5	Number of referrals to Gynaecologist in the six months prior to inclusion in the study.
GYNE1	Number of referrals to Gynaecologist in the first year after inclusion in the study.
GYNE2	Number of referrals to Gynaecologist in the second year after inclusion in the study.
GYNE3	Number of referrals to Gynaecologist in the third year after inclusion in the study.
ORTO_0_5	Number of referrals to Orthopaedic in the six months prior to inclusion in the study.
ORTO1	Number of referrals to Orthopaedic in the first year after inclusion in the study.
ORTO2	Number of referrals to Orthopaedic in the second year after inclusion in the study.
ORTO3	Number of referrals to Orthopaedic in the third year after inclusion in the study.
RHEU_0_5	Number of referrals to Rheumatologist in the six months prior to inclusion in the study.
RHEU1	Number of referrals to Rheumatologist in the first year after inclusion in the study.
RHEU2	Number of referrals to Rheumatologist in the second year after inclusion in the study.
RHEU3	Number of referrals to Rheumatologist in the third year after inclusion in the study.
GIIN_0_5	Number of referrals to GI investigation in the six months prior to inclusion in the study.
GIIN1	Number of referrals to GI investigation in the first year after inclusion in the study.
GII2	Number of referrals to GI investigation in the second year after inclusion in the study.
GII3	Number of referrals to GI investigation in the third year after inclusion in the study.
PHYS_0_5	Number of referrals to Physiotherapist in the six months prior to inclusion in the study.
PHYS1	Number of referrals to Physiotherapist in the first year after inclusion in the study.
PHYS2	Number of referrals to Physiotherapist in the second year after inclusion in the study.
PHYS3	Number of referrals to Physiotherapist in the third year after inclusion in the study.
RADI_0_5	Number of referrals to Radiologist in the six months prior to inclusion in the study.
RADI1	Number of referrals to Radiologist in the first year after inclusion in the study.
RADI2	Number of referrals to Radiologist in the second year after inclusion in the study.
RADI3	Number of referrals to Radiologist in the third year after inclusion in the study.

Clinical events (FILE 5, "CLINEV")

This file will contain information on incident clinical events such as fractures, cancers and cardiovascular events recorded in Problem or Note over the three years period of follow-up. The file will also contain information on when the first of these diagnoses were recorded after index date. The Read codes used to define the clinical events are outlined in appendix 1B.

The rule that was developed in datafile 1 applies in this datafile as well: If a new fracture is recorded in Read within a 3 month period following another recording, then it is likely that the second fracture is the same as the first and this will consequently be ignored. Counting for new fractures will start after month 3. This rule does not apply to hip, wrist or vertebral fractures, as the recording of these is likely to be site-specific.

Var. name	Definition and creation
HIPFN	Number of hip fractures noted in patient records after index date
DHIPP	Date of first hip fracture experienced after index date
YHIPP	Year of hip fracture
WRISTFN	Number of wrist fractures after index date
DWRISTF	Date of first wrist fracture experienced after index date
YWRISTF	Year of wrist fracture
VERTFN	Number of vertebral fractures after index date
DVERTF	Date of first vertebral fracture experienced after index date
YVERTF	Year of vertebral fracture
OTHFN	Number of other fractures
DOTHF	Date of first other fracture experienced after index date
YOTHF	Year of other fracture
ANYFN	Number of any fracture after index date
DANYF	Date of any fracture first experienced after index date
YANYF	Year of any fracture
BRCAN	Diagnosis of breast cancer
DBRCAN	Date of breast cancer experienced after index date
YBRCAN	Year of breast cancer
OVCAN	Diagnosis of ovarian cancer
DOVCAN	Date of ovarian cancer experienced after index date
YOVCAN	Year of ovarian cancer
STROKE	Diagnosis of stroke after index date
DSTROKE	Date of first stroke experienced after index date
YSTROKE	Year of stroke
MI	Diagnosis of myocardial infarction after index date
DMI	Date of first MI experienced after index date
YMI	Year of MI
CEREBR	Diagnosis of Cerebrovascular disease after index date
DCEREBR	Date of Cerebrovascular disease experienced after index date
YCEREBR	Year of Cerebrovascular disease
CHD	Diagnosis of coronary heart disease (CHD)
DCHD	Date of first recording of CHD experienced after index date
YCHD	Year of CHD

APPENDIX 5.2

READ CODES FOR THE CLASSIFICATION OF CLINICAL EVENTS

Osteoporosis

- .M61. Osteoporosis
- .1268 FH: Osteoporosis

Asthma

- .14B4 H/O: asthma
- .173A Exercise-induced asthma
- .663N Asthma disturbing sleep
- .663O Asthma not disturbing sleep
- .663P Asthma limiting activities
- .663Q Asthma not limiting activities
- .663U Asthma management plan given
- .663V Asthma severity
- .663W Asthma prophylaxis used
- .663d Emerg asthm adm since 1st appt
- .663e Asthma restricts exercise
- .663f Asthma never restricts exercise
- .663h Asthma – currently dormant
- .663j Asthma – currently active
- .8791 Further asthma – drug prevent.
- .8793 Asthma control step 0
- .8794 Asthma control step 1
- .8795 Asthma control step 2
- .8796 Asthma control step 3
- .8797 Asthma control step 4
- .8798 Asthma control step 5
- .8H2P Emergency admission asthma
- .9N1d Seen in asthma clinic
- .9OJ. Asthma monitoring admin.
- .9OJ1 Attends asthma monitoring
- .9OJ2 Refuses asthma monitoring
- .9OJ3 Asthma monitor offer default
- .9OJ4 Asthma monitor 1st letter
- .9OJ5 Asthma monitor 2nd letter
- .9OJ6 Asthma monitor 3rd letter
- .9OJ7 Asthma monitor verbal invite
- .9OJ8 Asthma monitor phone invite
- .9OJ9 Asthma monitoring deleted
- .9OJA Asthma monitoring check done
- .9OJZ Asthma monitoring admin.NOS
- .9Q21 Patient in Asthma study
- .H43. Asthma

- .H431 Extrinsic asthma – atopy
- .H432 Extrinsic asthma
- .H433 Status asthmaticus
- .H434 Asthma attack NOS
- .H43Z Asthma NOS
- .c7.. ASTHMA PROPHYLAXIS

Cervical/uterine cancer

- .685C Ca cervix screen abnormal
- .B1E. Carcinoma cervix uteri
- .B1E1 Ca cervix uteri – endocervix
- .B1E2 Ca cervix uteri – exocervix
- .B1EZ Ca cervix uteri NOS
- .B1F. Carcinoma body of uterus
- .B1G. Ca ovary/other uterine adnexa
- .B1GZ Ca ovary/uterine adnexa NOS
- .B1N1 Ca uterus NOS
- .B312 Carcinoma in situ cervix uteri

Breast cancer

- .B1D. Carcinoma breast
- .B1D1 Ca breast – nipple central
- .B1D2 Ca breast-upper,inner quadrant
- .B1D3 Ca breast-lower,inner quadrant
- .B1D4 Ca breast-upper,outer quadrant
- .B1D5 Ca breast-lower,inner quadrant
- .B1D6 Ca breast – axillary tail
- .B1DZ Ca breast - NOS
- .B311 Carcinoma in situ breast

Bronchitis and COPD

- .12D1 FH: Bronchitis/COAD
- .H16. Acute bronchitis/bronchiolitis
- .H161 Acute bronchitis
- .H16Z Acute bronchitis NOS
- .H4.. Chronic obstructive pulm. dis.
- .H41. Chronic bronchitis
- .H411 Simple chronic bronchitis
- .H412 Mucopurulent chr. bronchitis
- .H413 Obstructive chronic bronchitis
- .H414 Chronic bronchitis, acute exac
- .H415 Acute exacerbation COAD
- .H41Z Chronic bronchitis NOS
- .H4Z. Chronic obst. pulm. dis. NOS
- .H512 Acute chemical bronchitis

- .H6Z1 Bronchitis NOS
- .H46. Mild chron obstr pulm disease
- .H47. Mod chron obstr pulm disease
- .H48. Sev chron obstr pulm disease

Stroke

- .14A7 H/O: CVA/stroke
- .14AK H/O: Stroke in last year
- .662M Stroke monitoring
- .G75. Stroke/CVA – undefined

Cerebrovascular disease

- .1477 H/O: cerebrovascular disease
- .F366 Cerebral oedema
- .G7.. Cerebrovascular disease
- .G71. Cerebral haemorrhage
- .G712 Intracerebral haemorrhage
- .G71Z Cerebral haemorrhage NOS
- .G72. Precerebral arterial occlusion
- .G72Z Precerebral A. occlusion NOS
- .G73. Cerebral arterial occlusion
- .G731 Cerebral thrombosis
- .G732 Cerebral embolism
- .G73Z Cerebral A. occlusion NOS
- .G74. Transient cerebral ischaemia
- .G76. Other cerebrovascular disease
- .G761 Cerebral arteriosclerosis
- .G763 Cerebral aneurysm-non ruptured
- .G71Z Cerebral hemorrhage
- .G72. Precerebral arterial occlusion
- .G72Z Precerebral A. occlusion NOS
- .G73. Cerebral arterial occlusion
- .G731 Cerebral thrombosis
- .G732 Cerebral embolism
- .G73Z Cerebral A. occlusion NOS
- .G74. Transient cerebral ischemia
- .G76. Other cerebrovascular disease
- .G761 Cerebral arteriosclerosis
- .G763 Cerebral aneurysm-non ruptured
- .G76Z Other cerebrovascular dis.NOS
- .G77. Cerebrovasc. dis. late effects
- .G7Z. Cerebrovascular disease NOS
- .P312 Cerebral lacteration/contusion
- .P31Z Intracerebral injury NOS

Myocardial infarction (MI)

- .14A3 H/O: myocardial infarct <60
- .14A4 H/O: myocardial infarct >60
- .323. ECG: myocardial infarction
- .3232 ECG: old myocardial infarction
- .323Z ECG: myocardial infarct NOS
- .G41. Acute myocardial infarction
- .G43. Old myocardial infarction

Hysterectomy

- .1599 H/O: hysterectomy
- .15A9 H/O: myectomy/hysterotomy
- .685H No smear – benign hysterectomy
- .8L7O Hysterectomy planned
- .7CB. Hysterectomy
- .7CB1 Subtotal abdom. hysterectomy
- .7CB2 Total abdominal hysterectomy
- .7CB3 Vaginal hysterectomy
- .7CB4 Radical abdominal hysterectomy
- .7CB5 Radical vaginal hysterectomy
- .7CB7 Total hysterectomy & B.S.O.
- .7CBZ Hysterectomy NOS
- .7CD2 Uterine FB –vaginal remove NOS
- .7CF1 Hysterectomy

Hip fractures

- .M626 Pathological fracture - pelvis/thigh
- .M656 Fracture malunion-pelvis/thigh
- .P1E. #Femur
- .P1E1 #Neck of femur
- .P1E2 #Shaft of femur
- .P1E3 #Lower end of femur
- .P1EZ #Femur NOS

Wrist fractures

- .M624 Pathological fracture - fore arm
- .M625 Pathological fracture - hand
- .M654 Fracture malunion fore arm
- .P1A. #Radius/ulna
- .P1A1 #Radius/ulna - upper end
- .P1A2 #Radius/ulna -shaft
- .P1A3 #Radius/ulna-lower end-colles
- .P1AZ #Radius/ulna NOS

Vertebral fractures

- .P13. #Spine - no cord lesion
- .P131 #Cervical spine-no cord lesion
- .P132 #Thoracic spine-no cord lesion
- .P133 #Lumbar spine - no cord lesion
- .P134 #Sacrum/coccyx-no cord lesion
- .P13Z #Spine NOS - no cord lesion
- .P14. #Spine + cord lesion
- .P141 #Cervical spine + cord lesion
- .P142 #Thoracic spine + cord lesion
- .P143 #Lumbar spine + cord lesion
- .P144 #Sacrum/coccyx + cord lesion
- .P14Z #Spine NOS + cord lesion

Non-specific fractures

- .8F86 Convalesc. after fracture Rx
- .8HB9 Fracture therapy follow-up
- .9N0X Seen in fracture clinic
- .M62. Pathological fracture
- .M621 Path.fracture - multiple
- .M62Z Path.fracture - NOS
- .M65. Malunion/nonunion of fracture
- .M65Z Fracture malunion - NOS
- .P1K. Multiple fractures
- .P1Z. Fracture NOS
- .P1.. Fractures

Other specific fractures

- .M623 Path.fracture - upper arm
- .M624 Path.fracture - fore arm
- .M625 Path.fracture - hand
- .M627 Path.fracture - lower leg
- .M652 Fracture malunion - shoulder
- .M654 Fracture malunion - fore arm
- .M655 Fracture malunion - hand
- .M658 Fracture malunion - ankle/foot
- .O213 Fracture of clavicle - birth
- .P11. #Skull
- .P111 #Frontal bone
- .P112 #Parietal bone
- .P113 #Base of skull
- .P114 #Skull - multiple
- .P11Z #Skull NOS
- .P12. #Facial bones
- .P121 #Nasal bones

.P122 #Mandible
.P123 #Maxilla
.P124 #Zygoma
.P125 #Orbital floor (blow-out)
.P12Z #Facial bones NOS
.P15. #Rib/sternum/larynx/trachea
.P151 #Rib
.P152 #Sternum
.P153 #Flail chest
.P154 #Larynx/trachea
.P16. #Pelvis
.P161 #Acetabulum of pelvis
.P162 #Pubis of pelvis
.P16Z #Pelvis NOS
.P17. #Clavicle
.P18. #Scapula
.P19. #Humerus
.P191 #Humerus - upper end
.P192 #Humerus - shaft
.P193 #Humerus - lower end
.P19Z #Humerus NOS
.P1B. #Carpal bones
.P1B1 #Scaphoid
.P1BZ #Carpal bones NOS
.P1C. #Metacarpal bones
.P1C1 #Bennett's fracture
.P1CZ #Metacarpal bone NOS
.P1D. #Phalanx of finger/thumb
.P1D1 #Phalanx of finger
.P1D2 #Phalanx of thumb
.P1DZ #Phalanx of finger/thumb NOS
.P1F. #Patella
.P1G. #Tibia/fibula
.P1G1 #Tibia alone
.P1G2 #Fibula alone
.P1G3 #Tibia/fibula - upper end
.P1G4 #Tibia/fibula - shaft
.P1GZ #Tibia/fibula NOS
.P1H. #Ankle
.P1H1 #Ankle - medial malleolus
.P1H2 #Ankle - lateral malleolus
.P1H3 #Ankle - bimalleolar - Potts #
.P1H4 #Ankle - trimalleolar
.P1HZ #Ankle - NOS
.P1I. #Tarsal/metatarsal bones
.P1I1 #Calcaneus - heel
.P1IZ #Tarsal/metatarsal bone NOS
.P1J. #Phalanges of foot - toe

Hypertension

- .14A2 H/O: hypertension
- .6624 Borderline hyperten:yearly obs
- .6628 Poor hypertension control
- .6629 Hypertension:follow-up default
- .662F Hypertension treatm. started
- .662G Hypertensive treatm.changed
- .662H Hypertension treatm.stopped
- .68B1 Hypertension screen
- .68B4 Hypertension risk
- .8B26 Antihypertensive therapy
- .90D. Hypertension screen admin.
- .90DZ Hypertension screen admin. NOS
- .90I. Hypertension monitoring admin.
- .90I1 Attends hypertension monitor.
- .90I2 Refuses hypertension monnitor
- .90I3 Hyperten.monitor offer default
- .90I4 Hypertens.monitor.1st letter
- .90I5 Hypertens.monitor 2nd letter
- .90I7 Hypertens.monitor verbal inv.
- .90I8 Hypertens.monitor phone inv.
- .90I9 Hypertension monitor
- .90IA Hypertension monitor.chck done
- .90IZ Hypertens.monitoring admin.NOS
- .F363 Benign intracranial hyperten.
- .F522 Hypertensive retinopathy
- .G3.. Hypertensive disease
- .G31. Essential hypertension
- .G32. Hypertensive heart disease
- .G33. Hypertensive renal disease
- .G34. Hypertensive renal + heart dis
- .G35. Secondary hypertension
- .G36. Malignant hypertension
- .G3Z. Hypertensive disease NOS
- .G521 Primary pulmonary hypertension
- .G762 Hypertensive encephalopathy
- .I734 Portal hypertension
- .K24. Preg.+ pre-existing hypertensn

IHD/CHD

- .14A. H/O: cardiovascular disease
- .14A6 H/O: heart failure
- .14AA H/O: heart disease NOS
- .14AL H/O:Treatmen ischaem heart dis
- .14O7 At risk of heart disease
- .14S3 H/O: heart recipient
- .14S4 H/O: heart valve recipient
- .14T3 H/O: artificial heart valve

.18.. Cardiovascular symptoms
.329. ECG: heart block
.329Z ECG: heart block NOS
.5533 Angiocardiology abnormal
.68B3 Heart disease screen - risk
.77.. Cardiovascular operations
.771. Heart valve operations
.7711 Closed heart valvotomy
.7712 Open heart valvotomy
.7713 Heart valve replacement-graft
.7714 Heart valve replace-prosthesis
.7715 Heart valvuloplasty
.7716 Heart valve replacement NOS
.771Z Heart valve operation NOS
.772. Heart septum operations
.7723 Close heart septal defect
.772Z Heart septum operation NOS
.777. Heart repair
.778. Heart transplant
.77Z. Cardiovascular operations NOS
.7931 Cardioplasty
.8H2O Admit cardiothoracic emergency
.9Q22 Patient in heart disease study
.G4.. Ischaemic heart disease
.G45. Chr. ischaemic heart dis. NOS
.G452 Aneurysm of heart
.G45Z Chr. ischaemic heart dis. NOS
.G4Z. Ischaemic heart disease NOS
.G5.. Pulmonary heart disease
.G52. Chronic pulmonary heart dis.
.G52Z Pulmonary heart disease NOS
.G5Z. Other pulmonary heart disease
.G6.. Other heart disease
.G64. Cardiomyopathy
.G64Z Cardiomyopathy NOS
.G655 Heart block NOS
.G6A. Heart failure
.G6AZ Heart failure NOS
.G6B. Other heart disease
.G6B5 Cardiomegaly
.G6BZ Heart disease NOS
.G6Z. Other heart disease NOS
.N21. Bulbus/septum heart anomalies
.N22. Congenital heart anomaly NOS
.N22Z Congenital heart anomaly NOS
.RD20 Echocardiogram abnormal [D]
.RD21 X-ray heart shadow abnormal[D]
.RD22 Ultrasound cardiogram abn. [D]
.RE3. Cardiovascular function abn[D]
.RE31 ECG electrocardiogram abn. [D]
.RJ10 Cardiorespiratory failure [D]

- .14A5 H/O: angina pectoris
- .14AJ H/O: Angina in last year
- .662K Angina control
- .8B27 Antianginal therapy
- .G44. Angina pectoris

Arthritis

- .M2I7 Knee joint pain
- .M267 Osteoarthritis - knee joint
- .M261 Osteoarthritis -multiple joint
- .M2I2 Shoulder joint pain
- .M26. Osteoarthritis
- .M2I6 Hip joint pain
- .M2I8 Ankle/foot joint pain
- .M2I1 Multiple joint pain
- .M2.. Arthritis/arthrosis
- .M266 Osteoarthritis - hip joint
- .M296 Chondromalacia patellae
- .M2I5 Hand joint pain
- .M2I4 Wrist joint pain
- .M265 Osteoarthritis - hand joint
- .M2I. Pain in joint – coxalgia
- .M268 Osteoarthritis - ankle/foot
- .M29. Internal derangement of knee
- .M22. Rheumatoid arthritis
- .M2I3 Elbow joint pain
- .M2G7 Effusion - knee joint
- .M28Z Arthropathy NOS
- .M2Z. Joint disorder NOS
- .M221 Rheum. arth. - multiple joint
- .M2I9 Other joint pain
- .M262 Osteoarthritis -shoulder joint
- .M292 Derangement of medial meniscus
- .M29Z Internal knee derangement NOS
- .M269 Osteoarthritis - other joint
- .M26Z Osteoarthritis - NOS
- .M25. Polyarthropathy NOS -inflammat
- .M264 Osteoarthritis - wrist joint
- .M2IZ Joint pain NOS
- .M2L. Difficulty in walking
- .M293 Deranged lateral meniscus
- .M28. Arthropathy NOS
- .M285 Clicking joint
- .M263 Osteoarthritis - elbow joint
- .M284 Reactive arthropathy unspec
- .M2A. Articular cartilage disorder
- .M295 Loose body in knee
- .M225 Rheum. arth. - hand joint
- .M2J1 Multiple joint stiffness

.M294 Meniscus derangement NOS
.M2H7 Haemarthrosis - knee joint
.M2D2 Recur. disloc.- shoulder joint
.M2K. Other joint symptoms
.M2J5 Hand joint stiffness
.M251 Sero negative polyarthritis
.M22Z Rheumatoid arthritis NOS
.M24. Still's disease - juvenile R.A
.M2J. Joint stiffness NOS
.M227 Rheum. arth. - knee joint
.M217 Pyog. arthr. - knee joint
.M2J7 Knee joint stiffness
.M2G3 Effusion - elbow joint
.M2J2 Shoulder joint stiffness
.M224 Rheum. arth. - wrist joint
.M2B. Loose body in joint - not knee
.M2E5 Contracture - hand joint
.M291 Old bucket handle tear- medial
.M2D7 Recur. disloc. - knee joint
.M275 Traumatic arth. - hand joint
.M2G. Effusion of joint
.M226 Rheum. arth. - hip joint
.M277 Traumatic arth. - knee joint
.M278 Traumatic arth. - ankle/foot
.M2J9 Other joint stiffness
.M228 Rheum. arth. - ankle/foot
.M279 Traumatic arth. - other joint
.M283 Post-viral arthropathy
.M2J6 Hip joint stiffness
.M2J8 Ankle/foot joint stiffness
.M2G8 Effusion - ankle/foot
.M216 Pyog. arthr. - hip joint
.M2H. Haemarthrosis
.M213 Pyog. arthr. - elbow joint
.M2J4 Wrist joint stiffness
.M222 Rheum. arth. - shoulder joint
.M2D9 Recur. disloc. - other joint
.M218 Pyog. arthr. - ankle/foot
.M2E. Contracture of joint
.M2JZ Joint stiffness NOS
.M2E9 Contracture - other joint
.M2G5 Effusion - hand joint
.M2J3 Elbow joint stiffness
.M2H3 Haemarthrosis - elbow joint
.M272 Traumatic arth.-shoulder joint
.M2G2 Effusion - shoulder joint
.M27. Traumatic arthropathy
.M282 Climacteric arthritis
.M2D. Recurrent dislocation of joint
.M21. Pyogenic arthritis
.M215 Pyog. arthr. - hand joint

.M2D5 Recur. disloc. - hand joint
.M273 Traumatic arth. - elbow joint
.M274 Traumatic arth. - wrist joint
.M2D6 Recur. disloc. - hip joint
.M2E3 Contracture - elbow joint
.M229 Rheum. arth. - other joint
.M2H2 Haemarthrosis - shoulder joint
.M281 Allergic arthritis
.M2G4 Effusion - wrist joint
.M212 Pyog. arthr. - shoulder joint
.M2C2 Path. disloc. - shoulder joint
.M2G6 Effusion - hip joint
.M223 Rheum. arth. - elbow joint
.M2D3 Recur. disloc. - elbow joint
.M2E7 Contracture - knee joint
.M276 Traumatic arth. - hip joint
.M2DZ Recurrent dislocation NOS
.M23. Felty's syndrome
.M2D8 Recur. disloc. - ankle/foot
.M214 Pyog. arthr. - wrist joint
.M2C5 Path. disloc. - hand joint
.M2E8 Contracture - ankle/foot
.M2H5 Haemarthrosis - hand joint
.M219 Pyog. arthr. - other joint
.M2C6 Path. disloc. - hip joint
.M2F6 Ankylosis - hip joint
.M2G9 Effusion - other joint
.M21Z Pyogenic arthritis NOS
.M2C3 Path. disloc. - elbow joint
.M2C7 Path. disloc. - knee joint
.M2C9 Path. disloc. - other joint
.M2E4 Contracture - wrist joint
.M2E6 Contracture - hip joint
.M2EZ Contracture of joint NOS
.M2G1 Effusion - multiple joint
.M2GZ Effusion of joint NOS
.M2H8 Haemarthrosis - ankle/foot
.M211 Pyog. arthr. - multiple joint
.M27Z Traumatic arthritis NOS
.M2E1 Contracture - multiple joint
.M2H4 Haemarthrosis - wrist joint
.M2H9 Haemarthrosis - other joint
.M271 Traumatic arth.-multiple joint
.M2D4 Recur. disloc. - wrist joint
.M2F7 Ankylosis - knee joint
.M2FZ Ankylosis of joint NOS
.M2E2 Contracture - shoulder joint
.M2F8 Ankylosis - ankle/foot
.M2F9 Ankylosis - other joint
.M2C. Pathological dislocation
.M2C8 Path. disloc. - ankle/foot

- .M2CZ Pathological dislocation NOS
- .M2F. Ankylosis of joint
- .M2H1 Haemarthrosis - multiple joint
- .M2H6 Haemarthrosis - hip joint

GI Disease

- .12E1 FH: Peptic ulceration
- .14C1 H/O: peptic ulcer
- .1956 Peptic ulcer symptoms
- .1972 Epigastric pain
- .25C3 O/E - abd. pain - epigastrium
- .25C9 O/E - abd. pain - hypogastrium
- .363. Endoscopy: gastrointestinal
- .7956 Gastric fundusectomy
- .7981 Perf. peptic ulcer closure
- .7982 Gastric ulcer suture
- .7984 Gastric anastomosis revision
- .I22. Gastric ulcer
- .I22Z Gastric ulcer NOS
- .I24. Peptic ulcer NOS = PU
- .I24Z Peptic ulcer NOS
- .I25. Gastritis and duodenitis
- .I251 Acute gastritis
- .I252 Chronic (atrophic) gastritis
- .I253 Alcoholic gastritis
- .I255 Helicobacter gastritis
- .I25Z Gastritis/duodenitis NOS
- .I264 Dyspepsia, indigestion NOS
- .P784 Burn - oesophagus
- .P785 Burn - gastrointestinal tract
- .R906 Epigastric pain [D]

Rheumatism

- .M421 Shoulder syndrome
- .M422 Tennis elbow – epicondylitis
- .M4G. Leg cramps
- .M426 Tendinitis of ankle/tarsus
- .M4A7 Plantar fasciitis
- .M4E4 Pain in left leg
- .M434 Synovitis/tenosyn.- wrist
- .M4E5 Pain in right leg
- .M4A8 Muscle injury / strain
- .M42Z Peripheral enthesopathy NOS
- .M453 Bursitis - elbow
- .M457 Bursitis - knee
- .M464 Ganglion/synov.cyst - wrist
- .M44. Bunion

.M41. Polymyalgia rheumatica
.M435 Synovitis/tenosyn.- hand
.M4E2 Pain in left arm
.M4A6 Dupuytren's contracture
.M4AZ Muscle/ligament disorder NOS
.M4E3 Pain in right arm
.M4CZ Myalgia/myositis - NOS
.M4E9 Pain in limb NOS
.M465 Ganglion/synov.cyst - hand
.M458 Bursitis - ankle/foot
.M4E. Pain in limb
.M4B. Rheumatism/fibrositis NOS
.M468 Ganglion/synov.cyst-ankle/foot
.M46. Ganglion/synovial cyst
.M4DZ Neuralgia/neuritis - NOS
.M4A. Muscle/ligament disorder NOS
.M4E1 Pain in limb - multiple
.M4B2 Rheumatism NOS - shoulder
.M456 Bursitis - hip
.M43. Synovitis/tenosynovitis
.M4D. Neuralgia/neuritis NOS
.M4C. Myalgia/myositis NOS
.M438 Synovitis/tenosyn.- ankle/foot
.M477 Rupture of synovium - knee
.M4BZ Rheumatism/fibrositis NOS
.M425 Patellar tendinitis
.M4C1 Myalgia/myositis - multiple
.M4B1 Rheumatism NOS - multiple
.M4C2 Myalgia/myositis - shoulder
.M488 Tendon rupture - ankle/foot
.M437 Synovitis/tenosyn.- knee
.M4.. Nonarticular rheumatism
.M42. Peripheral enthesopathies
.M45. Bursitis
.M433 Synovitis/tenosyn.- elbow
.M467 Ganglion/synov.cyst - knee
.M4C6 Myalgia/myositis -pelvis/thigh
.M4C7 Myalgia/myositis - lower leg
.M432 Synovitis/tenosyn.- shoulder
.M485 Tendon rupture - hand
.M43Z Synovitis/tenosyn.- NOS
.M46Z Ganglion/synov.cyst - NOS
.M4A9 Fibromyalgia
.M4C3 Myalgia/myositis - upper arm
.M49. Synovial/tendon problem NOS
.M45Z Bursitis NOS
.M423 Periarthritis of wrist
.M4D8 Neuralgia/neuritis -ankle/foot
.M4A5 Hypermobility syndrome
.M436 Synovitis/tenosyn.- hip
.M4D2 Neuralgia/neuritis - shoulder

.M4A3 Muscle wasting/atrophy NOS
.M4D6 Neuralgia/neuritis.-pelvis/thigh
.M4F5 F.B. left in hand
.M48. Tendon rupture - non traumatic
.M4B5 Rheumatism NOS - hand
.M4B7 Rheumatism NOS - knee
.M48Z Tendon rupture - NOS
.M4D7 Neuralgia/neuritis - lower leg
.M482 Tendon rupture - shoulder
.M4A4 Laxity of ligaments
.M4D3 Neuralgia/neuritis - upper arm
.M424 Tendinitis of hip region
.M4B6 Rheumatism NOS - hip
.M4Z. Nonarticular rheumatism NOS
.M452 Bursitis - shoulder
.M463 Ganglion/synov.cyst - elbow
.M4D4 Neuralgia/neuritis - fore arm
.M4B8 Rheumatism NOS - ankle/foot
.M487 Tendon rupture - knee
.M4C4 Myalgia/myositis - fore-arm
.M4D5 Neuralgia/neuritis - hand
.M431 Synovitis/tenosyn.- multiple
.M4F8 F.B. left in ankle/foot
.M461 Ganglion/synov.cyst - multiple
.M4A2 Muscle calcific./ossification
.M462 Ganglion/synov.cyst - shoulder
.M4B3 Rheumatism NOS - elbow
.M4FZ F.B. left in tissue NOS
.M4F. Foreign body left in tissue
.M4B4 Rheumatism NOS - wrist
.M484 Tendon rupture - wrist
.M4A1 Infective myositis
.M4D1 Neuralgia/neuritis - multiple
.M483 Tendon rupture - elbow
.M454 Bursitis - wrist
.M455 Bursitis - hand
.M4C5 Myalgia/myositis - hand
.M4C8 Myalgia/myositis - ankle/foot
.M4F7 F.B. left in knee
.M481 Tendon rupture - multiple
.M451 Bursitis - multiple
.M4F2 F.B. left in shoulder
.M4F3 F.B. left in elbow
.M466 Ganglion/synov.cyst - hip
.M486 Tendon rupture - hip
.M47. Rupture of synovium
.M4F1 F.B. left in tissue - multiple
.M4F4 F.B. left in wrist
.M478 Rupture of synovium-ankle/foot

Vertebral Column Disease

- .M3Z5 Backache NOS
- .M3Z4 Sciatica
- .M342 Cervicalgia
- .M321 Cervical spond.- no myelopathy
- .M344 Torticollis NOS
- .M3Z3 Lumbago
- .M32. Spondyloses
- .M343 Brachial (cervical) neuritis
- .M3Z7 Coccygeal disorder NOS
- .M333 Lumbar disc lesion - displaced
- .M323 Lumbosacral spond-no myelopath
- .M312 Sacroiliitis
- .M322 Cervical spond.with myelopathy
- .M33. Intervertebral disc disorders
- .M336 Lumbar disc degeneration
- .M34. Cervical disorder NOS
- .M331 Cervical disc lesion-displaced
- .M34Z Cervical/neck disorder NOS
- .M311 Ankylosing spondylitis
- .M32Z Spondylosis NOS
- .M3.. Vertebral column syndromes
- .M325 Thoracic spondylosis
- .M334 Cervical disc degeneration
- .M33Z Intervertebral disc lesion NOS
- .M3Z1 Spinal stenosis excl. cervical
- .M3Z2 Pain in thoracic spine
- .M324 Lumbosacral spond + myelopathy
- .M3ZZ Back disorder/symptom NOS
- .M3Z. Back disorders - other
- .M332 Thoracic disc lesion-displaced
- .M3Z6 Sacral disorder NOS
- .M337 Disc disorder + myelopathy
- .M335 Thoracic disc degeneration
- .M338 Post laminectomy syndrome
- .M341 Cervical spinal stenosis
- .M31Z Inflam. spondylopathies NOS
- .M31. Inflammatory spondylopathies

APPENDIX 4.3

READ CODES FOR THE CLASSIFICATION OF ADMINISTRATIVE ACTIVITY

Hospital Admissions

- .8H1. Admit to intensive care unit
- .8H11 Admit to cardiac ITU
- .8H12 Admit to respiratory ITU
- .8H13 Admit to neurological ITU
- .8H14 Admit to metabolic ITU
- .8H15 Admit to burns unit
- .8H1Z Admit to intensive c.u. NOS
- .8H2. Emergency hospital admission
- .8H21 Admit medical emergency unsp.
- .8H22 Admit surgical emergency unsp.
- .8H23 Admit psychiatric emergency
- .8H24 Admit geriatric emergency
- .8H25 Admit paediatric emergency
- .8H26 Admit gynaecological emergency
- .8H27 Admit obstetric emergency
- .8H28 Admit orthopaedic emergency
- .8H29 Admit ENT emergency
- .8H2A Admit trauma emergency
- .8H2B Admit ophthalmological emerg.
- .8H2C Admit rheumatology emergency
- .8H2D Admit dermatology emergency
- .8H2E Admit neurology emergency
- .8H2F Admit urology emergency
- .8H2G Admit radiotherapy emergency
- .8H2H Admit haematology emergency
- .8H2I Admit plastic surgery emergenc
- .8H2J Admit diabetic emergency
- .8H2K Admit oral surgical emergency
- .8H2L Admit psychogeriatric emerg.
- .8H2M Admit renal medicine emergency
- .8H2N Admit neurosurgical emergency
- .8H2O Admit cardiothoracic emergency
- .8H2P Emergency admission asthma
- .8H2Z Admit hospital emergency NOS
- .8H3. Non-urgent hospital admission
- .8H31 Non-urgent hosp.admission unsp
- .8H36 Non-urgent medical admission
- .8H37 Non-urgent surgical admission
- .8H38 Non-urgent psychiatric admisn.
- .8H39 Non-urgent geriatric admission

.8H3A Non-urgent paediatric admisn.
.8H3B Non-urgent gynaecol.admission
.8H3C Non-urgent obstetric admission
.8H3D Non-urgent orthopaedic admisn.
.8H3E Non-urgent ENT admission
.8H3F Non-urgent trauma admission
.8H3G Non-urgent ophthalmolog.admisn
.8H3H Non-urgent rheumatology admisn
.8H3I Non-urgent dermatology admisn.
.8H3J Non-urgent neurology admission
.8H3K Non-urgent urology admission
.8H3L Non-urgent radiotherapy admisn
.8H3M Non-urgent haematology admisn.
.8H3N Non-urgent plastic surg.admisn
.8H3O Non-urgent diabetic admission
.8H3P Non-urgent respiratory admis.
.8H3S Non-urgent neurosurgical admis
.8H3U Non urgent oral surg admission
.8H3V Non-urgent cardiolog admission
.8H3Z Other hospital admission NOS
.8HM. Listed for admission to hosp.
.8HM1 Listed for Gen Medicine admiss
.8HM2 Listed for Paediatric admiss'n
.8HM3 Listed for Dermatology admissn
.8HM4 Listed for Cardiology admiss'n
.8HM5 Listed for Immunology admiss'n
.8HM6 Listed for Neurology admission
.8HM7 Listed for Geriatrics admiss'n
.8HM8 Listed for Gastroenterol admis
.8HM9 Listed for Psychiatric admissn
.8HMA Listed for Rheumatology admiss
.8HMB Listed for Chest Med admission
.8HMC Listed for Psychogeriatr admis
.8HMD Listed for Oncology admission
.8HME Listed for Diabetology admissn
.8HMF Listed for other medical admis
.8HMG Listed for Gen Surgery admissn
.8HMH Listed for Ophthalmology admis
.8HMJ Listed for E.N.T. admission
.8HMK Listed for Orthopaedic admissn
.8HML Listed for Neurosurgical admis
.8HMM Listed for Paed Surgery admiss
.8HMN Listed for Obstetric admission
.8HMO Listed for Gynae admission
.8HMP Listed for Plastic Surg admiss
.8HMQ Listed for Oral Surgery admiss
.8HMR Listed for Urology admission
.8HMZ Listed for hosp admission NOS
.8HO. Admission funding status
.8HO1 Admission fundable procedure
.8HO2 Admission non-fundable proced.

- .8HOZ Admission funding status NOS
- .8HX. Admission to hospice
- .8HX0 Urgent admission to hospice
- .8HX1 Routine admission to hospice
- .8HX2 Admission to hospice - respite

Respiratory Diseases

- .14B. H/O: respiratory disease
- .14B1 H/O: hay fever
- .14BZ H/O: respiratory disease NOS
- .17.. Respiratory symptoms
- .17Z. Respiratory symptoms NOS
- .17ZZ Respiratory symptom NOS
- .1827 Painful breathing -pleurodynia
- .23E. O/E - non-specific resp.lesion
- .23E5 O/E - fibrosis of lung present
- .3374 Lung function testing abnormal
- .3375 Lung function mildly obstruct.
- .3376 Lung function signific. obstr.
- .3377 Lung function restrictive
- .3394 Resp. flow rate abnormal
- .663B Resp. treatment changed
- .663C Resp.dis.treatment started
- .663D Resp.dis.treatment stopped
- .76.. Respiratory system operations
- .764. Excision of lung/bronchus
- .7668 Lung transplant
- .876Z Respiratory medication NOS
- .H... Respiratory system diseases
- .H231 Chronic rhinitis
- .H28. Hay fever
- .H29. Allergic rhinitis NOS
- .H6.. Other resp. system diseases
- .H6Z. Respiratory disease NOS
- .H6ZZ Respiratory disease NOS
- .HZ.. Respiratory diseases NOS
- .R6.. Respiratory/chest symptoms [D]
- .R60. Respiratory abnormalities [D]
- .R607 Respiratory insufficiency [D]
- .R60Z Respiratory abnormality NOS[D]
- .R651 Painful respiration [D]
- .R6Z. Respir/chest symptoms NOS [D]
- .RD1. Lung x-ray field abnormal [D]
- .RD10 Lung x-ray, coin lesion [D]
- .RD11 Lung x-ray, shadow [D]
- .RD1Z Lung x-ray abnormal NOS [D]
- .RE20 Lung scan abnormal [D]
- .RE21 Ventilatory capacity reduc.[D]
- .RJ1. Respiratory failure [D]

- .RJ10 Cardiorespiratory failure [D]
- .RJ11 Respiratory arrest [D]
- .RJ1Z Respiratory failure NOS [D]

Smoking

- .137. Tobacco consumption
- .1372 Trivial smoker - < 1 cig/day
- .1373 Light smoker - 1-9 cigs/day
- .1374 Moderate smoker - 10-19 cigs/d
- .1375 Heavy smoker - 20-39 cigs/day
- .1376 Very heavy smoker - 40+cigs/d
- .1377 Ex-trivial smoker (<1/day)
- .1378 Ex-light smoker (1-9/day)
- .1379 Ex-moderate smoker (10-19/day)
- .137A Ex-heavy smoker (20-39/day)
- .137B Ex-very heavy smoker (40+/day)
- .137C Keeps trying to stop smoking
- .137D Admitted tobacco cons untrue
- .137E Tobacco consumption unknown
- .137F Ex-smoker - amount unknown
- .137G Trying to give up smoking
- .137H Pipe smoker
- .137I Passive smoker
- .137J Cigar smoker
- .137K Stopped smoking
- .137L Current non-smoker
- .137M Rolls own cigarettes
- .137N Ex pipe smoker
- .137O Ex cigar smoker
- .137P Cigarette smoker
- .137Q Smoking started
- .137R Current smoker
- .137S Ex smoker
- .137T Date ceased smoking
- .137Z Tobacco consumption NOS
- .6791 Health ed. - smoking
- .67A3 Pregnancy smoking advice
- .9001 Attends stop smoking monitor.
- .9002 Refuses stop smoking monitor
- .9003 Stop smoking monitor default
- .9004 Stop smoking monitor 1st lettr
- .9005 Stop smoking monitor 2nd lettr
- .9007 Stop smoking monitor verb.inv.
- .9008 Stop smoking monitor phone inv
- .9009 Stop smoking monitoring delete
- .900A Stop smoking monitor.chck done
- .900Z Stop smoking monitor admin.NOS

All Referrals

- .8H.. Referral for further care
- .8H4. Referral to physician
- .8H41 General medical referral
- .8H42 Paediatric referral
- .8H43 Dermatological referral
- .8H44 Cardiological referral
- .8H45 Immunological referral
- .8H46 Neurological referral
- .8H47 Geriatric referral
- .8H48 Gastroenterological referral
- .8H49 Psychiatric referral
- .8H4A Referred to venereologist
- .8H4B Referred to rheumatologist
- .8H4C Referred to chest physician
- .8H4D Referral to psychogeriatrician
- .8H4E Referral to oncologist
- .8H4F Referral to diabetologist
- .8H4G Refer to radiotherapist
- .8H4H Refer to clin pharmacologist
- .8H4I Refer to geneticist
- .8H4J Referred to anaesthetist
- .8H4K Referred to endocrinologist
- .8H4L Referred to nephrologist
- .8H4M Ref to community paediatrician
- .8H4N Ref to paediatric cardiologist
- .8H4O Ref to paediatric neurologist
- .8H4Z Referral to physician NOS
- .8H5. Referral to surgeon
- .8H51 General surgical referral
- .8H52 Ophthalmological referral
- .8H53 ENT referral
- .8H54 Orthopaedic referral
- .8H55 Neurosurgical referral
- .8H56 Paediatric surgical referral
- .8H57 Obstetric referral
- .8H58 Gynaecological referral
- .8H59 Referred to plastic surgeon
- .8H5A Referred to oral surgeon
- .8H5B Referred to urologist
- .8H5C Referred to thoracic surgeon
- .8H5D Referred to vascular surgeon
- .8H5E Burns referral
- .8H5F Refer to maxillofacial surgeon
- .8H5G Refer Cardiothoracic surgeon
- .8H5Z Referral to surgeon NOS
- .8H6. Referral to other doctor
- .8H61 Referral to private doctor
- .8H62 Referral to G.P.
- .8H63 Refer to casualty officer

.8H64 Refer to house officer
.8H65 Refer to hospital registrar
.8H66 Refer to child medical officer
.8H67 Referred for radiotherapy
.8H68 Referred to haematologist
.8H69 Refer to pain clinic
.8H6A Refer to terminal care consult
.8H6B Refer to public health
.8H6Z Refer to other doctor NOS
.8H7. Other referral
.8H71 Refer to practice nurse
.8H72 Refer to district nurse
.8H73 Refer to health visitor
.8H74 Refer to mid-wife
.8H75 Refer to social worker
.8H76 Refer to dietician
.8H77 Refer to physiotherapist
.8H78 Refer to counsellor
.8H79 Refer to community day centre
.8H7A Refer to mental health worker
.8H7B Refer to community psych.nurse
.8H7C Refer diabetic liaison nurse
.8H7D Refer to osteopath
.8H7E Refer to chiropractor
.8H7F Referred to dentist
.8H7G Refer to speech therapist
.8H7H Refer to optician
.8H7I Refer to partner
.8H7J Refer to occupational therap.
.8H7K Refer to orthoptist
.8H7L Refer for terminal care
.8H7M Refer to stoma nurse
.8H7N Refer for colposcopy
.8H7O Refer to Radiology department
.8H7P Refer to pathology department
.8H7Q Refer to surgical fitter
.8H7R Refer to chiropodist
.8H7T Refer to psychologist
.8H7U Refer to homeopathist
.8H7V Refer to audiologist
.8H7W Refer to TOP counselling
.8H7X Refer to podiatry
.8H7Y Refer to acupuncture
.8H7Z Refer to other health worker
.8H7a Refer to hospital
.8H7b Refer to day hospital
.8H7c Refer to occup health dept
.8H7d Refer to school nurse
.8H7e Referral to nurse practitioner
.8H7f Referral to diabetes nurse
.8HA5 Follow-up refused

.8HC. Refer to hospital casualty
.8HC1 Refer to A. & E. department
.8HC2 Refer to hosp. eye casualty
.8HC3 Refer to hosp. paed's casualty
.8HCZ Refer to hospital casualty NOS
.8HD. Refer to hospital OPD
.8HH. Referred - other care
.8HH5 Refer domiciliary physio
.8HH6 Referral to Macmillan nurse
.8HI. Refer - no direct consultation
.8HJ. Self-referral to hospital
.8HJ1 General medical self-referral
.8HJ2 General surgical self-referral
.8HJ3 Psychiatric self-referral
.8HJ4 Geriatric self-referral
.8HJ5 Paediatric self-referral
.8HJ6 Gynaecological self-referral
.8HJ7 Obstetric self-referral
.8HJ8 Orthopaedic self-referral
.8HJ9 ENT self-referral
.8HJA Trauma self-referral
.8HJB Ophthalmology self-referral
.8HJC Rheumatology self-referral
.8HJD Dermatology self-referral
.8HJE Neurology self-referral
.8HJF Urology self-referral
.8HJG Radiotherapy self-referral
.8HJH Haematology self-referral
.8HJI Plastic surgery self-referral
.8HJZ Self-referral to hospital NOS
.8HP. Referral for laboratory tests
.8HP1 Referral for haematology test
.8HP2 Refer for microbiological test
.8HP3 Refer for biochemical test
.8HP4 Refer for radio-immune assay
.8HP5 Refer for thyroid test
.8HP6 Refer for cytological test
.8HP7 Refer for histology
.8HP8 Refer for serological testing
.8HP9 Refer for immunological test
.8HPZ Referral for lab test NOS
.8HQ. Refer for imaging
.8HQ1 Refer for X-Ray
.8HQ2 Refer for ultrasound investign
.8HQ3 Refer for NMR scanning
.8HQ4 Refer for CAT scanning
.8HQ5 Refer for medical photography
.8HQ6 Refer for angiogram
.8HQZ Refer for imaging NOS
.8HR. Refer for physiology investign
.8HR1 Refer for ECG recording

.8HR2 Refer for audiometry
.8HR3 Refer for nerve conduct study
.8HR4 Refer for lung function test
.8HR5 Refer for EEG
.8HR6 Refer to Urodynamic studies
.8HR7 Refer for vascular studies
.8HR8 Referral for 24 hour BP
.8HR9 Referral for 24 hour ECG
.8HRA Referral for exercise ECG
.8HRZ Refer for physiology test NOS
.8HS. Referral for endoscopy
.8HS0 Refer for sigmoidoscopy
.8HT. Referral to clinic
.8HT1 Referral to lipid clinic
.8HT2 Refer to hearing aid clinic
.8HT3 Referral to audiology clinic
.8HT4 Referral to orthodontic clinic
.8HT5 Referral hypertension clinic
.8HT8 Referral to mammography clinic
.8HT9 Referral to antenatal clinic
.8HTA Referral to postnatal clinic
.8HTB Referral to fertility clinic
.8HTC Referral to well woman clinic
.8HTD Referral fam planning clinic
.8HTE Referral to other clinic
.8HTF Referral to emergency clinic
.8HTG Refer acute chest pain clinic
.8HTH Referral to back pain clinic
.8HTI Referral to breast clinic
.8HTZ Referral to clinic NOS
.8HU. Referral other investigation
.8HU0 Referral for bronchoscopy
.8HU1 Referral for colonoscopy
.8HU2 Referral for sigmoidoscopy
.8HV. Private referral
.8HV0 Private general surg referral
.8HV1 Private ophthalmolog referral
.8HV2 Private ENT referral
.8HV3 Private orthopaedic referral
.8HV4 Private neurosurgical referral
.8HV5 Private paediat surg referral
.8HV6 Private obstetric referral
.8HV7 Private gynaecologic referral
.8HV8 Private referral to plast surg
.8HV9 Private referral to oral surg
.8HVA Private referred to urologist
.8HVB Privat referral to thorac surg
.8HVC Private referral to vasc surg
.8HVD Private referr maxillofac surg
.8HVE Private referr cardiothor surg
.8HVF Private referral to physician

.8HVG Private general medic referral
.8HVV Private referral paediatrician
.8HVI Private dermatology referral
.8HVJ Private cardiological referral
.8HVK Private immunological referral
.8HVL Private neurological referral
.8HVM Private geriatric referral
.8HVN Private gastroenterol referral
.8HVO Private psychiatric referral
.8HVP Private referral to venereolog
.8HVQ Private referr rheumatologist
.8HVR Private ref to chest physician
.8HVT Private referral to oncologist
.8HVU Private referral diabetologist
.8HVV Private ref to radiotherapist
.8HVV Private ref to radiotherapist
.8HVV Private ref to radiotherapist
.8HVV Private ref to radiotherapist
.8HVY Private referral anaesthetist
.8HVZ Private ref to endocrinologist
.8HW. Referral by nurse
.8HY. Referral to hospice
.8HZ. Further care referral NOS
.8HZ0 Referral needed

Osteopath Referrals

.8H7D Refer to osteopath

Physiotherapy Referrals

.8HH5 Refer domiciliary physio
.8H77 Refer to physiotherapist

Radiology Referrals

.8H4G Refer to radiotherapist
.8H67 Referred for radiotherapy
.8H70 Refer to Radiology department
.8HJG Radiotherapy self-referral
.8HP4 Refer for radio-immune assay
.8HQ1 Refer for X-Ray
.8HQ3 Refer for NMR scanning
.8HQ4 Refer for CAT scanning
.8HQZ Refer for imaging NOS
.8HVV Private ref to radiotherapist

Orthopaedic Referrals

.8H54 Orthopaedic referral
.8HJ8 Orthopaedic self-referral

.8HV3 Private orthopaedic referral

Gynaecological Referrals

.8H58 Gynaecological referral
.8HJ6 Gynaecological self-referral
.8HV7 Private gynaecologic referral

Gastro-Intestinal Referrals

.8H48 Gastroenterological referral
.8HU1 Referral for colonoscopy
.8HU2 Referral for sigmoidoscopy
.8HVN Private gastroenterol referral

Rheumatology Referrals

.8H4B Referred to rheumatologist
.8HJC Rheumatology self-referral
.8HVQ Private referr rheumatologist

Nursing Codes

.13F6 Nursing/other home
.13G1 District nurse attends
.13GA District nurse involv.stopped
.6314 Nursing home birth
.64Q6 Child ref. to school nurse
.66S5 Shared care: district nurse/GP
.66S7 Full care: nurse practitioner
.66S8 Shared care: practice nurse/GP
.6732 Counselling by a nurse
.6B0. Nurse health promotion
.8A31 Head injury - nursing supervis
.8C1. Nursing care
.8C11 Intensive nursing care
.8C12 Nursing supervision
.8C13 Intermediate nursing care
.8C15 Nursing care - dressing
.8C16 Nursing care - irrigation
.8C17 Nursing care - injections
.8C18 Nursing care-enema administrat
.8C19 Nursing care - bathing patient
.8C1A Nursing care - pressure areas
.8C1B Nursing care - blood taken
.8C1C Nursing care operation assist
.8C1Z Nursing care NOS
.8GA. Psychological nursing

.8H71 Refer to practice nurse
.8H72 Refer to district nurse
.8H7B Refer to community psych.nurse
.8H7C Refer diabetic liaison nurse
.8H7M Refer to stoma nurse
.8H7d Refer to school nurse
.8H7e Referral to nurse practitioner
.8H7f Referral to diabetes nurse
.8HE6 Delayed disch nursing home
.8HH6 Referral to Macmillan nurse
.8HW. Referral by nurse
.8O07 Prov special educ needs nursry
.9493 Patient died in nursing home
.9N1G Seen in nursing home
.9N22 Seen by practice nurse
.9N24 Seen by district nurse
.9N2J Seen by hospital nurse
.9N2Y Seen by community paed nurse
.9N64 Referred by nurse
.9NFA District nurse visit
.9NFE First annl visit distr nurse

APPENDIX 5.5.

DRUGS WITH AN EFFECT IN THE TREATMENT OF OSTEOPOROSIS

ALLYLESTRENOL
AREDIA
BONEFOS
CLIMAGEST
CLIMESSE
CONJ OEST/NORGESTR
CONJUGATED OESTROG
CRINONE WYE
CYCLO PROGYNOVA
CYCLOGEST
DEPO PROVERA
DERMESTRIL
DUPHASTON
DYDROGESTERONE
ELLESTE DUET
ELLESTE DUET CONTI
ELLESTE SOLO
ELLESTE SOLO MX
ESTRACOMBI
ESTRADERM
ESTRADERM MX
ESTRAPAK
ETHINYLOESTRADIOL
EVISTA
EVOREL
EVOREL CONTI
EVOREL PAK
EVOREL SEQUI
FEMAPAK
FEMATRIX
FEMOSTON
FEMSEVEN
GESTANIN
GESTONE
HARMOGEN
HORMONIN
HYDROXYPROGESTRONE
IMPROVERA
KLIOVANCE
LORON
MEDROXYPROGESTERON
MENOPHASE
MENOEST
MENZOL

MICRONOR
NORETHISTERONE
NUVELLE
NUVELLE CONTINUOUS
NUVELLE TS
OESTRADIOL
OESTRADIOL/NORETH
OESTRIOL
OESTROGEL
OVESTIN
PAMIDRONATE
PIPERAZINE OESTRON
PREMARIN
PREMPAK
PRIMOLUT N
PROGESTERONE
PROGYNOVA
PROGYNOVA TS
PROLUTON DEPOT
PROVERA
RALOXIFENE
SANDRENA
SKELID
SODIUM CLODRONATE
STILBOESTROL
TILUDRONIC
TRIDESTRA
TRISEQUENS
UTOVLAN
ZUMENON

APPENDIX 5.6.

STATISTICAL VARIABLES FOR THE SECOND STUDY

Var. name	Label	Definition and creation	Notes
PATNR	Patient ID	Unique identifier for a study subject as given in the source data.	
INXDRG	Index drug	The study drug prescribed to a study subject on the index date. The variable categories will be named as following: EVISTA FOSAMAX DIDRONEL PMO CONTROL	
INXDAT	Date of Index drug	The date on which the study drug was initiated during the study exposure period	+ DATE WILL BE ASSIGNED TO THE CONTROL PATIENTS ACCORDING TO THE INDEX DATE OF THE PATIENTS THEY WERE MATCHED TO
INXDIAG	Diagnosis for index drug	The diagnosis under which the index drugs was prescribed. Osteoporosis: OSTEO Other: OTHER If no diagnosis of osteoporosis was recorded, then missing value.	
PATAGE	Patient age in index year	The age of the patient as of the year of the index date, computed from patient's age as of December 1999	
HOST	History of osteoporosis	Mention of osteoporosis-related diagnosis in the patient record prior to index date	Problem or Note of either: .M61 Osteoporosis .1268 FH: Osteoporosis
STOP_3	Stop date - 3	Date at which the coverage of the last prescription of the first treatment episode with index drug ended. Date when using a definition of drug-free period of 3 months.	Patient must have at least 3 months drug free following stop date
STOP_4	Stop date - 4	Stop date when using a definition of drug-free period of 4 months.	Patient must have at least 4 months drug free following stop date
STOP_5	Stop date - 5	Stop date when using a definition of drug-free period of 5 months	Patient must have at least 5 months drug free following stop date
STOP_6	Stop date - 6	Stop date when using a definition of drug-free period of 6 months	Patient must have at least 6 months drug free following stop date
COST_3	Cost of	Total cost of the index therapy until stop	Cost of Index therapy may vary according to

Var. name	Label	Definition and creation	Notes
	index therapy	date for patients identified in STOP_3	the definition of Stop Date, so individual variables needed for each definition.
COST_4	Cost of index therapy	Total cost of the index therapy until stop date for patients identified in STOP_4	
COST_5	Cost of index therapy	Total cost of the index therapy until stop date for patients identified in STOP_5	
COST_6	Cost of index therapy	Total cost of the index therapy until stop date for patients identified in STOP_6	
A2B1-6	Cost of H2 antagonists prior	Cost of H2 antagonists (A02B1) in the 6 months prior to index date	
A2B1+6	Cost of H2 antagonists after	Cost of H2 antagonists in the 6 months after index date	
A2B2-6	Cost of PPI's prior	Proton pump inhibitors (A02B2) in the 6 months prior to index date	
A2B2+6	Cost of PPI's after	Cost of Proton pump inhibitors (A02B2) in the 6 months after index date	
N2A-6	Cost of narcotic analgesics prior	Cost of narcotic analgesics (N02A only) in the 6 months prior to index date	
N2A+6	Cost of narcotic analgesics after	Cost of narcotic analgesics (N02A only) in the 6 months after index date	
N2B-6	Cost of non-narcotic analgesics prior	Cost of non-narcotic analgesics (N2B) in the 6 months prior to index date	
N2B+6	Cost of non-narcotic analgesics after	Cost of non-narcotic analgesics (N2B) in the 6 months after index date	
M1A-6	Cost of NSAIDs prior	Cost of antirheumatic non-steroidals (M1A) in 6 months prior to index date	
M1A+6	Cost of NSAIDs after	Cost of antirheumatic non-steroidals (M1A) in 6 months after index date	

APPENDIX 6

ABSTRACT BASED ON RESULTS OF FIRST STUDY IN MEDIPLUS®

USE OF BISPHOSPHONATES AND THE COST OF TREATING GASTROINTESTINAL SIDE EFFECTS IN A UK CLINICAL PRACTICE SETTING.

Urdahl HF¹, Young P¹, McKendrick J², Freemantle N¹

¹Medicines Evaluation Group, Centre for Health Economics, University of York, UK

²Eli Lilly and Company Ltd, Dextra Court, Basingstoke, UK

Background

Gastrointestinal (GI) side effects have been associated with bisphosphonate use in clinical practice. This study assesses the cost implications for treating GI problems in women taking bisphosphonates in a UK primary care setting.

Material and Methods

Information was obtained from 140 general practices contributing data to the UK Primary Care Database from IMS Health, UK *Mediplus*®. Women included in the study were those initiating therapy on the bisphosphonates, alendronate or cyclical etidronate, between January 1996 through January 1999, with a prior period of 6 months of no bisphosphonate use. Age matched women provided control data. The costs of the use of Proton Pump inhibitors (PPIs) and H2-antagonists were determined each year for 3 years and were controlled for concomitant use of non-narcotic analgesics (including NSAIDs). Cost data were non-normally distributed and many patients incurred no cost. Therefore, the non-parametric bootstrap was used to assess differences in the average cost of PPIs and H2-antagonists.

Results

666 women taking bisphosphonates (162 cyclical etidronate, 504 alendronate) and 1332 controls were analysed. The average cost in the previous year for H2-antagonists was £5.12 in the control group and £11.94 in the bisphosphonate group, and similarly £10.08 and £30.10 for PPIs. During the first year of bisphosphonate use, the average cost of H2 antagonists increased by £9.40 (95% CI 4.30 to 15.00) when compared to the controls. Similarly, the average cost of PPI's increased by £10.10 during the first year (95% CI 2.70 to 27.70) when compared to controls. Similar increases were observed in subsequent years. There was no significant increase in the number of women experiencing GI side effects.

Conclusion

The use of cyclical etidronate and alendronate are associated with a constant yearly increase in the costs of H2 antagonists and PPIs of approximately £9.50 per annum. Concomitant medication costs may be an important consideration when undertaking an economic evaluation for osteoporosis in primary care.