DOCTOR OF PHILOSOPHY

Pharmacologically-based mechanistic modelling in health economic evaluation

Pink, Joshua

Award date:
2012

Awarding institution:
Bangor University

Pharmacologically-based mechanistic modelling in health economic evaluation

PhD Thesis

Joshua Pink

Centre for Health Economics and Medicines Evaluation, Institute of Medical and Social Care Research, Bangor University

2012

Contents

Summary

This thesis concerns the linking together of the currently distinct techniques of pharmacokinetic-pharmacodynamic and health economic modelling. These have limitations, with economic models being empirical and thus hard to extrapolate outside the evidence base where they are constructed, and challenging to implement early in the drug development process. Pharmacokinetic-pharmacodynamic models, by contrast, are mechanistic but produce a limited range of outputs, not generating all the values useful to inform decision making.

Pharmacokinetic-pharmacodynamic-pharmacoeconomic models, by linking these approaches, have the potential to overcome these limitations. The feasibility, validity and applicability of such an approach are assessed through two case studies. The first contains both retrospective and prospective simulations of rituximab for the treatment of follicular lymphoma. Retrospective analyses allow simulated results to be compared with trial-based data, and show an acceptable degree of concordance between the two methods. The prospective simulation of a trial currently recruiting will enable comparisons with the results of the trial, when these become available.

The second and larger case study uses anticoagulation and stroke prophylaxis for patients with atrial fibrillation as an example. The end result is a full, prospective simulation of genotype dosed warfarin compared with both standard clinical dosing and a number of newly available oral anticoagulants. To make such an analysis possible, necessary prerequisite work was undertaken with the construction of a discrete event simulation to extrapolate both trial and simulation results to a lifetime horizon and an indirect comparison of all available treatments, to ensure all possible alternatives are considered in the analysis.

The modelling approach described has the potential to allow the calculation of earlier estimates of cost-effectiveness than are currently available, which can be used to inform and improve the efficiency of the drug development process, and enable better extrapolations of trial-based analyses to different patient populations or dosing regimens.

Acknowledgements

Having read numerous other acknowledgement sections from different people's theses, I am still none the wiser as to what one should write in them. They seem to be either overly emotional to a point where it seems it must be fake or symptomatic of a lack of perspective, or simply run through a formulaic list of people who really should already know what is being said. I personally do not feel I have enough real sentiment or the capacity suitably mimic it for either of these approaches. Therefore, in the simplest possible terms, if anyone reading this (and in reality who really does read this section) feels they deserve to be acknowledged, consider it done, and if they do not then try harder next time.

Thesis Structure

This thesis is written as a collection of research papers following a common thread, rather than as a single body of work. This follows the style of the Institute of Medical and Social Care Research, and is designed to increase experience of paper writing and benefit from peer-review comments. As such, and due to the structure of chapters 2-5 as research papers, there will be a small amount of unavoidable repetition inherent in the need for each chapter to contain sufficient background information as to be publishable. At the time of writing, chapters 4 and 5 are currently in the process of publication, with chapters 2 and 3 already having been published as:

Pink J, Lane S, Hughes DA. Mechanism-based approach to the economic evaluation of pharmaceuticals: pharmacokinetic/Pharmacodynamic/pharmacoeconomic analysis of rituximab for follicular lymphoma. Pharmacoeconomics, 30(5):413-29 (2012). doi: 10.2165/11591540

Pink J, Lane S, Pirmohamed M, Hughes DA. Dabigatran etexilate versus warfarin in management of non-valvular atrial fibrillation in UK context: quantitative benefit-harm and economic analyses. BMJ, 343:d6333 (2011). doi: 10.1136/bmj.d6333

Work on this thesis (specifically chapter 2) was also rewarded with the Lewis Sheiner Prize at the 2011 PAGE (Population Approach Group in Europe) conference.

Chapter 1

Introduction

Health economics

Health economics is the branch of economics that concerns health, and the provision of healthcare[1]. There are a number of key features of healthcare systems, including an asymmetric relationship between physicians as the primary decision makers and patients as the nominal consumers, many externalities in the effects of healthcare interventions in other areas, and the high levels of government intervention (in much of the world) associated with efforts to overcome the market failures these cause[2,3]. One application of health economics is to inform decision making, from the overall structure of healthcare systems to micro-level comparisons of individual treatment options.

As in all problems of efficient resource allocation, two important tenets are that resources are limited but the potential demands on those resources are not. Specific challenges in healthcare include an ageing population and the associated increase in the burden of disease[4], technological advances and the development of new (and almost always more expensive) treatments[5]. This means that no healthcare system, however it is constituted, will ever be able to meet all the demands of the people it serves. This requires us to make trade-offs between the various goals we wish to achieve, leading to the notion of an opportunity cost. This represents the benefits forgone from not allocating resources to a specific purpose and arises both when additional money is committed to the healthcare system and when money is reallocated within it.

Two separate paradigms exist in which health economic evaluations can be conducted. The first is based on welfare economic theory[6] and views healthcare in exactly the same light as any other resource allocation decision. Its three central tenets are, firstly, individualism, in that consumers are deemed to be the best judges of their own welfare and behave in a utility maximising manner. This means that the impacts of changes are best judged as the person them self perceives them to be. Social welfare thus becomes the sum of the individual utilities of the people contained within that society. Secondly, consequentialism, in that utility is generated only by the consumption of goods and services. Finally, the underlying theoretical aim of welfare economics is to derive decision rules that enable us to rank all possible states of the world according to their total utility.

Specifically, these mean that each alternative choice should be judged as to whether it represents a potential Pareto improvement in social welfare, i.e. the overall balance of those

who gain and lose from a policy is positive. This means no preferential weight is given to changes in health outcome from a given policy, as opposed to any other utility changes. This perspective leads naturally to the use of cost-benefit analysis as an evaluative tool[7]. This involves assigning an economic value to each outcome of a policy decision, necessitating techniques to assign a monetary value to health and other changes that may result.

The second framework is based on an extra-welfarist approach, and looks at healthcare decisions as problems of constrained maximisation[8]. This approach will take a series of inputs (budget limitations, a defined maximisation objective etc) and look to maximise an objective function subject to those constraints. This leads naturally on to the process of cost-effectiveness analysis, where we study the incremental cost needed to generate a fixed additional unit of a defined benefit. In health economics, at least at a treatment selection rather than service structure level, we would typically (though not exclusively) have a purely health related objective function, subject to a budget constraint. In this form, cost-effectiveness analysis has become the most commonly used evaluative tool in health economics[9].

Pharmacoeconomics and drug development

Pharmacoeconomics refers to the branch of health economics specifically associated with pharmaceuticals. As a result, one of its principal uses is during and following the developments of new drugs, or the expansion of existing drugs into new indications, as they apply for regulatory approval and market access.

The standard development pathway for a new pharmaceutical involves four clinical phases, after preclinical research has found a promising candidate. It is now commonly thought of as containing two successive learning-confirming cycles[10]. Phase I trials typically involve a small number of healthy volunteers, with the intention to study the pharmacological effects of the drug in humans (which may differ greatly from the predicted responses from preclinical animal testing) and to determine the safety profile of the drug by studying the relationship between dose and adverse drug reactions. Data will also be gathered on the pharmacokinetic properties of the drug and, in those rare cases where phase I studies are conducted in real patients (e.g. terminal cancer studies where the treatment could be expected to make healthy volunteers ill), data may also be gathered on pharmacodynamics and early indications of effectiveness.

Phase II studies will take place on patients with the intended disease or condition for treatment. They are still relatively small scale and are primarily dose response studies, examining the link between dose and both efficacy and adverse drug reactions. The aim is usually to recruit the smallest number of patients necessary in order to gather sufficient information to appropriately design a phase III trial, or alternatively to decide the drug is not sufficiently efficacious (or has sufficiently serious adverse drug reactions so as not) to merit further study. Thus, phase I and early phase II studies complete the first learning-confirming cycle, as we first find the largest dose that can be administered without causing harm (the maximum tolerated dose), then confirm whether that dose provides some measurable therapeutic benefit (minimum therapeutic dose).

The second learning-confirming cycle begins in the later stages of phase II, as attempts are made to first optimise dosing regimens, looking for one or more doses which have promising harm-benefit balances, then obtain confirmation of this efficacy in phase III trials, and establish, for each patient subgroup, which of these doses has the optimal harm-benefit balance. Phase III trials are large scale trials of efficacy, designed to support applications for regulatory approval. They attempt to provide definitive evidence of efficacy, ideally comparing to the current gold standard treatment for a given condition, though comparisons against placebo and suboptimal alternatives do also occur. They take the form of randomised controlled trials (RCTs), interventional and prospective studies where, following patient recruitment, each person is randomised to one of the arms of the trial and, where possible, blinded to treatment allocation[11].

A number of different types of phase III trial can be performed, depending on the form of license the company intends to apply for. The first is a superiority trial[12], designed to show the new drug has superior efficacy to either placebo or the standard existing treatment. Improvement in a given outcome will be assessed through formal statistical hypothesis testing, and it will need to be proven that any improvement is both statistically significant and clinically meaningful. Another form of trial is a noninferiority trial[13],which is assessed in the same way but with the addition of a noninferiority margin i.e. the drug is shown to be no more than a pre-specified (and clinically insignificant) amount worse than the current treatment. The third form of trial is an equivalence trial, where we wish to confirm there are no statistically significant differences in efficacy between the new and old treatments, tested by whether the confidence intervals for the relevant parameters lie within a specified

equivalence margin. It used to be the case that almost all new drug submissions would require support from multiple phase III trials, but it is becoming increasingly common for applications to be based on a single study.

Finally, phase IV trials take the form of post marketing surveillance, and are conducted once a drug has already been approved for use. They study the long term safety profile of the drug over a large patient population (pharmacovigilance), and gather any additional necessary data that the time horizon of the phase III trial was too short to acquire. Modifications to the administration protocol may be made at this stage, and drugs may ultimately be withdrawn if sufficiently serious adverse reactions are identified.

As a result of this relatively structured licensing process, and the subsequent use of the same data to promote the adoption of the product and its subsequent reimbursement, there is a greater degree of uniformity surrounding questions of resource allocation in pharmacoeconomics than in many other health economic areas e.g. other health technologies, system changes, public health interventions etc. As regulatory authorities (FDA, EMA etc) have set evidential standards for market authorisation, it is likely that approximately equivalent amounts and types of data will be gathered at the same stage across different drugs and trials. Whilst these standards (efficacy, often based on short term endpoints) are not those desired for reimbursement decisions (comparative effectiveness, long term outcomes), they are nonetheless closer than the evidence available in many non-pharmacological interventions. There are a number of issues that may make interpretation of the results obtained difficult, for example the fact that RCTs may not represent standard care, as they are often carried-out in controlled environments, and thus the results obtained may not be entirely applicable in practice (something which has led to the development of pragmatic trials). Nevertheless, this commonality of design has led to the development of a number of standard techniques for use in pharmacoeconomic evaluation.

Economic evaluation

In most of what follows, the primary interest is in microeconomic decisions between individual treatment alternatives, which generally involve quantifying the total costs and benefits (health or otherwise) associated with each alternative option, and using these as the basis for decision making in a cost-effectiveness framework. The aim is to maximise the allocative efficiency of the system (directing resources at the areas which will produce the

greatest gain) at a given fixed level of technical efficiency (maximising the output from a given combination of resources in a certain area)[14]. Changes in technical efficiency will change an assessment of allocative efficiency, but it is often assumed that changes in allocation will not affect technical efficiency. Since costs and benefits are being measured in different units, to render these analyses practical to perform usually requires the definition of a set of decisions rules, often based on a willingness to pay. Such a threshold represents the maximum amount society is prepared to pay for a given additional unit of health.

Health outcomes can be measured in ways specific to the disease or intervention in question, usually based on a pre-existing clinical measure or scale, but there is often a desire to compare a range of conditions which may have entirely different clinical outcomes. This necessitates the use of generic measures of health which can be applied across conditions, by far the most common of which is the quality-adjusted life-year (QALY)[15]. Importantly, this is a preference-based measure based on choices and strength of preferences for different health outcomes[16]. It is constructed on the principle that health care interventions are designed to achieve two things, increasing both length and quality of life. One quality-adjusted life-year is equivalent to one year in perfect health, and then lengths of time spent in sub-optimal health are weighted according to population derived values[17]. Thus, if the health state associated with a given condition is deemed to have half the utility of being in full health, two years in this state will equate to one QALY. The total QALYs accrued by a patient over a time period is then equal to the length of time multiplied by the mean health state utility over that period[16]. The weight given to the gain of a QALY is normally assumed to be the same, regardless of any characteristics of the patients involved, and the total QALYs of different patients are assumed to be additive (even if certain people may lose QALYs as a result of a policy change), meaning the gain of one QALY by one individual is regarded as equivalent to the gains of 0.2 QALYs by five separate individuals.

The most common method of acquiring QALY values in a trial is the administration of patient questionnaires on preference-based outcome, each possible set of answers to which will equate to a given health state. An example of such a questionnaire would be the EQ-5D, a generic measure of health status created by the EuroQol group[18]. Cost-effectiveness analysis based on the QALY is often referred to as cost-utility analyses.

Costs in standard evaluations are generally considered from the perspective of the payer/funding body, though there is variation in how broad a range of costs are considered to

be applicable. For example, some analyses will consider purely health service costs; others will include social services, while others will expand still further into areas such as productivity gains as a result of health improvements. Resource use is measured through a variety of methods, including patient recall questionnaires, hospital notes and electronic records, and unit costs are attached to that resource use from the appropriate perspective.

In contrast to the medical literature, where trial evaluations generally adopt a Frequentist perspective, with p values and hypothesis tests the standard results reported, health economic evaluations are traditionally structured on something closer to a Bayesian approach. This stems from the argument that classical statistical inference (based around type I and type II error rates) is arbitrary and irrelevant to clinical decision making[19]. If one were to accept the null hypothesis in cases where a new treatment has a statistically insignificant incremental net health benefit, this would, across a range of treatments, impose additional costs on the system (in healthcare or money). Decisions should thus be made purely on the basis of the optimal probabilistic choice. The variance of the benefit is irrelevant to the correct treatment choice, though it may influence decisions concerning the collection of additional data to reduce that uncertainty[19].

Following calculation of average, per patient, costs and utilities (QALYs) in each arm of the trial, the simplest summary statistic to report is an incremental cost-effectiveness ratio (ICER), which is the amount of additional money needed to produce a given additional unit of health improvement. Thus, for two treatments A and B, with per patient costs C_A and C_B and per patient outcomes O_A and O_B respectively, the ICER for treatment B versus treatment A is given by:

$$ICER = \frac{C_B - C_A}{O_B - O_A}$$

We can then compare ICERs of multiple treatment options to see which is the most cost-effective, and compare an individual ICER to a specified willingness to pay threshold to decide if we deem it worth funding. In a similar vein, for a specified willingness to pay threshold λ, we can calculate the net monetary benefits for treatments A and B, defined as:

$$NMB(A) = \lambda \times O_A - C_A$$

$$NMB(B) = \lambda \times O_B - C_B$$

Treatments can then be ranked in order of net monetary benefit. As this involves converting health outcomes into monetary values, the ranking of treatments by net monetary benefit can be thought of as a special case of cost-benefit analysis, where health benefits are the only ones being considered.

Parameter uncertainty in a health economic evaluation is usually addressed through sensitivity analyses. These can either be univariate, looking at the change in output from a given change in input, or probabilistic, where all uncertainty is considered simultaneously. In probabilistic sensitivity analyses, each parameter in the model where there is uncertainty will be assigned a parametric probability distribution, then multiple samples will be taken from the joint distributions of these parameters and the model run for each of these parameter sets. This is possible because of the Bayesian framework of health economic evaluations, and from these simulations we can calculate summary results such as cost-effectiveness acceptability curves, which show the probability of a given treatment being cost-effective (derived from the proportion of simulations where that treatment is cost-effective) at different willingness to pay thresholds. Such a probabilistic approach is necessary in well conducted analyses, as it is important the full uncertainty present in the data should be reflected in the end results[20]. An important structural framework within which such an approach is possible is decision modelling.

Decision modelling

Decision analysis is a set of mathematical and statistical techniques designed for the analytical evaluation of decision making under uncertainty[21]. It has been used across a wide range of disciplines outside of health care, including engineering and finance. Within healthcare, it has been applied in clinical decision making, population analyses and in economic evaluations. A decision model consists of a set of mathematical relationships that, based on a set of possible inputs representing alternative scenarios under consideration, produces a series of possible sets of consequences. Each possible outcome will be associated with a probability, and costs and outcomes can be assigned. From this we can calculate the expected costs and expected outcomes of each scenario being considered, by weighting the costs and outcomes of each alternative by the probability. Such modelling is an unavoidable part of the evaluation process, in order to perform the necessary extrapolations from trial results to the required endpoints for decision making[22].

Decision models allow for the explicit incorporation of uncertainty and variability[23]. This enables them to more accurately represent the real world fact that two seemingly identical patients will react in very different ways to the same treatment. The representation of this in terms of event probabilities (and associated distributions) allows us to collate and summarise the total effects of all the uncertainties present in the data. These models can then be used to inform resource allocation decisions, based on a set of decision rules defining the allocation priorities of the funding body in question (health maximisation, equity etc). As the use of economic evaluations by health care decision makers has increased[24], so has the use of decision modelling as a means of carrying out that analysis. An example of this would be the methodology guidelines published by the UK National Institute for Health and Clinical Excellence[25].

Decision models have a number of important features which make them suitable tools in economic evaluations. First, they enable us to synthesise all the available evidence in one evaluation. The Bayesian framework in which most decision models operate enables us to use a variety of meta-analytical tools to synthesise evidence from different sources. Many different types of data may need to be synthesised in a single analysis, including resource use, quality of life and effectiveness evidence, all of which are consistent with the principles of evidence-based medicine[26]. Second, as previously stated, decision models allow for the explicit incorporation of parameter uncertainty. The use of probabilistic sensitivity analysis enables us to produce outcomes such as the probability that a given decision is the correct one, summarising all the uncertainties in the evidence base we have.

Thirdly, there is the need to compare in a single evaluation all possible treatment options, even when these may have been tested across a number of different trials and thus head to head comparisons between some of the agents may not be available. The use of Bayesian indirect treatment comparisons across these different trials enables us to produce appropriate adjusted comparisons, rather than simply relying on naïve single arm comparisons. Finally, such models enable us to extrapolate the results outside of the constraints of the trials they are based on. Specifically, the limited time horizon of trials will often not be sufficient to represent the long-term differences that may result, particularly when there is a potential difference in mortality rates between arms of the trial. These extrapolations can be modelled within a decision analytic framework, both by using available information on the natural history of the condition being studied to project people's events, costs and outcomes to a

longer horizon, and second to estimate how the differences in treatment effect between the different arms will change over an extended horizon.

Three specific types of decision analytic model are in common usage in economic evaluations. The first and simplest is a decision tree, which contains a fixed number of pathways which patients can follow, with probabilities attached to each bifurcation of the tree. Costs and utilities are assigned to each possible pathway, which can then be summed across the whole tree to obtain total expected costs and utilities. Decision trees are simple to both use and explain, but have a number of limitations including the small number of health states it is possible to model, a need to define a separate probability for each branch of the tree and no time dependency in the model, a clear limitation in a field such as medicine where we clearly would expect a patients' condition to change over time. Decision trees were among the earliest decision analytic models used in health economics, but have increasingly been superseded in recent times by more complex, and thus more realistic, methodologies.

A second, somewhat more complex modelling framework is a Markov model. These contain a fixed number of health states patients can be in at any point in time, thus making them suitable for the assessment of conditions which can be divided into discrete health states. Patients remain in a state for a fixed period of time (the defined cycle length of the model), and a transition matrix (which can either be static or vary with time) then gives probabilities of what state the person will move to in the next cycle. Models can be run for an indefinite number of cycles (often until all patients have moved to the death state), and again costs and utilities attached to both states and the transitions between states. The key structural feature of a Markov model is the memorylessness property. Formally, for a random variably X:

$$\Pr(X > m + n | X \geq m) = \Pr(X > n)$$

In practice, this means that the future course of a patients' events is defined only by their current state, and not by their past event history. This assumption significantly simplifies the computation complexity of a given problem, but again produces limitations as with many health conditions, we would expect a patient's full event history to be relevant to their likely future progression. Despite this, Markov models have grown to become probably the most frequently used modelling technique in pharmacoeconomic evaluations, with the development of additional features such as tunnel states (a set of states which contain information not just on a patient's health state but how long they have been in that state) to

counteract the limitations listed above. Whilst the number of states in a Markov model can, in theory, be expanded indefinitely to cover any plausible health states, the need to identify and specify a separate parameter for each transition between states in the model creates a ceiling on the number of states it is practical to implement.

Markov models can be run as either cohort or individual patient simulations. In cohort models, a fixed number of patients are entered into the model at the start and then fixed proportions transition to each possible alternative state at every cycle, in exact concordance with the parameters of the transition matrix. In individual patient simulations, patients are entered into the model one at a time and their transitions are randomly simulated given the parameters of the model. This individual simulation approach is more computationally time consuming, but has the advantage of being able to generate variances and confidence intervals surrounding output values.

The third and most recent model to come into prominent use in economics is the discrete event simulation (DES)[27]. These will usually take the form of a continuous time Markov process (as opposed to the discrete time Markov model above), and will be constructed as an individual patient simulation. Each simulated person will have a set of demographic characteristics and a treatment history, with probabilities of future events defined as a function of those parameters. This gives DES a greater flexibility than either decision trees or Markov models, with the ability to model a considerably greater number of states (the total number of possible states is the product of the possible numbers of states for each individual parameter, and can thus be infinite if some parameters are continuous). However, the downside is that their structure as individual patient simulations can lead to them being more computationally complex and time consuming than the other methods.

It has been claimed that discrete event simulations, whilst having many advantages, suffer from a potential lack of transparency due to both the number of calculations required and number of relationships within the model. However, if this complexity is a result of the underlying complexity of the system, then any attempt to use simpler modelling techniques would simply result in a loss of accuracy in the results obtained[28]. Complexity in model structure should be possible to overcome with sufficient clarity of reporting.

Weaknesses of traditional economic evaluations

Despite the well-developed methodology described above, there are certain limitations inherent in economic evaluations as they are usually conducted in pharmacoeconomics. Many of these stem from the use of a single or, at best, a limited set of clinical trials as the sole source of clinical data. Whilst a randomised controlled trial has a high degree of internal validity, making it a very effective method of establishing the efficacy of a drug within the constraints of the trial, it lacks generalisability and may provide considerably less information about what the real world effectiveness will be. Whilst long term follow up data from routine use should eventually be able to answer this question, it will not in most cases become available until considerably after both licensing and reimbursement decisions have been made, and may never become available for treatments that were not approved.

Substantial differences between a treatment's efficacy and effectiveness may result as a consequence of non-adherence, which is more prevalent in routine practice than in trials, a different patient casemix, including co-morbidities or co-medications, or simply lower levels of patient monitoring in routine care. Traditional health economic evaluations have no rigorous way to address these concerns, as they are dependent on the trial data as a source of effectiveness evidence. Various approaches have been suggested for the incorporation of non-adherence data into economic evaluations, including regression modelling and the explicit incorporation of non-adherence as a parameter in decision modelling, all of which necessitate the collection of reliable data on adherence alongside a trial[29]. One ad hoc method that has been used to try and address this limitation is the adjustment of the control arm of the trial via the inclusion of data from routine practice[30]. This can partially control for protocol deviations in the control arm, but since no such data is likely to be available at this stage for the new intervention drug, it is not clear if adjusting just one of the arms of a trial in this way will reduce or increase the level of bias in the analysis.

Standard economic models also struggle when attempts are made to extrapolate the results of evaluations outside of the context of the initial trial[31]. Some level of extrapolations is almost always necessary (e.g. modelling patient lifetimes from the much shorter time horizon of a trial) and accepted as standard practice in economic evaluations[22]. However, other extrapolations, such as modelling different patient populations from that of the trial, studying different dosing schedules or examining patient heterogeneity are much more difficult. The empirical nature of most health economic models essentially precludes such attempts to

extrapolate outside the boundaries of the trial evidence, or at least considerably limits the confidence with which we can do so. This is particularly a problem as both the cost and time involved in conducting phase III trials mean there is little prospect of additional trials being conducted to answer such questions.

There is also an increasing desire to perform economic evaluations earlier in the drug development process, for example at the end of phase II[32]. It would be desirable to obtain early estimates of cost-effectiveness that could be used to inform the design of phase III trials, including sample size calculations, appropriate endpoints, appropriate trial populations etc. They would also be able to suggest which treatments would ultimately not prove to be cost-effective and thus help to inform stop/go decisions during drug development.

Full scale economic evaluations are traditionally only conducted once phase III data has been gathered, and this is the first reliable estimate of cost-effectiveness that will be derived. This means that many phase III trials are conducted (at considerable expense in both time and money) on drugs that ultimately turn out not to be clinically effective (or to be toxic) and thus not cost-effective, a problem that explains a substantial part of the reason why so many phase III trials ultimately end in failure (over 50% of recent phase III trials have either been terminated early or ended in failure[33]). However, it is often deemed impossible to conduct early analyses with standard modelling techniques, as there is simply not sufficient data available at the end of phase II trials to be able to undertake a useful economic analysis. Even where such analyses are conducted, the small sample sizes available from phase II trials mean that there will be extremely high levels of uncertainty in the estimates that are generated.

The root cause of all these problems is the fact that standard economic evaluations only make use of a limited subset of the data that is collected during drug development (specifically only the cost and outcome data collected during phase III trials), something which contradicts one of the tenets of evidence based medicine, namely using all available evidence in decision making. By making use of other data collected during phase II trials, particularly pharmacological data, and combining this with the standard data collected in economic evaluations, we can look at addressing these issues.

Pharmacological modelling

Pharmacokinetics (PK) is the study of the absorption, distribution, metabolism and excretion of substances administered to the body. A pharmacokinetic model will typically produce the

concentration of a given compound (or its metabolites) in various components of the body over time. Population pharmacokinetic models look at inter-patient variability in how the body processes the drug over time. They take as inputs the dose (and administration schedule) of the drug, as well as patient demographic, genetic variations, clinical measurements and other therapeutic characteristics (e.g. co-medications), giving as outputs estimates of the pharmacokinetic parameters (absorption, clearance, volume of distribution etc) and the variables that influence these parameters.

Whilst non-compartmental pharmacokinetic models do exist, we are here most interested in compartmental models, as these can be used to predict concentration at any time point[34]. Key concepts in pharmacokinetics include bioavailability (the proportion of the administered drug available to the body), clearance (the rate at which the drug is excreted) and the volume of distribution (the apparent volume over which a drug is distributed). The structure of the model used will depend on the mode of administration of the drug. For example, with a bolus administration where the drug can all be assumed to enter the body instantaneously, less parameters will be necessary than with an oral medicine, where additional parameters will be needed to model the absorption of the drug (e.g. an additional compartment to represent the gut). Figure 1 below shows the structure of one, two and three compartment PK models, with the arrows representing drug entry, drug elimination and the movement between compartments. It becomes necessary to use two compartments when the time course of drug concentration shows two distinct phases, i.e. when the drug is absorbed by the soft tissue, and likewise for the addition of further compartments (e.g. effect compartments, a compartment or set of compartments modelled as the effect site of the drug).

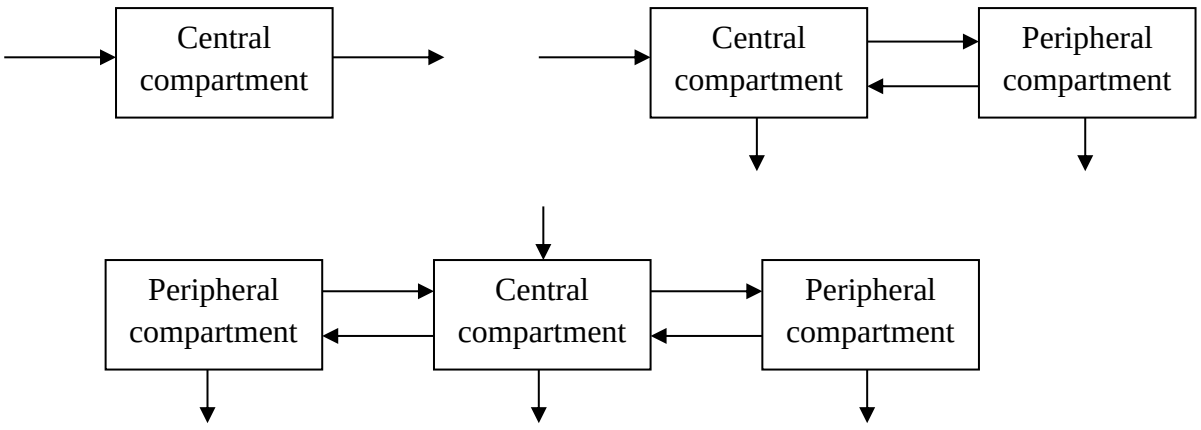


Figure 1

The simplest possible model is for a single IV bolus dose, in a one compartment linear PK model. Here the time course of the concentration of the drug is given by:

$$C(t) = \frac{D}{V_d} \exp(-k_{el}t)$$

where D is the dose, V_d the volume of distribution and k_{el} the eliminations rate constant. In the slightly more complex example of an orally administered dose, the concentration over time is given by:

$$C(t) = \frac{F.D.k_a}{V_d(k_a - k_{el})} [\exp(-k_{el}t) - \exp(-k_a t)]$$

where D is the dose, F the bioavailability, V_d the volume of distribution, k_a the absorption rate constant and k_{el} the elimination rate constant. In these cases, it is possible to analytically derive an equation for concentration over time, but in more complex situations we have to rely on numerical methods. This is particularly the case where nonlinear PK models are used.

Pharmacodynamics (PD) is the study of the effects of a drug on the body, and quantitatively, the link between drug concentration, patient characteristics, and pharmacological effect. PD models are usually time independent and thus describe an equilibrium relationship between concentration and effect. The simplest useful model in pharmacodynamics is the E_{max} model:

$$E = \frac{E_{max} \times C}{EC_{50} + C}$$

where E is effect, C the concentration of the drug (or the active moiety), E_{max} the maximum attributable effect of the drug and EC_{50} the concentration of drug producing 50% of the maximum effect. More generalised formulae also exist to cover situations which are not accurately modelled by the equation above. For example, we can create a new set of models by introducing a shape parameter, the Hill coefficient γ giving:

$$E = \frac{E_{max} \times C^{\gamma}}{EC_{50}{}^{\gamma} + C^{\gamma}}$$

We can also introduce additional compartments into the model, representing the concentration of a drug at multiple effect sites. Models can, in principle, be made increasingly complex to cope with whatever the true concentration-effect relationship is, but

in practice we soon hit limits introduced by both computational efficiency and parameter identifiability.

Pharmacokinetic-pharmacodynamic (PKPD) models, by combining these two together, thus give a model for pharmacological effect over time (figure 2). Population-PKPD models, considering inter-individual variability in patient response, describe the time course of pharmacological effect, conditional on a patient's characteristics (age, sex, bodyweight, co-medications etc) and dosing schedule. In some cases, PK and PD models will be derived separately from different data sets and then combined together, and in other cases the parameters for the model will be co-estimated from the data set. Importantly for our purposes, regardless of which of these approaches is used, the models are predictive and can be used to simulate predicted effects in specific patient subgroups, for different doses and dosing schedules.

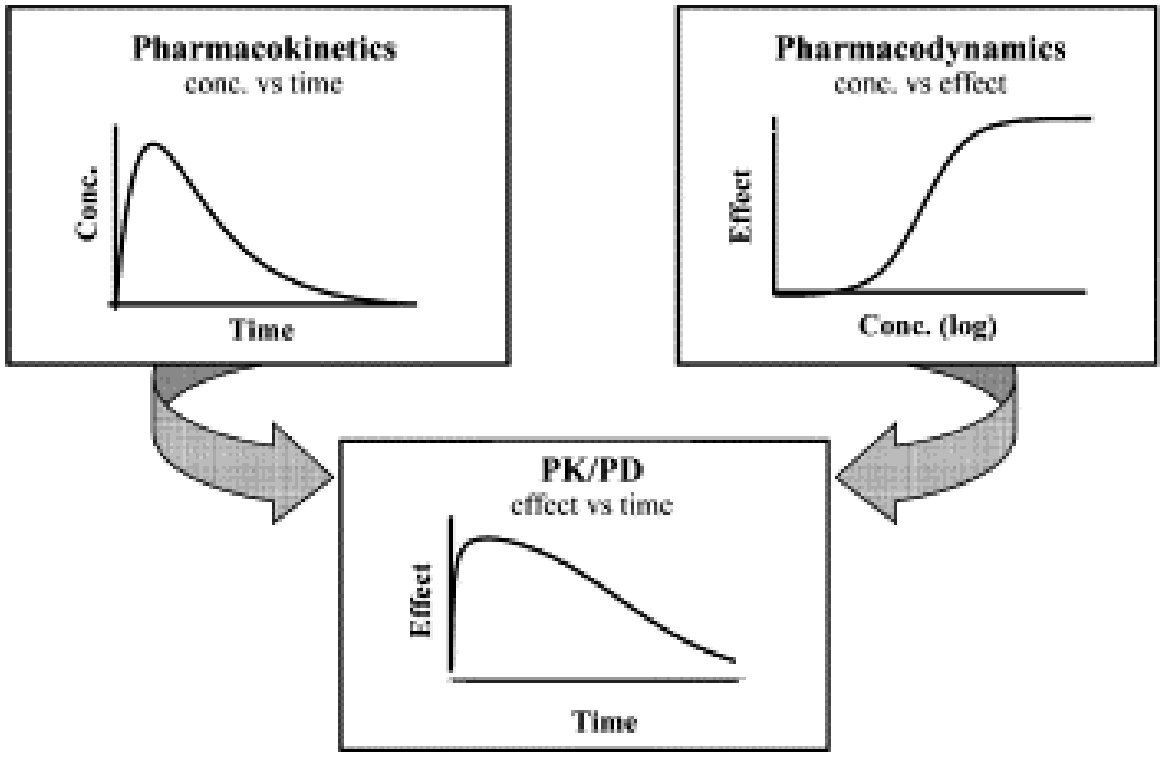


Figure 2

Pharmacogenetics is the study of how genetic variations, and the accompanying variability in metabolic pathways, can affect individual response to drugs, both in terms of efficacy and adverse events. This is becoming an increasingly important area of study, as it links in with the field of personalised medicine, where treatments can be tailored to a person's individual

characteristics. PKPD models can be extended to include how genetic variations can affect individual drug response. The standard way this is implemented into PKPD models is to have genetic information included as a covariate, i.e. the parameters in a PKPD model vary according to the genotype of the patient.

Population-PKPD models are often constructed from data collected during phase II clinical studies, as this provides the optimum balance between the number of patients studied and the intensity of monitoring (phase I trials are generally not sufficiently large, in phase III trials there is not sufficient monitoring). These models are already used extensively to study the properties of the drug, for example the forgiveness to missed doses, or the prediction of the appropriate steady state concentration of the drug. Such modelling has been used to inform licensing decisions for different patient subgroups and is accepted as evidence to support such decisions by licensing bodies (e.g. FDA)[35]. A more recent development is the use of these models as the basis for simulation studies.

Clinical trial simulation

Clinical trial simulations based on population-PKPD models are now widely used during the drug development process[36]. These will typically consist of simulating the outcomes of various possible phase III trials, and using these outputs to optimise trial design, looking at such areas as dosing schedules, patient demographics and the effects of patient non-adherence on effectiveness.

A full CTS model is made up of three components; an input-output model, a covariate distribution model and a trial execution model[37]. The standard input-output model in most CTS is a population-PKPD model, though additional models may also be added at this stage, such as models of expected disease progression over time. The covariate distribution model contains the details of the population being modelled, and will need to specify values and distributions for all parameters that affect the PKPD model. Finally, the trial execution model represents both the intended trial protocol and expected deviations from that, e.g. patient withdraw, non-adherence, missing data etc. Whilst there will often be considerable parameter uncertainty in the PKPD models used, this uncertainty can be explicitly incorporated into the simulations through parameter distributions, and thus quantified in terms of its impact on the results.

One of the uses of CTS is in assessing the effect protocol deviations will have on the outcome of a trial[38]. During the planning of a trial it is necessary to assume it will proceed according to a prescribed experimental design. If this protocol is appropriately implemented and adhered to, sufficient power should be obtained to estimate treatment efficacy. However, deviations from this protocol will often lead to a trial not having a sufficient sample size or appropriately constructed sample to reach these levels, meaning the trial will result in failure. Whilst this can be compensated for by deliberately overpowering trials, this is an inefficient and unreliable process. We can however, given a sufficiently well specified trial execution model, use individual patient CTS to estimate the combined effects of a multitude of potential protocol deviations, and thus quantify the expected change these will cause in the outcome of the trial[39].

Taking non-adherence as an example of a protocol deviation we may wish to correct for, various different modelling approaches can be adopted. A simple method is to use a two state Markov model, where a person's probability of missing a given dose is contingent only on whether they missed their previous dose. This can then be made more precise by the addition of further covariates to the model, linking non-adherence rates to other patient demographic characteristics. Another approach is to have non-adherence modelled as a function of a person's disease and demographic characteristics, or events (either due to lack of efficacy or adverse drug reactions) they may experience whilst on the medication, creating a feedback cycle between the PKPD and adherence models. Discontinuations can also be modelled in this way, or we can perform a survival analysis of any pre-existing data we may have, creating a parametric function for discontinuation over time. Whilst in any individual case it will be necessary to perform some form of model selection to decide which of these options to use, the important point here is that all of these approaches are consistent with and implementable within a PKPD based CTS[38].

Another use is in the optimisation of the design for phase III trials, to aim for greater efficiency through utilising the minimum number of patients/doses to make the same inferences. Only a finite number of potential dosing schedules will have been studied in phase II, and often at this stage new information on dose-exposure-response relationships will be discovered which may, as an example, imply dosing based on body-surface area will be more effective than fixed dosing. Simulations can attempt to both optimise the algorithms for such dosing and provide support for the implementation of a dosing regimen that may not have been tested in exactly that format in phase II. As a post phase III extension of this,

simulations may also have a role as supporting data for new drug submissions to regulatory bodies, particularly in cases where only a single trial has been conducted[40]. A simulation model that accurately predicts real world results can be used as evidence the pharmacology of the new agent is well understood.

Subgroup analyses are particularly important at phase II, to study heterogeneity in patient response. If it is the case a new drug has considerable variability in response, then the inclusion of a large number of patients in a phase III trial in whom there is no incremental effect will not only reduce the effective sample size of the trial, but also introduce extra variability in results through the confounding effect of these patients who ideally should not have been included. However, since one of the key goals of phase II trials is to make decisions based on the minimum possible numbers of patients, we are extremely unlikely to have sufficiently many patients in any subgroup of interest to perform traditional subgroup analyses. CTS can fill this is gap by allowing us to simulate the differential outcomes of various group of interest. A key example of when this might be desirable is where there are differences in response based on genetics, as this may well only become known during phase II trials when PK and PD data is collected, thus any analysis of the effects of this will need to be post-hoc as the phase II trial was not designed with it in mind. All such simulations would be hypothesis generating, and then ideally these hypotheses would be tested during the phase III trial.

A final key use of CTS is in internal strategic decision making by drug companies during the development process. Previously we have discussed ways of optimising the design of future trials, but it is also important to identify early on those drugs where a phase III trial is liable to give a negative response, and prevent the unnecessary waste of both time and money in conducting it[41]. If simulations can provide an increase in the success rate of clinical trials, by preventing some trials that would ultimately prove negative from starting without simultaneously preventing trials that would give a positive result from being conducted, this will greatly improve the efficiency of the process.

Weaknesses of PKPD modelling and traditional clinical trial simulations

The use of CTS has already been shown to lead to an increase in the likelihood of achieving a trial's prescribed objectives[42]. However, there are a number of limitations to their usefulness in pharmacoeconomics as currently conducted. First there is the difficulty in deciding on the

complexity of the PKPD model which should be used. In many cases simple linear models with small numbers of compartments will not generate good fits to the data, which we can only achieve by introducing much more complex features, such as nonlinear dynamics or multiple compartment chains. Whilst these models may well fit the data better, we lose the ability to relate the structure of the model to the biological features of the underlying system; for example we may not have any indication as to what each compartment of the model actually represents, if anything. Given that much of the current use of such models is to gain a greater understanding of the underlying dynamics of the system, these models will not be able to fulfil that purpose.

Physiologically-based pharmacokinetic-pharmacodynamic (PBPKPD) models are an attempt to fill this gap, where each compartment in a model has anatomical and physiological significance, representing specific tissues or sets of tissues[43]. The derived parameters then directly correspond to the underlying biological processes of the system. Such models are still at an early stage of their acceptance into mainstream use, and there remaining problems with both computational complexity and sufficiently accurate model fitting, but nonetheless they represent a promising potential developmental path[43].

A significant limitation of standard clinical trial simulations is that the endpoints they produce as outputs are in many cases not the same outputs as those by which the drug will ultimately be judged. First, often PKPD models only have intermediate endpoints or biomarkers as their output. These may well be useful for informing various aspects of the earlier development process, but are limited somewhat in their usefulness by not linking to longer term clinical outcomes. Even when the final output of a simulation is a useful clinical endpoint, it will generally only be one out of a number that are clinically important. The proportion of published PKPD models which allow you to simulate all relevant efficacy and safety endpoints is extremely small. Whilst it is possible to use different PKPD models and thus different simulations to prioritise different endpoints, joint optimisation across these different simulations is exceedingly complex, meaning that in practice simulations are usually only used to optimise one primary endpoint.

Whilst this may be sufficient for traditional sample size calculations, it misses a significant part of the picture. As many countries now require evidence of cost-effectiveness in order to secure reimbursement, decision making should also be informed by this important endpoint. A drug may well appear to be effective, compared to standard care, on the basis of

simulations and may well indeed be so in reality. However, the less cost-effective the drug is, the less it will be utilised and thus the less money it will ultimately make for the company. The ability for phase III trial decisions to be assessed on the basis of cost-effectiveness could help to ensure potential reimbursement likelihoods are also taken into account in decision making.

Standard clinical trial simulations do not, however, provide a great deal of help in making such decisions, as they do not create a sufficiently large set of parameters, with only some effectiveness parameters and no resource use data being generated. Whilst it is possible to perform very simple cost-effectiveness analyses from the data they generate, essentially using only the measures of effectiveness simulated and only the expected cost differential in drug price and simulated event treatment cost, and such analyses are better than nothing in informing decisions, they are nonetheless extremely unreliable as a data source to use. Thus, there is currently no easy way to synthesise the available economic data with CTS of efficacy[39].

Finally, at present, CTS have not been put to great use post phase III, with the exception of adjusting for protocol deviations. This is generally because the data from phase III trials, given that it is of much larger samples and longer follow up, has tended to be used for all subsequent analysis. This ignores the fact that the considerably more intensive monitoring undertaken during phase II means that the this data, whilst it can clearly never replace the need for phase III trials, can still be extremely valuable in supplementing it even when phase III trials have already been fully completed.

PKPDPE modelling

Pharmacokinetic-pharmacodynamic-pharmacoeconomic (PKPDPE) modelling is the synthesis of these two techniques (pharmacological and economic modelling) in an attempt to overcome the inherent shortcomings each separately possesses.

Drug-disease trial models are already recommended by the FDA Critical Path document as a valuable tool to improve the productivity of the drug development process[44], and a number of publications have suggested the extension of these models to incorporate economic information[45,46]. Despite this, virtually no publications have actually attempted to implement this idea in practice. This is principally due to the cross-disciplinary nature of such an idea, meaning only a limited number of people have sufficient knowledge of both fields to make

this practical. This has not, however, stopped interest in the underlying principle of the idea from increasing. Many companies are now considering implementing such models, but without any publications emanating from this it is difficult to say how far the idea has progressed.

The general structure of a PKPDPE model is that estimates of efficacy (and other relevant and available outcomes, e.g. adverse drug reactions) derived from clinical trial simulations can then be used as inputs to traditional decision-analytic pharmacoeconomic models. The PKPD model would be used to simulate the full expected outcome of a clinical trial, and then this would be fed into a health economic model as if it were the data from a trial. The stochastic nature of the output of the CTS matches neatly with the distributions necessary for probabilistic sensitivity analysis in economic evaluations. A full population-PKPD model should contain all the necessary data to simulate different patient subgroups and compare them in terms of cost-effectiveness. Previously published or prospectively collected cost and QALY data can then be used complete the necessary parameter sets to run the model. Ultimately we should be able to construct exactly the same set of results we would for clinical trial data, and whilst there would obviously be considerable uncertainty, we are able to quantify that uncertainty into easily interpretable summary measures such as the probability of cost-effectiveness.

The first advantage of this is that we are able to run these simulations and obtain estimates of cost-effectiveness as soon as a PKPD model has been built, i.e. during phase I/II. As that model is improved by the collection of additional data we are able to rerun our simulation to refine our cost-effectiveness estimates. We are thus able to address all the standard questions that CTS are used for, but from the standpoint of cost-effectiveness rather than simply effectiveness. Such an approach might reasonably be hoped to facilitate more accurate decision making than is currently the case.

Such an approach is also compatible with value of information (VOI) analysis, which is an alternative method for optimising trial designs. This looks at the value attached to the reduction of uncertainty in a model or certain parameters within that model, measured as the benefits lost through making the wrong allocation decision due to uncertainty[47]. This value can then be compared to the cost of gathering additional data to reduce that uncertainty. Trial designs can then be optimised based on the value of the information they collect. The practical implementation of such an approach takes as its input the output of the probabilistic

sensitivity analysis of an economic evaluation, meaning it can also be conducted based on our simulated data sets. Thus, with such an approach we are able to implement full scale VOI analyses at an earlier stage than is currently possible.

Whilst all the analyses we have discussed will require the collection of certain additional data, such as the costs and health state utilities of various events, this is all data the company will ultimately be required to find anyway, it simply brings the need for it earlier in the process. The same is true for the building of the economic model itself; it does not necessarily represent additional work, merely a change in when that work takes place. Importantly, in such analyses, we are only interested in the predictive power of the PKPD model being used, and not in pharmacological interpretation of the results. This means we are free to make use of the best fitting model we have available to us, even if it does not immediately seem to be a direct representation of the underlying biological processes. This provides a clear distinction between situations where we wish to develop our understanding of the underlying dynamics of the system (where PBPKPD models may offer the most promising path for development), and our situation where we wish to use such models for purely simulative purposes.

The mechanistic nature of PKPD modelling also helps address one of the weaknesses of purely empirical health economic models. Even after the completion of phase III trials, simulations can still have an important role in aiding extrapolations outside the scope of the trial, be they extrapolations of time, population or treatment. Such extrapolations are often criticised in the health economics literature as introducing additional uncertainty to a data set, the effect of which cannot be quantified in standard evaluations[48]. However, once again the stochastic nature of CTS, combined with probabilistic analysis in health economics, allows us to quantify any uncertainty we may introduce from our use of simulated as opposed to trial data.

In summary, PKPDPE modelling has the potential to address shortcomings in both standard pharmacological and economic modelling, and can do so primarily using tools that have already been well developed in the other field. However, considerable work on validating this approach would be necessary, looking at comparisons of simulated and real world results, to ascertain whether the predictions obtained are sufficiently accurate as to be useful to decision makers. However attractive the principle of the idea may be, this is of no use if its practical implementation would not produce real changes in decision making, be they internal drug company decisions (economic viability, price setting) or those by reimbursement bodies.

Intentions

Three principal questions around PKPKDPE need to be addressed; namely the feasibility, validity and applicability of such an approach. We here present two case studies designed to assess these questions, and consider how such an analysis might be conducted in practice.

The first case study in chapter 2 is primarily retrospective, and based on the use of rituximab for the treatment of follicular lymphoma. By simulating the results of already conducted trials, we can compare how well these simulations match the trial derived results, and thus how much confidence we feel we can have in the process. We then provide one example of a prospectively simulated economic analysis, both as a demonstration of how such analyses might be conducted, and as a basis for comparison when that trial is ultimately completed. Since it is virtually impossible to construct any reasonable way to internally validate these models, the best approach seems to be for a number to be constructed which can then be compared with real data as it is collected. We will then be able to address if there are specific features of the modelling approach or conditions being studied that appear to lead to greater congruence between trial and simulated results.

Chapters 3-5 present a full prospective analysis of a currently topical condition, namely drugs for stroke prophylaxis in atrial fibrillation. Chapter 3 presents an economic model of atrial fibrillation, used for an economic comparison of warfarin and dabigatran, derived fully from the published literature and based on a large scale multinational phase III clinical trial. All the cost and quality of life data used were available prior to the completion of the trial, meaning it would have been available to use in a phase II economic analysis had one been conducted.

Chapter 4 is an indirect comparison of currently available treatment alternatives, performed so as to include all currently available treatments (warfarin, dabigatran, rivaroxaban, apixaban and aspirin) in the analysis. This is a key step in any attempt to gain early indications of cost-effectiveness, as it is inherently a comparative process; a drug can only ever be cost-effective in comparison to another. This form of meta-analysis falls within the same tenets of evidence based medicine, namely making use of all available data, in which we have justified our use of CTS.

Finally, chapter 5 is a complete prospective population-PKPD simulation, based on genotype-guided warfarin dosing, which is used as an input to these previously constructed models, in order to gain estimates of cost-effectiveness in situations where trials have not be conducted

and are unlikely ever to be conducted. Whilst some trials of warfarin pharmacogenetics have taken place, they have been too small scale to provide definitive results, and only a small number of the myriad possible warfarin algorithms have been tested. The sheer number of potential algorithms makes it certain that only a small number will ever be tested in trials, all of which are likely to be comparisons only with warfarin as opposed to the newer anticoagulants available, leaving simulations as the only possible way to assess the cost-effectiveness of these. The particular simulation used here has three stages. First, a CTS simulates the outcome of an important biomarker (International Normalised Ratios - a measure of the clotting properties of the blood), a meta-analysis of published literature is performed that links this intermediate outcome to clinical endpoints, and these clinical event rates are fed into an economic model.

Feasibility will be assessed simply through the process of constructing these models; specifically whether simulations provide sufficient data to fully supply necessary clinical parameters, and whether it is possible to obtain the necessary cost and quality of life data at these earlier stages. Validity will ultimately be judged by comparing the simulated results we obtain with real world data as it is collected. Some such data is already available for comparison, and more will become available as some of the trials which we are currently simulating are completed. Finally, we will need to consider what specific questions we are ultimately able to address with these techniques, and the best methods for implementing them in practice.

Chapter 6 concludes by giving an overall summary of the conclusions of the thesis, its strengths, limitations and the avenues for further research, as well as exploring the place of the techniques discussed within the context of the expanding field of model-based drug development.

Chapter 2

Mechanism-Based Approach to the Economic Evaluation of Pharmaceuticals: Pharmacokinetic/Pharmacodynamic/Pharmacoeconomic Analysis of Rituximab for Follicular Lymphoma

Summary

Introduction: Economic value is an important consideration during all phases of the drug development process. Building on previously published work describing a mechanism-based economic modelling approach that incorporates data obtained during phase II clinical studies on the relationships between dose, exposure and response, this chapter describes case studies of rituximab for the treatment of follicular non-Hodgkin's lymphoma based on this methodology.

Methods: A population pharmacokinetic and pharmacodynamic model was used linking serum rituximab concentration to progression-free survival, to simulate the effectiveness of rituximab in various clinical contexts. These served as inputs to economic models of follicular lymphoma, based on National Institute for Health and Clinical Excellence (NICE) appraisals, to assess the cost-effectiveness of rituximab. Results were compared with trial-based estimates from the NICE appraisals. In a further analysis, the results of an on-going trial were simulated to generate predictions of cost-effectiveness.

Results: The analyses suggest an acceptable degree of concordance between simulation- and trial-based estimates of cost-effectiveness. For first-line and maintenance therapy, deviations of £2,099 and £1,355 per QALY, respectively, from trial-based incremental cost-effectiveness ratio estimates of £8,290 and £7,721 per QALY gained would not affect reimbursement decisions. The probability of rituximab-containing regimens being cost effective at £20,000 and £30,000 per QALY thresholds was 1 for both first-line and maintenance therapy in both simulated and trial-based analyses.

Discussion: These analyses demonstrate the feasibility of mechanism-based economic analyses, which may have applications during drug development to the following: (i) directing future research based on the cost of reducing uncertainty; (ii) assessing subgroups, dosing schedules and protocol deviations; and (iii) informing strategic research and development and pricing decisions.

Introduction

As value-based pricing extends from smaller markets (e.g. Australia, Canada and Sweden) to the UK market[49], which has international significance as a pricing reference, pharmaceutical industries are becoming increasingly mindful of the importance of the early determination of cost-effectiveness. Although pricing has always influenced decisions throughout the drug development process, the role and methods of pharmacoeconomics are less well defined in the early phases[50,51].

Various papers have previously proposed a mechanism-based modelling approach to help inform clinical phases of drug development while also directly considering the resource constraints of payers of healthcare[45,46]. Our approach is reliant on the mathematics of the structural relationships between dose, internal exposure and response (which describe the pharmacokinetics and pharmacodynamics of a drug), and how uncertainties in related parameters may be explained by clinical, demographic and other covariates.

Pharmacokinetics is the study of drug absorption, distribution, metabolism and excretion. These processes determine the fate of a drug when administered to humans, and may be quantified by determining concentrations, usually in blood plasma, urine or saliva. PK models are mathematical representations of the time course of drug concentration. This fluctuates according to several factors, including the dose; dosage form and dosing schedule; use of concomitant interacting medications; and patient demographic and pathophysiological characteristics, e.g. liver and kidney function and adherence. Pharmacodynamics is the study of the pharmacological effects of drugs, and in particular the relationship between concentration and response. PD models thus link drug concentration and response, again taking into account individual patient characteristics. Combined PK/PD models quantify the time course of drug effects and may be used to simulate clinical data for a given population, drug and dosing schedule. This enables us, through our simulations, to study different populations or dosing schedules from those in which research has previously been carried out[52-55].

Population pharmacokinetic/pharmacodynamic (PK/PD) models are being recognized increasingly for their utility in expediting the development process by the pharmaceutical industry and are accepted by regulatory authorities[44,56]. Applications include simulations of drug effectiveness and safety; optimization of trial design in the latter phases of development; and exploration of the effects of different dosing regimens and patient demographics[36]. As

these simulations are usually based on the data from phase II trials, there is necessarily considerable uncertainty surrounding the parameters of such models. Nevertheless, they provide a framework to quantify both the available information and the uncertainty therein.

Using the output of population PK/PD models, to serve as the input of economic decision analyses, allows these same questions to be addressed, but from the perspective of cost-effectiveness rather than merely effectiveness. We explored the feasibility of the method in proof-of-concept case studies of rituximab, a chimeric monoclonal antibody used in the treatment of follicular non-Hodgkin's lymphoma. Rituximab was chosen because of the availability of the necessary data – PK and PD models, as well as economic evaluations for a number of indications – in the public domain. The pharmaceutical industry would typically have comparable PK and PD data available following phase II studies.

The first case study relates to the use of rituximab as a maintenance treatment following induction chemotherapy for patients with recurrent or refractory stage III or IV follicular non-Hodgkin's lymphoma[57]. The second explores rituximab as a first-line treatment[58]. In both cases, PK/PD-based decision analyses are compared with trial-based decision analyses that informed decisions made by the National Institute for Health and Clinical Excellence (NICE) for the use of rituximab in England and Wales. A third case study forecasts the clinical and economic outcomes of the PACIFICO trial (Purine-Alkylator Combination in Follicular Lymphoma Immuno-Chemotherapy for Older Patients; ClinicalTrials.gov identifier: ISRCTN99217456)[59], a phase III randomized controlled trial currently recruiting elderly patients to compare rituximab, fludarabine and cyclophosphamide (R-FC) and rituximab, cyclophosphamide, vincristine and prednisolone (R-CVP) induction chemotherapies for the treatment of stage II-IV follicular lymphoma.

Methods

Pharmacokinetic Simulations

A two-compartment linear PK model was utilized, in which body surface area (*BSA*) and gender were the significant covariates for both clearance (*Cl*) and central volume of distribution (*Vc*)[60]. The population PK model was based on a phase II study of rituximab in 102 patients with rheumatoid arthritis[60]. Individual (*Cl, Vc*) parameters are predicted from population parameters (*CL, VC*) according to:

$$Cl = CL \times \left(\frac{BSA}{1.79}\right)^{\theta_{BSA\_CL}} \times (1 + \theta_{SEX\_CL})$$

$$Vc = VC \times \left(\frac{BSA}{1.79}\right)^{\theta_{BSA\_VC}} \times (1 + \theta_{SEX\_VC})$$

where θ are the effects of covariates on the parameters. *CL* and *VC* were sampled from a log-Normal distribution with coefficients of variation (*CV*) of 28.2% and 12.3%, respectively[60]. Simulations of rituximab serum concentrations assume BSA follows a Normal distribution, BSA(m^2) ~ N(1.71, 0.2) for men and BSA(m^2) ~ N(1.91, 0.2) for women[61], and that patients were equally likely to be male or female. A mixed (additive and proportional) residual error model was used[60]. Mean concentrations since the last infusion (*Cm*) were calculated as follows:

$$Cm(t) = \frac{\int_{t_n}^{t} C(\varphi)d\varphi}{t - t_n}$$

where *C* is the concentration of rituximab at time *t*, t_n is the time of the most recent rituximab infusion and φ is a dummy integration variable. Mean rituximab concentration since the last infusion was used as this was found to give more realistic PD parameters than using the actual concentration at a given time[62].

Pharmacodynamic Model

Following Ternant et al.[62], an exponential hazard model was used to relate progression-free survival (*PFS*) at time t to mean rituximab concentration:

$$PFS(t) = e^{-\lambda_{max}\left(1-\frac{Cm(t)^{\gamma}}{Cm_{50}{}^{\gamma}+Cm(t)^{\gamma}}\right)t}$$

where λ_{max} is the median hazard in the absence of rituximab, Cm_{50} is the *Cm* value leading to a 50% decrease of λ_{max} and γ is a shape factor. Cm_{50} and γ are assumed to be the same in all situations, whilst λ_{max} varies depending on the adjuvant chemotherapy[62]. With no data on the inter-individual variability of Cm_{50}, the random effects were assumed to be log-Normally distributed (50% CV). In the original publication by Ternant et al., this model was built by fitting it to data from two studies of rituximab pharmacokinetics and validated by using it to predict the results of two additional separate studies[62]. Values of λ_{max} were obtained by fitting exponential survival curves to the PFS data from relevant trials of treatment without rituximab[63-65].

γ and Cm_{50} were both assumed to follow Normal distributions and the uncertainty in parameter estimates were accounted for as follows: $\gamma \sim N(0.486, 0.052)$ and $Cm_{50} \sim N(35.06, 2.2)$.[61] For the PACIFICO case study, where patients who respond follow a different treatment path from those who do not, responding patients were sampled as follows: $\gamma \sim N(1.5, 0.11)$ and $Cm_{50} \sim N(18, 0.87)$[62].

Simulated PFS data for cohorts of 1,000 patients receiving rituximab were generated from PK/PD analyses relating to each case study. This was done by first generating a set of 1,000 hypothetical patients by sampling from the aforementioned parameter distributions. PK simulation, using these patient profiles and the dosing schedule from the appropriate trial, generated 1,000 individual mean concentrations since the infusion. PFS data were subsequently simulated by propagating the PK data through the PD model.

Rituximab Maintenance Therapy

The pivotal European Organisation for Research and Treatment of Cancer (EORTC) 20981 trial randomized patients to either 3-weekly cycles of cyclophosphamide, doxorubicin, vincristine and prednisolone (CHOP) or rituximab plus CHOP (R-CHOP) as induction therapy[63]. Responding patients were randomized to either rituximab maintenance (375 mg/m^2 BSA per cycle), or no further treatment. λ_{max} was calculated by fitting an exponential function to 1500-day PFS data for patients randomized to CHOP, obtained by digitizing the published Kaplan-Meier PFS curve using Engauge Digitizer V4.1[66].

First-Line Treatment

Evidence supporting the use of rituximab as a first-line treatment for stage III–IV follicular lymphoma comes from the M39021 clinical trial[64], in which patients received either eight 3-

weekly cycles of cyclophosphamide, vincristine and prednisolone (CVP) chemotherapy, or rituximab (375 mg/m^2 BSA per cycle) plus CVP. The value of λ_{max} was calculated using digitized PFS data relating to CVP as described in the previous section.

PACIFICO

Values of λ_{max} for patients responding (complete or partial response) to induction with CVP and fludarabine and cyclophosphamide (FC) were derived following digitization of published data from a trial that randomized patients to each for a maximum of eight cycles, the E1496 trial[65]. This was identified from a search of PubMed, the Cochrane Central Register of Controlled Trials database and the American Society of Hematology database for trials that compare FC and CVP for the management of follicular lymphoma. λ_{max} for non-responders was obtained from the proportional difference in hazard rate between responders and non-responders[63], and assumed to be the same, irrespective of prior treatment.

Health Economic Modelling

Models that determine the incremental cost-effectiveness, expressed as cost (in £; year 2004 values for the maintenance and first-line models and year 2010 values for the PACIFICO model) per QALY gained, of regimens including rituximab were based on manufacturer submissions to NICE[57,58]. These adopted a UK NHS costing perspective. Reported methodologies were replicated, and models analysed according to the parameter values in table 1 (P43). Incremental cost-effectiveness ratios (ICERs) were calculated (difference in total expected costs divided by the difference in total expected QALYs), as were cost-effectiveness acceptability curves that present the probability of the rituximab-containing regimen being cost effective at different cost-effectiveness thresholds. For all models, utility values were calculated using the EQ-5D questionnaire from a UK study of a cohort of 222 patients with follicular lymphoma[67], and both costs and benefits were discounted at 3.5% per annum[25].

Rituximab Maintenance Therapy

The NICE report of rituximab for recurrent/refractory follicular lymphoma describes a Markov process approach, with three states representing PFS, progressed follicular lymphoma (PFL) and death[57] (figure 3a, P42). All patients enter in the PFS state, and transit among these states (1-month cycle duration) according to probabilities determined from observed values and modelled extrapolations of the EORTC 20981 trial[63]. For the first 24

months, overall survival (OS) and PFS were taken from Kaplan-Meier data, obtained by digitizing the published OS and PFS curves. Extrapolations beyond 24 months were based on Weibull parametric functions fitted to 1500-day data[67], a Weibull distribution being preferred to a log-logistic, log-Normal, exponential or Gompertz distribution following model selection by the Akaike information criterion and Schwarz's Bayesian criterion[67]. The number of patients in the PFL state was taken as the difference between the numbers in the PFS and dead states. In the PK/PD-based health economic analysis, trial-based PFS data were replaced with simulated data.

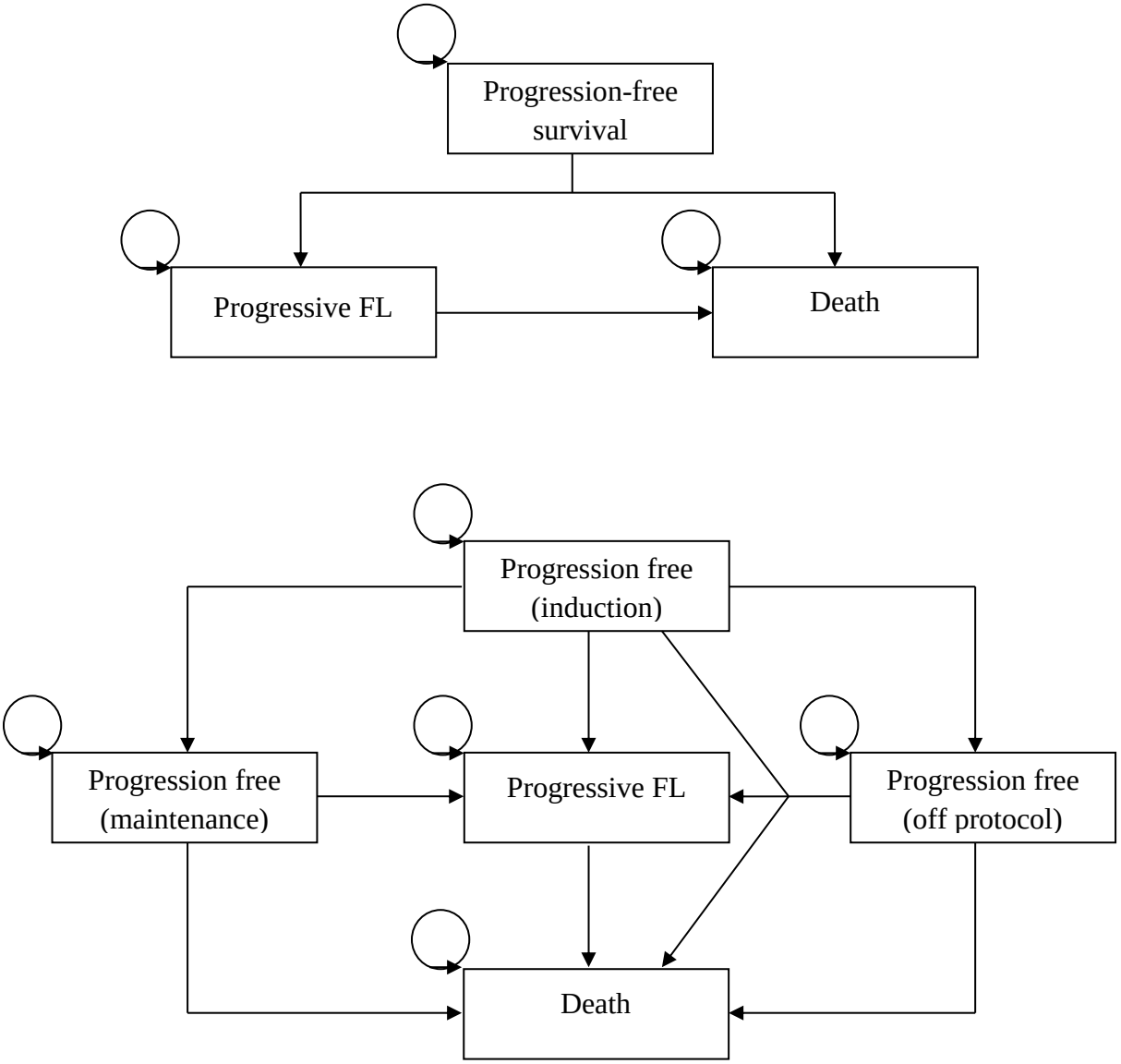


Figure 3

Table 1 - Parameters and distributions for models of first-line and maintenance therapy

Parameter	**Maintenance therapy**		**First-line therapy**	
	Mean (estimate for sensitivity analysis)	PSA distribution	Mean (estimate for sensitivity analysis)	PSA distribution
λ_{max}	0.0451	NA	0.0269	N(0.0269, 0.001)
Overall survival				
Weibull scale parameter	0.5274	N(0.5274, 0.0367)	NA	NA
Weibull intercept – control	7.7576	N(7.7576, 0.1257)	NA	NA
Weibull intercept – treatment	8.1555	N(8.1555, 0.1257)	NA	NA
Transition probability [PFS to PFL]	NA	NA	0.0170 (0.0085, 0.0255)	N(0.0170, 0.001)
Costs				
PFS state	£28.67 (86)	NA	£32.33 (16.16, 64.66)	Gamma(32.33, 1)
PFL state [routine management]	£86 (28.67)	Tri(60, 86, 141)	NA	NA
Treatment of PFL state – control	£285.77 (571.54)	[58]	£193.33 (96.67, 386.66)	Gamma(193.33, 1)
Treatment of PFL state – treatment	£286.27 (572.54)	[58]	£193.33 (96.67, 386.66)	Gamma(193.33, 1)
Drug – control	£0	NA	£330.96	Tri(220.96, 420.96)
Drug – treatment	£7739	1325 × N(5.9254, 0.263)	£10 110.24	Tri(7110.24, 13 110.24)
Drug administration – treatment	£502	86 × N(5.9254, 0.263)	£800 (400, 1200)	Tri(600, 800, 1000)
SAE – control	£7.05 (0)	Uniform(0, 2000) × Beta(1, 166)	NA	NA
SAE – treatment	£188.90 (0)	Uniform(0, 2000) × Beta(30,	NA	NA

		137)		
NSAE – control	£124.11 (62.05, 248.22)	1.443 × Uniform(0, 200)	NA	NA
NSAE – treatment	£138.01 (69.01, 276.02)	1.605 × Uniform(0, 200)	NA	NA
Other				
PFS utility	0.805 (0.618)	Trimmed N(0.805, 0.018)	0.805 (0.618)	Trimmed N(0.805, 0.018)
PFL utility	0.618 (0.805)	Trimmed N(0.618, 0.056)	0.618 (0.805)	Trimmed N(0.618, 0.056)
Duration of treatment benefit	5 years (2, 30)	NA	5 years (3, 25)	NA
Time horizon of analysis	30 years (4, 50)	NA	25 years (5, 50)	NA
Cost discount rate	3.5% (0, 6)	NA	3.5% (1.5, 6)	NA
QALY discount rate	3.5% (0, 6)	NA	3.5% (1.5, 6)	NA

λ_{max} = maximum value of median hazard; N = Normal; NA = not applicable; NSAE = non-serious adverse event; PFL = progressed follicular lymphoma; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; SAE = serious adverse event; Tri = triangular.

The costs and utilities associated with each health state, taken from the NICE report[57], are reproduced in table 1 (P43). PFS costs are based on expected numbers of visits to health clinics, whilst PFL costs are based on the cost and frequency of expected treatments following progression. Drug costs, taken from the British National Formulary[68], were based on BSA and assumed wastage from unused vials. Costs for individual patients accrue according to time spent in each state, number of cycles of chemotherapy, duration of treatment effect and the incidence of adverse events. QALY calculations were based on utilities in the PFS and PFL states and patients' survival.

The analytic time horizon was set to 30 years and the duration of treatment benefit from rituximab was set to 5 years, whereupon transition probabilities reverted to those of the control group[57].

First-Line Treatment

The same three-state health economic model formed the basis of a NICE report of rituximab as first-line treatment[58] (figure 3a, P42). Transitions from PFS to PFL were determined by a log-logistic extrapolation of the trial data, following model selection as described in the earlier section on rituximab maintenance therapy[58]. Transitions from PFS to death were taken from all-cause mortality data[69], and from PFL to death by a parametric extrapolation of data from the Scottish National Lymphoma Group (SNLG) Vanguard database[58], which records survival data pertaining to second-line chemotherapy. An exponential survival model fitted to these data was used to obtain transition probabilities, following the methodology of the NICE submission[57,58]. For the economic evaluation based on trial simulation, observed PFS data were replaced by simulated data from the PK/PD analysis.

Costs and utilities for each health state, and the costs of treatment, were taken from the published report[58]. PFS costs consisted of drug and administration costs for eight cycles of chemotherapy, and regular maintenance costs; PFL costs were based on treatments patients are expected to receive following progression.

The treatment benefit of rituximab was assumed to persist for the lifetime of the model, with the analytic time horizon set to 25 years, in line with the NICE model[58].

PACIFICO

In PACIFICO, patients are randomized to receive either eight cycles of R-CVP or four cycles of R-FC followed by four cycles of rituximab alone, with treatment terminated after four cycles if there is no response to induction[58]. Patients who continue to respond after eight cycles are given rituximab maintenance every 3 months for 2 years or until progression. A five-state Markov model was developed to predict the cost-effectiveness of regimens used in the trial[67] (figure 3b, P42). The model specifies states for PFS, PFL and death, with the PFS state subdivided into patients undergoing induction, those who have responded (and hence were assigned to the maintenance group), and those who did not respond (and hence were taken off protocol). Parameter estimates are presented in table 2 (P47).

The probability of remaining in the PFS state during each 1-month cycle was calculated directly from the PK/PD simulation. In line with a previous analysis[58], transition probabilities between PFS and death were calculated from all-cause mortality data[69], and an exponential survival model was applied to data from the SNLG database for transitions from PFL to death.

Response rates for induction were taken from a trial comparing FC and CVP chemotherapies[65], and adjusted for the effect of rituximab using data from a meta-analysis of published trials of both induction chemotherapies[63,65,70-73], identified from the literature review. Response rates from each study were weighted by population size, yielding a mean probability of 0.558 for non-responders being converted to responders by the addition of rituximab. The assumption of this probability being constant, irrespective of chemotherapy regimen, was tested in sensitivity analysis.

Table 2 - Parameters and distributions for PACIFICO model

Parameter	**R-CVP value**	**R-FC value**	**PSA distributions**	**Source(s)**
λ_{max}	0.02238	0.01943	N(Mean, 0.001)	[65]
λ_{max} – non-responders	0.025	0.025	N(0.025, 0.001)	[57,65]
Response rate – 4 cycles	0.8984	0.9381	Beta(107, 12) Beta(108, 7)	[65] and meta-analysis
Response rate – 8 cycles	0.9426	0.9735	Beta(112, 7) Beta(112, 3)	[65] and meta-analysis
Transition probability [PFL to death]	0.0170	0.0170	N(0.017, 0.001)	[58]
FC induction mortality rate	NA	0.0025	Uniform(0, 0.005)	Meta-analysis
Cost of PFS state	£31.25	£31.25	Gamma(31.25, 1)	[57]
Cost of PFL state	£405.23	£405.23	Gamma(405.23, 1)	[57]
Mean cost per dose				
Rituximab	£1,325.01	£1,325.01	None	[75]
Fludarabine	NA	£726.42	None	[75]
Cyclophosphamide	£9.47	£9.47	None	[75]
Vincristine	£26.46	NA	None	[75]
Prednisolone	£3.98	NA	None	[75]
Proportion neutropenic	0.3108	0.5890	Beta(442, 980) Beta(493, 344)	Meta-analysis
Cost of managing neutropenia	£3,773	£3,773	Gamma(3773, 1)	[74]
Proportion thrombocytopenic	0.0338	0.1195	Beta(48, 1374) Beta(100, 737)	Meta-analysis
Cost of managing thrombocytopenia	£1634	£1634	Gamma(1634, 1)	[74]
Proportion anaemic	0.0169	0.0335	Beta(24, 1398) Beta(28, 809)	Meta-analysis
Cost of managing anaemia	£1,634	£1,634	Gamma(1634, 1)	[74]

Proportion with infection	0.0422	0.0681	Beta(60, 1362) Beta(57, 780)	Meta-analysis
Cost of managing infection	£344	£344	Gamma(344, 1)	[74]
Proportion suffering other events	0.0253	0.0227	Beta(36, 1384) Beta(19, 818)	Meta-analysis
Mean cost of managing other events	£1,326	£1,326	Gamma(1326,1)	[74]
Mean number NSAE	2	2	4 × Beta(2, 2)	[57]
Unit cost of NSAE	£86	£86	Gamma(86, 1)	[57]
Health state utility – PFS	0.805	0.805	1-Gamma(117, 0.00166)	[57]
Health state utility – PFL	0.618	0.618	1-Gamma(46.5, 0.00821)	[57]
Cost discount rate	3.5%	3.5%	None	[25]
Outcome discount rate	3.5%	3.5%	None	[25]
Model horizon	30 years	30 years	None	Assumption
Duration treatment benefit	5 years	5 years	None	[57]

λ_{max} = maximum value of median hazard; FC = fludarabine and cyclophosphamide; N = Normal; NA = not applicable; NSAE = non-serious adverse event; PFL = progressed follicular lymphoma; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; R-FC = rituximab, fludarabine and cyclophosphamide; R-CVP = rituximab, cyclophosphamide, vincristine and prednisolone.

Adverse events were assumed to be additive for induction and maintenance regimens, and obtained from the same literature review. Where available, the proportions of patients experiencing neutropenia, thrombocytopenia, anaemia, infection or other adverse events related with R-FC or R-CVP regimens were extracted directly. Where data were only available for FC or CVP, incidences were increased by 54%, 68%, 27%, 51% or 69%, respectively, as calculated from a meta-analysis of relevant studies[63,65,73]. Data were weighted by sample size to derive the final estimates (table 2, P47**).**

The number of patients undergoing induction at each cycle was taken from the number in the PFS-induction state, and used to calculate the expected cost of induction therapy. Together with published costs for adverse events[74], the expected total cost per patient was calculated. The overall cost consisted of drug costs (taken from the British National Formulary[75]), administration costs, routine maintenance costs based on health state and the costs of adverse events.

Patients were assumed to enter the model at 60 years of age. The analytic time horizon was set to 30 years and, following the model for rituximab maintenance therapy[67], the duration of treatment benefit from rituximab was set to 5 years, whereupon transition probabilities reverted to those of the control group.

Sensitivity Analysis

Univariate and probabilistic sensitivity analyses (PSA) were conducted to explore parameter uncertainty. For the maintenance and first-line models, ranges and distributions were replicated from the NICE reports[57,58] wherever possible (table 1, P43). Where these were not available, and for the PACIFICO model (table 3, P50), univariate analysis involved varying parameters by ±50%. For PSA, proportional data were modelled with Beta distributions. Unless indicated otherwise, clinical parameters were assigned Normal distributions. Cost data were modelled with Gamma distributions and, when variances were not known, they were assumed to be equal to the mean (table 2, P47). In all cases, 2000 simulations were performed.

Table 3 - Sensitivity analysis for the model of the PACIFICO trial

Parameter	**Sensitivity range**	**Parameter estimate**	**Incremental cost per QALY gained (£)**
Base case	NA	NA	19,950
Monthly cost of PFS health state	−50%	£14.34	19,672
	+50%	£57.34	20,228
Monthly cost of PFL health state	−50%	£185.89	20,982
	+50%	£743.54	18,916
Utility values	NA	PFS utility = PD utility = 0.618	28,217
		PFL utility = PD utility = 0.805	21,721
Mortality rate for PFL state	−50%	0.0085	25,562
	+50%	0.0340	18,842
Cost of rituximab (per dose and administration – first line and maintenance)	−50%	£662.51	16,103
	+50%	£2,650.02	23,819
Cost of fludarabine (per dose and administration) - first line	−50%	£363.21	14,801
	+50%	£1,452.84	24,984
Cost of adverse events	−50%	Neutropenia: £1,886.50 Thrombocytopenia: £817 Anaemia: £817 Infection: £172 Other: £663	17,812
	+50%	Neutropenia: £7,546 Thrombocytopenia: £3,268 Anaemia: £3,268 Infection: £688 Other: £2,652	22,069
Discount rates	NA	Costs = 0%; QALYs = 3.5%	18,788
		Costs = 6%; QALYs = 3.5%	20,620

		Costs = 3.5%; QALYs = 0%	13,058
		Costs = 3.5%; QALYs = 6%	25,855
Time horizon of analysis	NA	10 years	49,796
Treatment benefit	NA	2 years	23,658
		30 years	19,437
Response rates (at 4 and 8 cycles, respectively) assumed to be equal in treatment arms	Rate for R-FC taken from R-CVP	(0.8984, 0.9426)	21,466
	Rate for R-CVP taken from R-FC	(0.9381, 0.9735)	20,394
Increase in response rate due to rituximab – FC arm	−50%	0.279	20,561
	+50%	0.837	19,319
λ_{max} for non-responders	NA	0.02	21,714
		0.03	18,742
λ_{max} for FC chemotherapy	NA	0.02238	31,181
λ_{max} for CVP chemotherapy	NA	0.01943	35,474
FC induction mortality	None	0	18,229
	Double	0.005	21,390

λ_{max} = maximum value of median hazard; CVP = cyclophosphamide, vincristine and prednisolone; FC = fludarabine and cyclophosphamide; NA = not applicable; PFL = progressed follicular lymphoma; PFS = progression-free survival; R-CVP = rituximab, cyclophosphamide, vincristine and prednisolone; R-FC = rituximab, fludarabine and cyclophosphamide.

One particular source of uncertainty that is not, in this context, captured in a standard PSA is the accuracy of cost estimates. Because current unit costs are used in order to estimate the probability of the intervention being cost effective at some future point, the uncertainty as to what costs will be at that time is greater than that indicated by the uncertainty in current cost estimates. This problem was mitigated to an extent by increasing the variance of the cost estimates in our PSA. In the PACIFICO example, we thus assessed the effect on the probability of cost-effectiveness of doubling the variance for each of the costs.

Value-of-Information Analysis

The expected value of perfect information (EVPI) for R-FC versus R-CVP was calculated using Monte Carlo simulation following methods described elsewhere[76], to provide an estimate of the upper limit on returns on future research. This was based on an annual UK patient population of 1,204 for 10 years[77].

Congruence of Modelling Approaches

For the economic models of first-line and maintenance rituximab, comparisons were made between the NICE reports[57,58], and the analyses in which the trial-based PFS data were substituted for simulated data. The congruence of both approaches was assessed by calculating the difference in mean output values (and associated 95% central range [CR]); by comparing the cost-effectiveness acceptability curves; and by comparing the frequency of dissimilar pairings of simulated ICERs. For the latter approach, the proportion of simulations that yielded a different outcome (i.e. whether or not they were deemed to be cost effective) was plotted for each threshold value of cost-effectiveness. This provides a representation of the width of the range of thresholds over which the PK/PD-based simulation results may be untrustworthy (>5% difference).

PK analyses were performed using NONMEM 7.1.0 (ICON, Ellicott City, MD)[78]. PD and health economic modelling were both done using R, version 2.9.2[79].

Results

Rituximab Maintenance Therapy

The simulation-derived ICER for R-CHOP versus CHOP is £9,076 per QALY gained, compared with £7,721 from the trial-based economic analysis (a 17.6% deviation; table 4, P55). Based on PSA, the mean difference in ICER between simulated and trial data is £1,896 per QALY gained (95% CR −£2,086, £5,773). Figure 4 (P54) presents the cost-effectiveness acceptability curves for simulated- and trial-based analyses, and the likelihood that the PK/PD-based analysis leads to a decision about the regimen's cost- effectiveness, which is different from trial-based analysis. Between cost-effectiveness thresholds of £3,247 and £16,256 per QALY, more than 5% of simulation pairs give different results. However, both approaches give a decision uncertainty of zero at the NICE threshold range of £20,000 to £30,000 per QALY[25].

The deviation in ICER values is related to the mean times in the PFS state, which differ from trial-based estimates, over a lifetime, by 4.7 weeks (2.6%) and 17.7 weeks (18.9%) for the rituximab and control groups, respectively. Differences in mean life expectancy are 1 day (0.1%) and 3 days (0.2%), respectively, and QALYs deviations are 0.063 (1.5%) and 0.363 (10.9%). Differences in modelled total cost are £1,120 (5.2%) for the rituximab group and £2,633 (17.9%) for the control group (table 4, P55).

First-Line Treatment

Using PFS derived from the PK/PD model, the ICER for first-line R-CVP versus CVP is £10,389 per QALY gained (table 4, P55). This compares with a trial-based estimate of £8,290 per QALY gained, a difference of 25.3%. Based on PSA, the mean difference in ICER between simulated and trial data is £2,062 per QALY gained (95% CR −£74, £3,972). At the NICE cost-effectiveness threshold, both models yielded a probability of 1.0 for R-CVP being cost effective. More than 5% of simulations give different results between cost-effectiveness thresholds of £6,168 and £13,872 per QALY (figure 4, P54).

Differences in mean PFS are 10.1 weeks (3.9%) and 8.8 weeks (5.8%) for the rituximab and control groups; and differences in mean life expectancy are 11.6 weeks (2.3%) and 10.4 weeks (2.6%), respectively. Differences in total QALYs were 0.199 (3.5%) and 0.062 (1.4%) for the rituximab and control groups, respectively, and differences in total costs were £1,193 (5.8%) and £1,368 (13.7%).

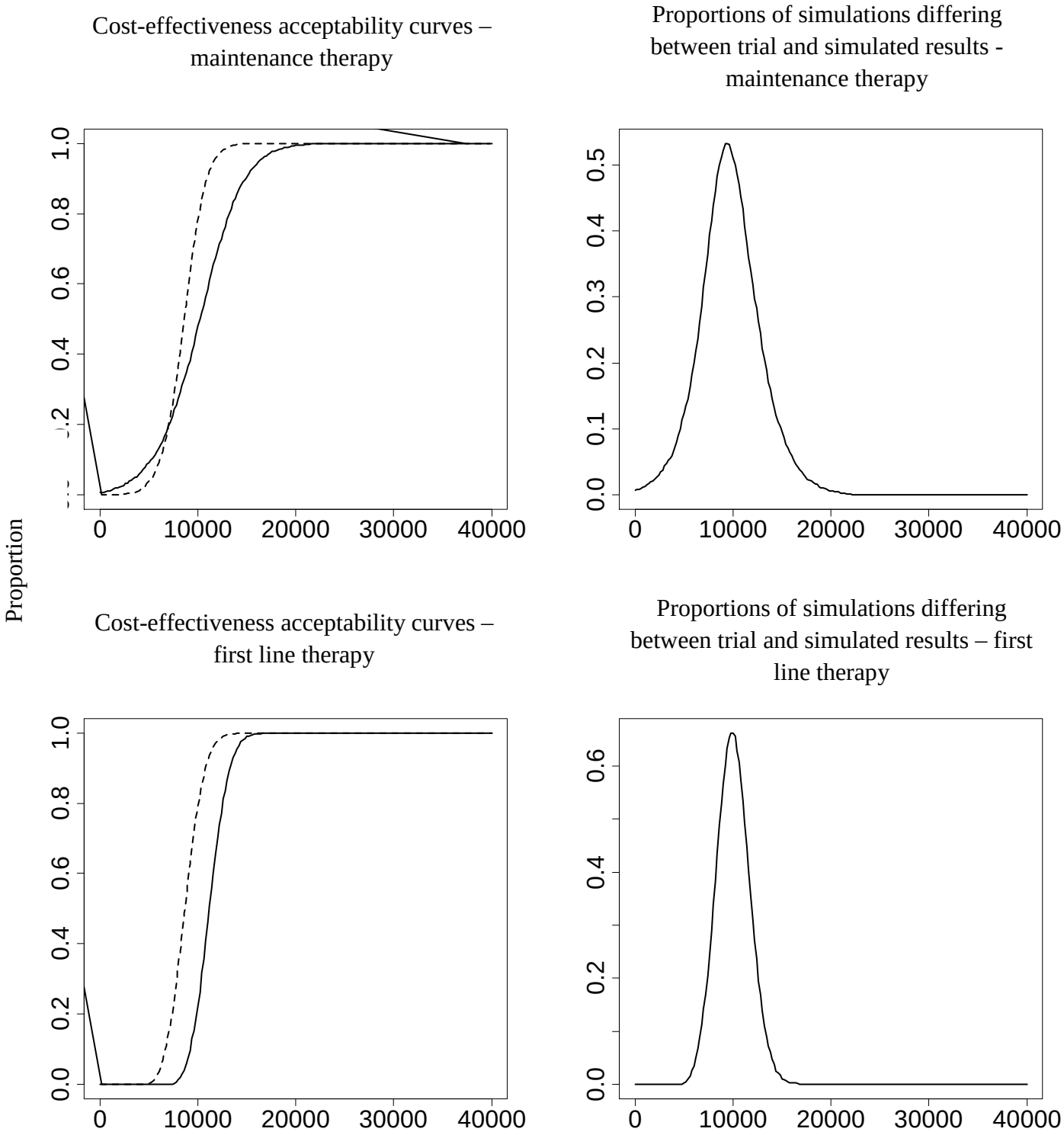

Cost-effectiveness acceptability curves – maintenance therapy
Proportions of simulations differing between trial and simulated results - maintenance therapy
Cost-effectiveness acceptability curves – first line therapy
Proportions of simulations differing between trial and simulated results – first line therapy
Proportion
0.0
0.2
0.4
0.6
0.8
1.0
0.1
0.3
0.5
0
10000
20000
30000
40000
Cost-effectiveness threshold (£/QALY)

Figure 4

Table 4 - Comparison of the results of the simulation- and trial-based economic evaluations of maintenance and first-line use of rituximab, and of the PACIFICO simulated trial[b]

Value	**Maintenance therapy**		**First-line treatment**		**PACIFICO**[a]
	Simulation	Original	Simulation	Original	Simulation
Costs – treatment (£)					
Study drug	8,241	8,241	10,910	10,910	25,016
Adverse event	327	327	NA	NA	2,746
PFS state	4,023	3,858	1,651	1,630	1,702
PFL state	10,137	9,182	9,069	7,807	12,326
Total	22,728	21,608	21,630	20,437	41,790
Costs – control (£)					
Study drug	0	0	331	331	19,991
Adverse event	131	131	NA	NA	546
PFS state treatment	4,319	3,833	1,062	891	1,634
PFL state treatment	12,905	10,758	9,952	8,755	13,013
Total	17,355	14,722	11,345	9,977	36,184
Benefits – treatment					
Mean life expectancy [undiscounted]	6.599	6.600	9.392	9.616	10.659
Median survival	6.249	6.221	9.685	9.801	9.538
Mean time in PFS state [undiscounted]	3.507	3.417	4.829	5.024	6.361
Total QALYs	4.288	4.225	5.512	5.711	4.982
Benefits – control					
Mean life expectancy [undiscounted]	5.398	5.409	7.815	7.615	10.162
Median survival	5.273	5.214	6.959	6.788	8.988
Mean time in PFS state [undiscounted]	2.139	1.799	3.088	2.919	5.642
Total QALYs	3.696	3.333	4.522	4.460	4.701
Incremental analysis – treatment and control					
Incremental cost (£)	5,373	6,886	10,285	10,370	5,606
Incremental life-years gained	1.011	1.000	1.094	1.327	0.331
Incremental QALYs gained	0.592	0.892	0.990	1.251	0.281

Incremental cost per life-year (£)	5,315	6,885	9,401	6,929	16,937
Incremental cost per QALY (£)	9,076	7,721	10,389	8,290	19,950

a. In PACIFICO, treatment is R-FC; control is R-CVP. b. Year of costing is 2004 for maintenance and first-line treatment and 2010 for the PACIFICO analysis. NA = not applicable; PFL = progressed follicular lymphoma; PFS = progression-free survival; R-CVP = rituximab, cyclophosphamide, vincristine and prednisolone; R-FC = rituximab, fludarabine and cyclophosphamide.

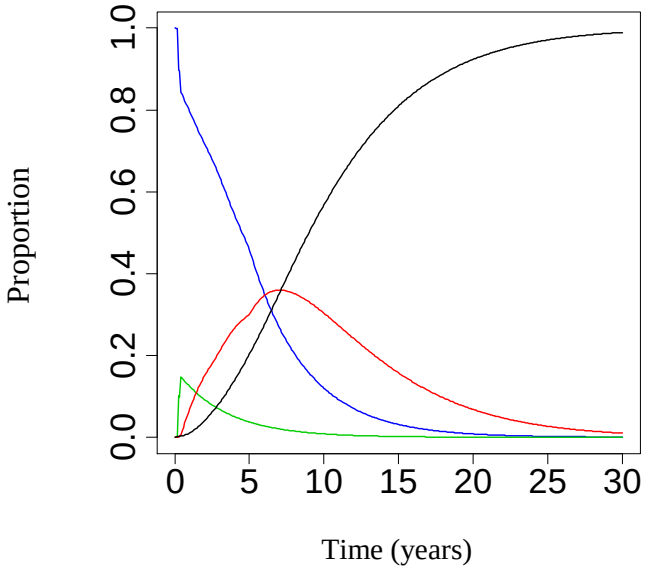
PACIFICO

Figure 5 (P58) shows the proportion of patients in each state over time, as predicted by the simulation of PACIFICO. At 3 years, in the R-FC arm of the trial, 72.2% of patients are in the PFS state and still on protocol; 3.9% are in the PFS state but are off protocol; 16.0% are in the PFL state; and 7.8% have died. The corresponding percentages for the R-CVP arm are 64.9%, 6.9%, 20.1% and 8.1%, respectively. Compared with R-CVP, R-FC increased PFS at the planned 3-year follow-up (primary endpoint) by 4.3%, from 71.8% to 76.1% (hazard ratio 0.822; $p < 0.01$).

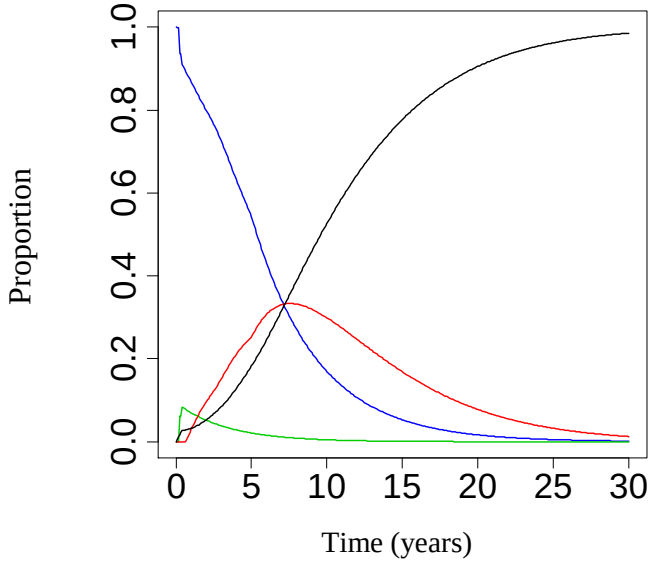
The ICER for R-FC versus R-CVP is £19,950 per QALY gained (table 4, P55). The mean (SE) simulated incremental cost and QALY gains are £5,606 (1,227) and 0.281 (0.133), respectively. The PSA suggests an 80% probability of R-FC being cost effective at the higher £30,000 per QALY threshold (figure 5, P58). For the modified PSA analysis using doubled variances for the cost data, the probability of R-FC being cost effective reduces to 76%. The results of univariate sensitivity analyses indicate that changes in the time horizon of analysis and the median hazard in the absence of rituximab (λ_{max}) for CVP chemotherapy have the most significant impact on the ICER, increasing it to £49,796 and £35,474 per QALY gained, respectively (table 3, P50).

Subgroup analyses indicate that significant covariates from the PK model had little impact on the ICER. When all patients are male (female) the ICER changes to £18,705 (£20,656) per QALY gained; and when mean BSA is reduced by 0.1 (increased by 0.1), the ICER changes to £19,279 (£20,423) per QALY gained.

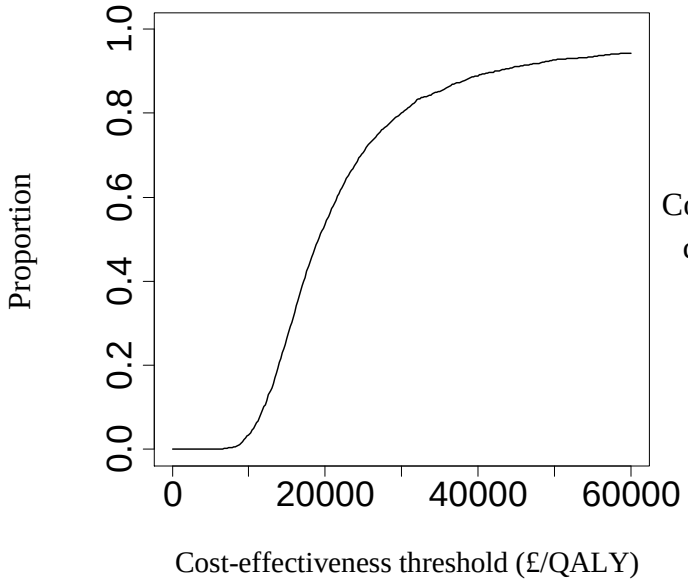
The UK EVPI for the simulation is £3,916,472 if one QALY is valued at £30,000.



Clinical simulation – R-CVP



Clinical simulation – R-FC



Cost-effectiveness acceptability curve – R-FC versus R-CVP

Figure 5

Discussion

We conducted cost-effectiveness analyses, based on PK/PD/pharmacoeconomic (PE) models, to calculate incremental costs per QALY gained for different rituximab-containing regimens for follicular lymphoma. Our choice of rituximab as a case study was facilitated by previously published population PK/PD analyses and conventional economic models. These enabled direct comparisons to be made between cost-effectiveness estimates derived from simulations and from actual trials when rituximab is used as first-line or maintenance therapy. Our simulation of PACIFICO represents an example of how such an approach might be applied in practice.

The simulated results for first-line and maintenance therapy gave higher costs per QALY than the original data, but have sufficient accuracy to result in the same economic decisions, given that the ICERs are well below NICE's cost-effectiveness threshold range of £20,000 to £30,000 per QALY. We acknowledge that factors other than cost-effectiveness will affect the reimbursement decision and so the actual decision may differ, irrespective of the probability of cost-effectiveness. Nevertheless, the deviations of our simulation-based analyses from trial-based analyses are no greater than the inter-model variation in the estimates of ICERs in other published economic evaluations of rituximab in follicular lymphoma[67,80-83].

For the maintenance rituximab model, the discrepancy between simulated and trial-based estimates results from patients spending less time in the PFS and more in the PFL state, hence accruing higher costs whilst having a lower health state utility. For the first-line rituximab model, simulated times in the PFS state are lower, thereby raising the ICER. The range of cost-effectiveness thresholds over which the simulations are an unreliable predictor (incongruence >5%) are approximately £13,000 and £8,000 per QALY for maintenance and first-line treatment, respectively. The impact this has on cost-effectiveness decisions depends on the proximity of the ICER to the cost-effectiveness threshold.

Our simulation of PACIFICO indicates an ICER for R-FC (vs. R-CVP) much closer to NICE's cost-effectiveness threshold. Therefore, there is less confidence in the expectation that R-FC will represent good value for money (i.e. decision uncertainty is high). As PACIFICO is not due to report until 2017, it remains to be seen whether the simulation is a reliable forecast of the planned trial-based economic evaluation[59].

The approach taken necessitates a number of assumptions, creating limitations in our analysis. First, models may be misspecified. The PK/PD model was developed from separate analyses of pharmacokinetics and pharmacodynamics, in different populations, and validated for rituximab in combination with other chemotherapies[62]. The relationship between exposure and response is assumed to apply to our analyses, but we have no evidence to support this. However, our deviations in PFS from observed trial results at 24 months (rituximab arms: 4.2% maintenance, 5.6% first-line) are well within the 3-52% range in the original study, which validated the PK/PD model[62]. Our economic models extrapolate short-term PFS estimates to lifetime. This is necessary to reduce bias in cost-effectiveness estimates, as interventions that impact differentially on survival will accrue costs and QALYs beyond the trial time horizon. Within the proposed analytic framework, extrapolations of this nature will depend on how well the PK/PD model is verified and on the parametric form adopted. Our disease model, which represents follicular lymphoma as a finite number of defined health states, represents a further potential source of model misspecification. In line with the reporting of clinical trials in oncology, health economic models are almost invariably based on a Markov architecture defining states of PFS, progressive disease and death. Our model follows those assessed by NICE in this respect.

Second, our reliance on the (unpublished) E1496 trial[65] to simulate PACIFICO has many caveats. The posology differed in terms of the doses administered (higher in E1496) and the dose schedules used (cyclophosphamide is no longer administered as a single dose on day 1), and no dose allowances were made for patients with reduced renal function, which collectively may have contributed to the higher mortality in the FC group of E1496[65]. Consequently, the efficacy and toxicity of chemotherapies during the induction phase of E1496 might not be generalizable to the PACIFICO trial, although sensitivity analysis suggests minimal impact of mortality in the FC group on the ICER.

Third, the model requires extensive parameterization (n = 47 for the PACIFICO model). Whereas conventional economic models are reliant on direct measures of treatment effect, an analysis based on PK/PD modelling requires many more parameters to describe the dose-exposure-response relationship. Whilst parsimony is an important consideration, the balance between models that are empirical on the one hand, and mechanistic (e.g. based on systems biology) on the other, depends on the context of use and scenarios being analysed.

A fourth limitation, the uncertainty around parameter estimates, was addressed in terms of its contribution to decision uncertainty, expressed as the probability of rituximab being cost effective given a particular cost-effectiveness threshold. It is important to also consider any additional uncertainties that may result from the fact the analysis is being performed at a very early stage. One example we consider is that the uncertainty surrounding cost data – because current unit costs are being used as a proxy for future costs – will be considerably higher than that captured by increasing the uncertainty around current unit costs. It will therefore be necessary, on a case-by-case basis, to modify the model parameters to take account of this fact. First, if the long-term inflation rates of two types of cost in the model are potentially different (e.g. if the costs of medicines increase on average at a faster rate than hospital unit costs), this will need to be incorporated in the analysis. Second, the increase in total uncertainty needs to be considered by appropriate modifications to the parameters in the PSA analysis.

A natural extension to the method presented in this paper is value-of-information (VOI) analysis, which quantifies the cost of reducing parameter uncertainty. The EVPI represents the maximum price that a pharmaceutical company (for instance) should be willing to pay for the perfect prediction of an uncertain outcome. In the case of R-FC, the global EVPI exceeds the actual cost of PACIFICO, giving no evidence the trial cost is unacceptable.

However, an EVPI alone is a very crude measure of the value of research. In practice, more sophisticated techniques such as the expected value of perfect parameter information (EVPPI) or the expected value of sample information (EVSI) should be preferred. These also allow questions such as the optimum design of a phase III trial or the key areas for future research to be addressed. We did not calculate these for this example as the focus of this paper was on demonstrating the feasibility of linking PK/PD and health economic models, rather than the particular way any such model should be used. Importantly, from the data generated in this model, one could perform an EVPPI or EVSI calculation, as appropriate for a given situation.

Our analysis lends support to the plausibility of a population PK/PD-based PE evaluation of rituximab for follicular lymphoma. Although we position such an approach for the early determination of cost-effectiveness, the case studies relied on a PK/PD model that was based on later-phase trials. Drug developers would have access to data that were not available to us in the public domain. Nevertheless, in simulating PACIFICO, the analysis demonstrates its potential utility.

Early-phase decisions are subject to considerable uncertainty, simply as a result of the paucity of data available. However, it is useful to be able to explicitly quantify the uncertainty to enable the identification of which particular parameters are principally responsible, e.g. using a Bayesian analytical framework to update *a priori* probabilities as new evidence becomes available, and VOI analyses to inform the design of phase III trials[84]. A range of methods have been proposed for incorporating economic analyses during early-phase clinical development[51], with applications covering strategic research and development decisions; pre-clinical market assessment; decisions for progressing between clinical phases of research; price determination; and reimbursement assessment. However, the population PK/PD-based approach described here is consistent with Sheiner's 'learning and confirming' paradigm for the clinical phases of drug development[10], and consequently might help facilitate a co-ordinated modelling approach across pharmaceutical industry R&D, Pricing and Reimbursement, Health Economic and Outcomes Research and Strategic Planning sections.

The pervasiveness of pharmacoeconomics – from drug development through to market access[85] – requires that cost-effectiveness is an important consideration in clinical research investment, and reliable methods for determining value for money during each of the clinical phases of drug development are necessary. The mechanistic or semi-mechanistic features of PK/PD models afford them the qualities that are desirable for integration within economic evaluations. Applications might feasibly include the impact of protocol deviations (e.g. non-adherence[86]) and subgroups (e.g. based on significant PK or PD co-variates) on cost-effectiveness, and the assessment of different doses and dose schedules[46]. However, further studies are necessary to determine the applicability of such an approach in different clinical contexts.

Preface to Chapter 3

After the completion of the initial rituximab case study, the intention was to conduct a second, to investigate a different aspect of when PKPDPE modelling might be used, in particular looking at the modelling of patient non-adherence. The models used for rituximab (both population PKPD and health economic) were comparatively simple, and the fact rituximab was already in common usage in a number of other indications meant there was published information available that would often not be the case when conducting such analyses. Our choice of warfarin for stroke prophylaxis in patients with non-valvular atrial fibrillation for this second study was motivated by the fact that that whilst warfarin is a well-established drug, meaning we have accessible PKPD and long tern efficacy data, genotyping to inform dose selection prior to warfarin initiation is still in the very early stages of development, with no definitive trials as yet having reported. Pharmacogenetics is an area which is likely to expand considerably over the coming years, making it a logical choice for a second case study.

The fact this analysis was fully prospective meant that we needed to build an extrapolative health economics model from scratch, rather than being able to base it on pre-existing evaluations as was done with rituximab. The key data source necessary was a large randomised control trial with warfarin as one of the treatment arms, which we would later be able to populate with data from our planned population PKPD simulations. The largest and most recent such study was the RE-LY study, a multinational trial of 18,113 patients comparing warfarin with dabigatran etexilate, a new oral anticoagulant. Dabigatran (and other newer anticoagulants) provide alternatives both to standard warfarin therapy and genotype-guided warfarin, meaning they must also be included in any evaluation of possible alternative anticoagulants.

The model constructed from this needed to both extrapolate the results of the trial data to a lifetime horizon and attach appropriate costs and utilities to the outcome data derived. Conducting quantitative risk-benefit and cost-effectiveness analyses of dabigatran versus warfarin would not only produce the necessary model framework but also enable us to compare our simulated warfarin results (genotype and non-genotype) not just with each other, but also with another competitor now available. Chapter 3 thus presents the model produced and the cost-effectiveness analysis of dabigatran and warfarin that resulted from it.

Chapter 3

Dabigatran etexilate versus warfarin in management of non-valvular atrial fibrillation in UK context: quantitative benefit-harm and economic analyses

Summary

Introduction: We wish to determine the incremental net health benefits of dabigatran etexilate 110mg, 150mg bid and warfarin in patients with nonvalvular atrial fibrillation; and to estimate the cost-effectiveness of dabigatran in the United Kingdom.

Methods: We conducted quantitative benefit-harm and economic analyses (from the perspective of the UK National Health Service) using a discrete event simulation model to extrapolate the findings of the RE-LY study to a lifetime horizon. Cohorts of 50,000 patients at moderate to high risk of stroke with a mean baseline $CHADS_2$ score of 2.1 were simulated with the main outcomes produced being quality-adjusted life-years (QALY) and the incremental cost per QALY of dabigatran versus warfarin

Results: Compared with warfarin, low- and high-dose dabigatran were associated with positive incremental net benefits of 0.094 and 0.146 QALYs. Positive incremental net benefits resulted for high-dose dabigatran in 94% and 76% of simulations versus warfarin and low-dose dabigatran. In the economic analysis, high-dose dabigatran dominated the low-dose, had an incremental cost-effectiveness ratio of £23,082 per QALY gained versus warfarin, and was more cost-effective in patients with baseline $CHADS_2$ score ≥ 3. However, at centres achieving good INR control, such as those in the UK, dabigatran 150mg was not cost-effective, at £42,386 per QALY gained.

Discussion: Our analysis supports regulatory decisions of dabigatran offering a positive benefit-to-harm ratio when compared with warfarin. However, we were unable to identify a sub-group for which dabigatran 110mg offered any clinical or economic advantage over 150mg. High-dose dabigatran will only be cost-effective for patients at increased risk of stroke or for whom INR is likely to be less well controlled.

Introduction

Atrial fibrillation is the most common sustained cardiac arrhythmia[87], with an estimated UK prevalence of 10% in patients aged 75 or over[88], and an associated five-fold increase in the risk of ischaemic stroke. Bed days for patients with a primary or secondary diagnosis of atrial fibrillation cost the National Health Service (NHS) £1.9bn in 2008, with outpatient and other inpatient costs (e.g. procedure and inpatient medication costs) totalling £329m[89].

Warfarin is the mainstay of oral anticoagulation thromboprophylactic therapy[90]. However, patients exhibit considerable variability in their response to warfarin, which, coupled with a narrow therapeutic range, necessitates frequent monitoring and dose adjustment to ensure optimal anticoagulation. Deviations outside the therapeutic range (international normalised ratio [INR] 2.0-3.0) increase the risks of both strokes and haemorrhagic events[91].

Dabigatran etexilate is a new oral direct thrombin inhibitor that may provide an alternative to warfarin, having the advantage of not requiring regular monitoring. In the multinational, Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) study, 18,113 patients with nonvalvular atrial fibrillation and at least one stroke risk factor were randomised to one of two doses of dabigatran (110mg or 150mg, twice daily) or dose-adjusted warfarin[92]. After a median follow-up of 2 years, the rates of the primary outcome (stroke or systemic embolism) were similar to warfarin among patients assigned the lower dose, but were lower among patients assigned the higher dose (1.11% vs. 1.71% per year; RR, 0.66; 95% CI, 0.53 to 0.82; p=0.0001). Compared with warfarin, the annual rate of major bleeding was lower among patients assigned 110mg dabigatran (2.71% vs. 3.36%; RR, 0.80; 95% CI, 0.69 to 0.93; p=0.003) but similar among patients assigned 150mg. Dabigatran was associated with higher rates of myocardial infarction, but these were not statistically significant[93].

The US Food and Drug Administration was satisfied of the positive benefit to harm balance of dabigatran, but failed to identify a patient subgroup in which the benefit-harm profile was superior for the 110mg dose compared with the 150mg dose[94], and consequently approved only the higher dose. However, both doses have been approved by other regulatory authorities, including the European Medicines Agency, which specifies 150mg bid for patients <80 years; and 110mg bid for patients ≥80 years, or as an option when the thromboembolic risk is considered to be low and the bleeding risk is high[95].

Against this background, we describe a quantitative analysis of the trade-off between thrombotic and bleeding risks – events which have differential impacts on life expectancy and quality of life – as a basis to guide clinicians' prescribing. We also develop a health economic evaluation to estimate the cost-effectiveness of dabigatran in patients with nonvalvular atrial fibrillation, given the considerable uncertainty of its cost-effectiveness within the UK healthcare setting.

Methods

We modelled the net benefits and expected lifetime clinical event rates of each dose of dabigatran and warfarin to quantify the benefits and harms of competing treatments, while accounting for parameter uncertainty[96,97]. Incremental net benefits were estimated as the difference among treatments in quality-adjusted life-years (QALYs), a preference-based outcome measure which combines two dimensions of health: life expectancy, and health-related quality of life.

In the economic analysis, the model was extended to estimate resource use and costs from the perspective of the UK NHS. The primary outcome was the incremental cost per QALY gained.

We developed a discrete event simulation model which considers individual patients, their characteristics and their experience of clinical events and outcomes according to the passage of time[27]. Following each event, a patient's health profile is updated, leading to a new set of probabilities for future events. Costs and QALYs are accrued from the patient's health states and events that occur.

For each treatment, identical cohorts of 50,000 patients were generated, each assigned an age and health profile, defined by the presence/absence (according to trial protocol[98]) of any of the following characteristics at baseline: hypertension, diabetes mellitus, congestive heart failure, prior stroke, prior transient ischaemic attack, prior myocardial infarction and prior intracranial haemorrhage (Table 5, P69)[92]. Health characteristics were assumed to be independent in the base-case analysis but a sensitivity analysis was conducted to assess the effect of correlation.

All analyses were performed in R[79]. Ethics approval was not required.

Clinical parameter estimates

A search of Medline, Embase, the Cochrane library and the FDA and ClinicalTrials.gov websites was undertaken during July 2010 to identify relevant trials of dabigatran in atrial fibrillation. We used "dabigatran", "BIBR 1048", "atrial fibrillation" as search terms, and identified three phase II trials (PETRO[99], PETRO-Ex[99] and NCT01136408[100]) and a single phase III trial (RE-LY)[92]. The phase II studies included too few patients receiving the licensed dose and were of too limited a duration (12 weeks) to provide useful data on

Table 5 - Baseline patient characteristics, costs, health state utilities and discount rate parameters used in the model.

Parameter	**Value**	**Probabilistic sensitivity analysis distribution**	**References**
Baseline characteristics			
Hypertension*	14283/18113	Beta (14283,3830)	[92,100]
Diabetes*	4221/18113	Beta (4221,13892)	[92,100]
Heart failure*	5793/18113	Beta (5793,12320)	[92,100]
Prior stroke*	2273/18113	Beta (2273,15840)	[92,100]
Prior transient ischaemic attack*	1663/18113	Beta (1663,16450)	[92,100]
Prior myocardial infarction*	3005/18113	Beta (3005,15108)	[92,100]
Prior intracranial haemorrhage*	713/18113	Beta (713,17400)	[92,100]
Health state utilities			
Atrial fibrillation (age 67)	0.774	1-Gamma (43.06,0.0052)	[101]
Stroke (permanent disutility)†	0.233	Normal (0.233,0.0032)	[102]
Stroke (temporary disutility) †	0.1385	Normal (0.1385,0.01)	[101,103]
Stroke (temporary duration, years)†	1/12	Uniform (0,0.183)	[103]
Myocardial infarction (permanent disutility)	0.0409	Normal (0.0409,0.002)	[101]
Myocardial infarction (temporary disutility)	0.1247	Normal (0.1247,0.01)	[101,103]
Myocardial infarction (temporary duration, years)	1/12	Uniform (0,0.183)	[103]
Intracranial haemorrhage (permanent disutility)	0.0524	Normal (0.0524,0.001)	[101]
Pulmonary embolism (temporary disutility)	0.1385	Normal (0.1385,0.01)	[101,103]
Pulmonary embolism (temporary duration, years)	1/12	Uniform (0,0.183)	[103]
Transient ischemic attack (temporary disutility)	0.1032	Normal (0.1032,0.01)	[101,103]
Transient ischemic attack (temporary duration, years)	5/365	Uniform (0,0.027)	[101]
Major bleed (temporary disutility)	0.1385	Normal (0.1385,0.01)	[101,103]
Major bleed (temporary duration, years)	1/12	Uniform (0,0.183)	[103]
Minor bleed (temporary disutility)	0.06	Normal (0.06,0.01)	[103]
Minor bleed (temporary duration, years)	5/365	Uniform (0.0.027)	[103]
Warfarin disutility	0.013	Gamma (1.3,0.01)	[102]
Dabigatran disutility	0.002	Gamma (0.2,0.01)	Assumption
Aspirin disutility	0.002	Gamma (0.2,0.01)	[102]

Costs			
Stroke – year 1†	£10,543.36	Gamma (102.68,102.68)	[104]
Stroke – subsequent years†	£2,781.22	Gamma (52.74,52.74)	[104]
Myocardial infarction – year 1	£2,357.13	Gamma (58.26,40.46)	[104]
Myocardial infarction – subsequent years	£828.90	Gamma (34.55,23.99)	[104]
Pulmonary embolism	£1,543.27	N/A	[105]
Transient ischaemic attack	£839.62	N/A	[105]
Major bleed	£1,684.58	N/A	[106]
Minor bleed	£93.17	N/A	[106]
Proton pump inhibitors (1 year)	£185.20	N/A	[107]
Warfarin – drugs (1 year)	£41.23	Uniform (32.98,49.48)	[75,107]
Warfarin – monitoring (1 year)	£198.39	Gamma (202.59,0.979)	[108]
Dabigatran – both doses (1 year)	£919.80	N/A	[109]
Aspirin (1 year)	£7.39	Gamma (73.9,0.1)	[75,103]
Discount rate			
Utilities	3.5%	N/A	[25]
Costs	3.5%	N/A	[25]

*Proportion in initial population

†Includes both strokes and systemic emboli, excluding pulmonary emboli

reduction in stroke event rate. The 5-year open-label extension to PETRO did not include warfarin as a comparator. We therefore used the RE-LY study for annualised clinical event rates[91,100] (Table 6, P72), and the patients being modelled consequently represented those of RE-LY.

This reflected patients in the RE-LY study[91], who were: 63.6% male, 70% Caucasian, with a mean age of 71.5 years, weight of 82.6kg and no contraindication to anticoagulation. Mean baseline $CHADS_2$ score was 2.1 with 32.4% of patients having a score of 3 or more. 50.4% of patients were vitamin-K antagonist naïve.

Our analysis considered the probability of (and reasons for) treatment discontinuation to better reflect the real world use of oral anticoagulants. This occurred in 21% and 17% of patients randomised to dabigatran and warfarin at 2 years, respectively[91,100]. Patients who discontinued dabigatran because of a bleed, or who discontinued warfarin (for any reason) were assumed to be switched to aspirin. Patients who discontinued dabigatran for reasons other than bleeds were assumed to be switched to warfarin, but this was tested in a sensitivity analysis.

Age-specific mortality rates from non-vascular causes were taken from general population data[69], as were incidence rates for hypertension[110] and diabetes mellitus[111]; all with the assumption that these adequately reflect the RE-LY population (Table 6, P72). The relative risks of thromboembolic events and bleeds with aspirin (versus warfarin) were taken from a published meta-analysis of comparative trials[112].

Utility estimates

The permanent utility decrement associated with stroke was taken from the results of the European Stroke Prevention Study, using the proportions of disabling and non-disabling strokes from RE-LY (45% of non-fatal strokes are non-disabling). The baseline health state utility for a person with atrial fibrillation (adjusted by age), as well as the decrements associated with other cardiovascular sequelae and haemorrhagic adverse events, were taken from a report of EQ-5D utility scores elicited from several thousand respondents to the US Medical Expenditure Panel Survey[101,103]. Utility losses in patients receiving warfarin (e.g. as a consequence of regular monitoring) and aspirin (assumed to be the same for dabigatran e.g. because of gastrointestinal upset), were obtained from a study of 83 patients with atrial

Table 6 - Clinical parameters used in the model

Parameter‡	**Aspirin**	**Warfarin**	**Dabigatran 110mg**	**Dabigatran 150mg**	**References**
Clinical event rates					
Stroke ($CHADS_2$ score ≤1)*†	0.0177	0.0109	0.0112	0.0068	[92,100,112]
Stroke ($CHADS_2$ score 2)*†	0.0222	0.0138	0.0145	0.0084	[92,100,112]
Stroke ($CHADS_2$ score ≥3)*†	0.0441	0.0273	0.0212	0.0189	[92,100,112]
Pulmonary embolism*	0.0016	0.0010	0.0012	0.0015	[92,100,112]
Transient ischaemic attack*	0.0135	0.0084	0.0062	0.0072	[92,100,112]
Congestive heart failure*	0.0062	0.0062	0.0070	0.0048	[92,100]
Probability of death from stroke†	0.1887	0.1887	0.1887	0.1887	[92,100]
Probability of death from pulmonary embolism	0.1591	0.1591	0.1591	0.1591	[92,100]
Vascular death (excluding stroke, systemic and pulmonary embolism)*	0.0228	0.0228	0.0216	0.0208	[92,100]
Probability major bleed is an intracranial haemorrhage	0.2191	0.2191	0.0839	0.0960	[92,100]
Adverse events					
Major bleed ($CHADS_2$ score ≤1)*	0.0127	0.0290	0.0188	0.0220	[92,100,112]
Major bleed ($CHADS_2$ score 2)*	0.0145	0.0331	0.0298	0.0304	[92,100,112]
Major bleed ($CHADS_2$ score ≥3)*	0.0202	0.0461	0.0380	0.0486	[92,100,112]
Minor bleed*	0.0718	0.1637	0.1316	0.1485	[92,100,112]
Non-bleed adverse events	N/A	0.4600	0.4596	0.4725	[92,100]
Proportion of patients using proton pump inhibitor	0.2317	0.1840	0.2126	0.2164	[92,100,113]
Myocardial infarction*	0.0064	0.0064	0.0082	0.0081	[92,100,104]
Co-morbidities					
Diabetes*	0.0122	0.0122	0.0122	0.0122	[111]
Hypertension*	0.0271	0.0271	0.0271	0.0271	[110]
Discontinuations					
Probability major bleed leads to discontinuation	N/A	0.1425	0.1801	0.2133	[92,100]

Probability adverse event leads to discontinuation	N/A	0.0194	0.0298	0.0292	[92,100]
Probability discontinue year 1 (other reasons)	N/A	0.0832	0.1160	0.1226	[92,100]
Probability discontinue year 2 onwards (other reasons)	N/A	0.0459	0.0475	0.0432	[92,100]

*Figures presented as rates per 100 person-years

†Includes both strokes and systemic emboli but not pulmonary emboli

‡See table 8 for parameters specifying the distributions for the probabilistic sensitivity analysis

Fibrillation[102]. All utility values are presented in Table 5 (P69), with multiple utility decrements for an individual patient assumed to be additive.

Resource use and cost estimates

All costs (besides those of dabigatran) are reported in 2009 GBP(£). Costs incurred during the first and subsequent years following stroke or myocardial infarction were inflated from 2006/7 prices[104]. Costs included in this figure were ward costs (staffing, equipment, consumables and overheads) and procedure costs (which also included the cost of hospital medicines), inpatient and outpatient costs, GP and district nurse visits and the costs of other medicines[104]. The costs of pulmonary emboli and transient ischaemic attacks were taken from NHS reference costs[105], as were those for managing major and minor bleeds, following the methodology and definitions of a NICE report on the costing of atrial fibrillation[106].

Incidences of other adverse events did not differ significantly between treatment groups, so it was not deemed necessary to attach a cost to such events. The exception to this is the higher incidence of dyspepsia in the dabigatran groups, 11.8% and 11.3% (110mg and 150mg), compared with 5.8% for warfarin, which was accounted for by including the cost of proton pump inhibitors. The proportion of patients taking proton pump inhibitors came from RE-LY, and the number of capsules per patient from a published cost-effectiveness analysis[107]. The relative proportion of patients using proton pump inhibitors in conjunction with aspirin was taken from a randomised, controlled trial of antithrombotic therapies[113].

The costs of warfarin and associated monitoring were based on a micro-costing analysis of 165 patients with atrial fibrillation included in a 6-month prospective cohort study, with the cost of warfarin initiation excluded from the long-term maintenance cost[108]. The average use of aspirin in practice was obtained from a published costing study[114].

Drug acquisition costs were taken from the British National Formulary[75] and the NICE appraisal consultation document for dabigatran[109]. All costs are presented in Table 5 (P69).

Discounting

We applied an annual discount rate of 3.5%[25] to costs, life-years and QALYs to reflect time preference but not to discreet clinical events[115].

Assumptions in model

- The risk of future cardiovascular events for each simulated patient, at any given time, is determined by their age, current treatment and $CHADS_2$ score (a stroke risk index based on age, diagnosis of hypertension, diabetes or congestive heart failure, and prior stroke or transient ischaemic attack[116]), according to probabilities determined from RE-LY (Table 6, P72)
- Clinical event rates (including myocardial infarction) would remain constant over time, unless a change occurred in one or more of the risk factors
- The rates of treatment discontinuation in the second year of the RE-LY study persisted for the lifetime of treatment
- The incidence (though not prevalence) of hypertension and diabetes was the same in patients with atrial fibrillation as the general population; as were deaths from non-vascular causes

Age-adjusted dosing

Patients initially below the age of 80 years start on the 150mg dose, and those above on the 110mg dose. If a person reaches 80 and is still continuing with the 150mg dose, they are then switched to the 110mg dose. We modelled this regimen in two different ways. Our primary method used the results of a post-hoc subgroup analysis[100], which subdivided people by age. The secondary method used the event rates from the full trial for patients taking either dose.

Sensitivity and scenario analyses

Univariate sensitivity analyses of each model parameter were performed to assess the stability of the results when key assumptions are tested. Parameter ranges were based on 95% confidence intervals where available, or alternatively, plausible percentage ranges (Table 7, P76). The possibility that the cost of managing intracranial haemorrhage and gastrointestinal bleeding may be higher with dabigatran than warfarin, because of the lack of an appropriate reversal agent, was tested by increasing the costs to consider the potential use of prothrombin complex concentrates (non-activated or activated)[117,118].

Our base-case assumes treatment benefit to persist for the lifetime of patients, but we tested two further scenarios, one where the benefit persisted for two years, the second where the benefit decreased linearly to zero over the ten years following the trial.

Table 7 - Combination of parameter estimates used to specify high and low ranges for the univariate sensitivity analyses

Analysis	Parameter	Low Value	High Value
Discount rates	Cost discount rate	0%	6%
	Utility discount rate	0%	6%
Discontinuation rates	Probability major bleed leads to discontinuation (warfarin)	0.1092	0.1794
	Probability major bleed leads to discontinuation (dabigatran)	0.1734	0.2561
	Probability adverse event leads to discontinuation (warfarin)	0.0159	0.0232
	Probability adverse event leads to discontinuation (dabigatran)	0.0250	0.0337
	Probability discontinue year 1 other reasons (warfarin)	0.0764	0.0903
	Probability discontinue year 1 other reasons (dabigatran)	0.1145	0.1310
	Probability discontinue year 2 other reasons (warfarin)	0.0404	0.0517
	Probability discontinue year 2 other reasons (dabigatran)	0.0378	0.0490
Bleed rates - dabigatran	Major bleed ($CHADS_2$ score ≤1)	0.0175	0.0270
	Major bleed ($CHADS_2$ score 2)	0.0254	0.0358
	Major bleed ($CHADS_2$ score ≥3)	0.0421	0.0556
	Probability major bleed is an intracranial haemorrhage	0.0683	0.1277
	Minor bleed	0.1422	0.1549
Medication cost - warfarin	Cost warfarin – drugs (1 year)	£33.39	£49.07
	Cost warfarin – monitoring (1 year)	£171.96	£226.56
Bleed rates - warfarin	Major bleed ($CHADS_2$ score ≤1)	0.0236	0.0349
	Major bleed ($CHADS_2$ score 2)	0.0280	0.0386
	Major bleed ($CHADS_2$ score ≥3)	0.0396	0.0531
	Probability major bleed is an intracranial haemorrhage	0.1799	0.2611
	Minor bleed	0.1571	0.1705
Event utility losses	Stroke (permanent disutility)	0.1703	0.2957
	Stroke (temporary disutility)	0.1189	0.1581
	Myocardial infarction (permanent disutility)	0.0370	0.0448

	Myocardial infarction (temporary disutility)	0.1051	0.1443
	Intracranial haemorrhage (permanent disutility)	0.0504	0.0544
	Pulmonary embolism (temporary disutility)	0.1189	0.1581
	Transient ischemic attack (temporary disutility)	0.0836	0.1228
	Major bleed (temporary disutility)	0.1189	0.1581
	Minor bleed (temporary disutility)	0.0502	0.0698
Medication utility losses	Warfarin disutility	0	N/A
	Dabigatran disutility	0	N/A
	Aspirin disutility	0	N/A
Vascular death rates - dabigatran	Probability of death from stroke	0.1561	0.2236
	Probability of death from pulmonary embolism	0.0681	0.2793
	Vascular death (excluding stroke, systemic and pulmonary embolism)	0.0183	0.0234
Vascular death rates - warfarin	Probability of death from stroke	0.1561	0.2236
	Probability of death from pulmonary embolism	0.0681	0.2793
	Vascular death (excluding stroke, systemic and pulmonary embolism)	0.0202	0.0256
Stroke rates - dabigatran	Stroke ($CHADS_2$ score ≤ 1)	0.0044	0.0097
	Stroke ($CHADS_2$ score 2)	0.0058	0.0114
	Stroke ($CHADS_2$ score ≥ 3)	0.0148	0.0234
Stroke rates - warfarin	Stroke ($CHADS_2$ score ≤ 1)	0.0077	0.0147
	Stroke ($CHADS_2$ score 2)	0.0105	0.0175
	Stroke ($CHADS_2$ score ≥ 3)	0.0224	0.0328

We conducted a probabilistic sensitivity analysis, implementing a Monte Carlo simulation of 2000 sets of simulated parameters (Tables 5 & 8, P69 & P79); to estimate the 95% central ranges of clinical event rates and net health benefits. In the economic analysis, the probabilistic sensitivity analysis was used to consider the joint uncertainty in costs and QALYs to estimate the probabilities of dabigatran being cost-effective at different thresholds, presented as a cost-effectiveness acceptability curve, and in different clinical scenarios.

We performed sub-group analyses to calculate the net health benefits (and associated 95% central ranges), the incremental cost-effectiveness ratios, and the probability of cost-effectiveness, in the following, pre-specified populations[100,119-121]:

1. Patients aged 75 or older.
2. Patients with a $CHADS_2$ score of 2, or a $CHADS_2$ score ≥ 3.
3. Patients who have previously suffered a stroke or transient ischaemic attack.
4. Patients attending trial centres (clinics) reporting mean INR time within therapeutic range greater (or lower) than 65.5%.
5. Patients on warfarin whose time within therapeutic range was greater (or lower) than 66.8% were compared with the full dabigatran populations. Only summary information was available for this calculation.
6. Patients with poor renal function as indicated by a low (30-50mL/min) creatinine clearance.
7. Patients who were vitamin-K antagonist naïve.

Table 8 - Parameters used for specifying the distributions for the probabilistic sensitivity analysis

Parameter	Aspirin	Warfarin	Dabigatran 110mg	Dabigatran 150mg
Clinical event rates				
Stroke ($CHADS_2$ score ≤1)* †	Lognormal(0.478,0.137)*Beta(37,3342)	Beta(37,3342)	Beta(41,3629)	Beta(25,3667)
Stroke ($CHADS_2$ score 2)* †	Lognormal(0.478,0.137)*Beta(60,4290)	Beta(60,4290)	Beta(59,4001)	Beta(35,4143)
Stroke ($CHADS_2$ score ≥3)* †	Lognormal(0.478,0.137)*Beta(102,3629)	Beta(102,3629)	Beta(82,3786)	Beta(73,3795)
Pulmonary embolism*	Lognormal(0.478,0.137)*Beta(12,11782)	Beta(12,11782)	Beta(14,11885)	Beta(18,12015)
Transient ischaemic attack*	Lognormal(0.478,0.137)*Beta(99,11695)	Beta(99,11695)	Beta(74,11825)	Beta(87,11946)
Congestive heart failure*	Lognormal(0.478,0.137)*Beta(73,11721)	Beta(73,11721)	Beta(83,11816)	Beta(58,11975)
Probability of death from stroke†	Beta(97,417)	Beta(97,417)	Beta(97,417)	Beta(97,417)
Probability of death from pulmonary embolism	Beta(7,37)	Beta(7,37)	Beta(7,37)	Beta(7,37)
Vascular death (excluding stroke, systemic and pulmonary embolism)*	Beta(269,11525)	Beta(269,11525)	Beta(257,11642)	Beta(250,11783)
Probability major bleed is an intracranial haemorrhage	Beta(87,310)	Beta(87,310)	Beta(27,295)	Beta(36,339)
Adverse events				
Major bleed ($CHADS_2$ score ≤1)*	Lognormal(-0.824,0.400)*Beta(98,3281)	Beta(98,3281)	Beta(69,3601)	Beta(81,3601)
Major bleed ($CHADS_2$ score 2)*	Lognormal(-0.824,0.400)*Beta(144,4206)	Beta(144,4206)	Beta(121,3939)	Beta(127,4051)

Major bleed ($CHADS_2$ score ≥3)*	Lognormal(-0.824,0.400)*Beta(172,3559)	Beta(172,3559)	Beta(147,3721)	Beta(188,3680)
Minor bleed*	Lognormal(-0.824,0.400)*Beta(1931,9683)	Beta(1931,9863)	Beta(1566,10333)	Beta(1787,10246)
Non-bleed adverse events	N/A	Beta(5425,6369)	Beta(5469,6430)	Beta(5685,6348)
Proportion of patients using proton pump inhibitor	1.26* Beta(1108,4914)	Beta(1108,4914)	Beta(1279,4736)	Beta(1315,4761)
Myocardial infarction*	Beta(66,11728)	Beta(66,11728)	Beta(87,11812)	Beta(89,11944)
Co-morbidities				
Diabetes*	Normal(0.0122,0.001)	Normal(0.0122,0.001)	Normal(0.0122,0.001)	Normal(0.0122,0.001)
Hypertension*	Normal(0.0271,0.002)	Normal(0.0271,0.002)	Normal(0.0271,0.002)	Normal(0.0271,0.002)
Discontinuations				
Probability major bleed leads to discontinuation	N/A	Beta(54,325)	Beta(58,264)	Beta(80,295)
Probability adverse event leads to discontinuation	N/A	Beta(105,5320)	Beta(163,5306)	Beta(166,5519)
Probability discontinue year 1 (other reasons)	N/A	Beta(501,5521)	Beta(698,5317)	Beta(745,5331)
Probability discontinue year 2 onwards (other reasons)	N/A	Beta(242,5036)	Beta(242,4852)	Beta(220,4872)

*Figures presented as rates per 100 person-years

†Includes both strokes and systemic emboli but not pulmonary emboli

Results

The results of our simulation at 2-years matched the results of the trial. No value deviated by more than 2.1%, a level of variability that would be expected given the stochastic nature of the simulation.

Clinical outcomes and net health benefit

In the base-case analysis, dabigatran 110mg bid and 150mg bid, extended life by 1.1 and 2.4 months, respectively, when compared with warfarin (Table 9, P82). The corresponding incremental net benefits were 0.094 (95% central range, -0.083 to 0.267) and 0.146 (95% central range, -0.029 to 0.322) QALYs. Compared with the low-dose dabigatran, the higher dose was associated with a positive incremental net benefit in 76% of simulations and with a mean value of 0.052 QALYs (95% central range, -0.122 to 0.228). Compared with warfarin, dabigatran 110mg bid and 150mg bid were associated with positive incremental net benefits in 86% and 94% of simulations, respectively.

Lifetime incidences of stroke or systemic embolism were 12.5% lower with dabigatran 110mg bid than warfarin; and 27.4% lower with dabigatran 150mg bid. Incidences of major haemorrhagic events were lower for low-dose dabigatran (by 4.0%), but higher for high-dose dabigatran (by 8.8%). There were no discernible differences in lifetime incidences of myocardial infarction between either doses of dabigatran, but these were about 19% higher than warfarin.

While being associated with lower bleeding rates, the higher rates of thrombotic events resulted in age-adjusted dabigatran dosing being inferior to the 150mg dose with respect to QALYs and life-years gained.

Costs and cost-effectiveness

Total, discounted, lifetime costs for dabigatran 110mg bid, 150mg bid and warfarin were, respectively, £10,529, £9,850 and £6,480. These were comprised mainly of drug and monitoring costs, which accounted for 47.3% and 44.2% of the overall costs of both doses of dabigatran compared with 22.4% for warfarin. The costs of managing strokes or systemic emboli accounted for 39.1%, 40.2% and 57.6% of total costs, respectively, with the remainder comprised of the costs of managing other events.

Table 9 - Lifetime estimates of event rates, net benefits and incremental differences versus comparator, derived from probabilistic sensitivity analysis

Outcome Referent	**Mean estimate**			**Mean difference**			**Comparator**
	Mean	2.5%	97.5%	Mean	2.5%	97.5%	
Quality-adjusted life-years (QALYs)†							
Warfarin	6.390	6.265	6.517	-0.094	0.083	-0.267	Dabigatran 110mg bid
Dabigatran 110mg bid	6.484	6.360	6.634	-0.049	0.126	-0.221	Dabigatran age-adj.‡
Dabigatran age-adj.‡	6.531	6.401	6.664	-0.005	0.171	-0.180	Dabigatran 150mg bid
Dabigatran 150mg bid	6.536	6.413	6.662	0.146	-0.029	0.322	Warfarin
Life-years†							
Warfarin	10.851	10.687	11.018	-0.089	0.142	-0.323	Dabigatran 110mg bid
Dabigatran 110mg bid	10.940	10.776	11.111	-0.102	0.129	-0.338	Dabigatran age-adj.‡
Dabigatran age-adj.‡	11.042	10.873	11.221	-0.009	0.243	-0.232	Dabigatran 150mg bid
Dabigatran 150mg bid	11.051	10.885	11.220	0.200	-0.035	0.429	Warfarin
Stroke or systemic embolism (excluding pulmonary emboli)							
Warfarin	0.2408	0.2010	0.2841	0.0302	-0.0260	0.0875	Dabigatran 110mg bid
Dabigatran 110mg bid	0.2107	0.1698	0.2538	0.0308	-0.0268	0.0893	Dabigatran age-adj.‡
Dabigatran age-adj.‡	0.1799	0.1401	0.2245	0.0044	-0.0476	0.0511	Dabigatran 150mg bid
Dabigatran 150mg bid	0.1755	0.1354	0.2196	-0.0654	-0.0092	-0.1226*	Warfarin
Ischaemic stroke							
Warfarin	0.1718	0.1484	0.1982	-0.0045	-0.0565	0.0493	Dabigatran 110mg bid
Dabigatran 110mg bid	0.1763	0.1507	0.2067	0.0331	-0.0189	0.0822	Dabigatran age-adj.‡
Dabigatran age-adj.‡	0.1432	0.1167	0.1708	0.0044	-0.0502	0.0570	Dabigatran 150mg bid
Dabigatran 150mg bid	0.1388	0.1121	0.1662	-0.0330	0.0261	-0.0803	Warfarin
Transient ischaemic attack							
Warfarin	0.1643	0.1281	0.2074	0.0218	-0.0280	0.0712	Dabigatran 110mg bid
Dabigatran 110mg bid	0.1425	0.1057	0.1791	0.0273	-0.0237	0.0762	Dabigatran age-adj.‡
Dabigatran age-adj.‡	0.1152	0.0791	0.1509	0.0042	-0.0449	0.0580	Dabigatran 150mg bid
Dabigatran 150mg bid	0.1110	0.0744	0.1476	-0.0533	-0.0035	-0.1027*	Warfarin

Intracranial haemorrhage							
Warfarin	0.0756	0.0655	0.0835	0.0479	0.0347	0.0614*	Dabigatran 110mg bid
Dabigatran 110mg bid	0.0277	0.0240	0.0308	-0.0062	-0.0191	0.0077	Dabigatran age-adj.‡
Dabigatran age-adj.‡	0.0339	0.0298	0.0372	-0.0017	-0.0133	0.0116	Dabigatran 150mg bid
Dabigatran 150mg bid	0.0356	0.0322	0.0391	-0.0400	-0.0271	-0.0578*	Warfarin
Major bleed (including intracranial haemorrhage)							
Warfarin	0.3313	0.2942	0.3766	0.0133	-0.0409	0.0673	Dabigatran 110mg bid
Dabigatran 110mg bid	0.3180	0.2811	0.3623	-0.0379	-0.0902	0.0257	Dabigatran age-adj.‡
Dabigatran age-adj.‡	0.3559	0.3180	0.3985	-0.0048	-0.0561	0.0512	Dabigatran 150mg bid
Dabigatran 150mg bid	0.3607	0.3233	0.4017	0.0294	0.0835	-0.0247	Warfarin
Non-fatal myocardial infarction							
Warfarin	0.0612	0.0434	0.0813	-0.0109	-0.0346	0.0126	Dabigatran 110mg bid
Dabigatran 110mg bid	0.0721	0.0560	0.0895	-0.0006	-0.0251	0.0256	Dabigatran age-adj.‡
Dabigatran age-adj.‡	0.0727	0.0560	0.0914	-0.0003	-0.0250	0.0255	Dabigatran 150mg bid
Dabigatran 150mg bid	0.0730	0.0561	0.0934	0.0119	0.0356	-0.0116	Warfarin

Columns 2-4 are means and central ranges for the given referent, columns 5-7 are means and central ranges for the difference from the comparator group.

*Incremental difference 95% central range not crossing zero

†Discounted at 3.5% per annum

‡Age-adjusted dabigatran dosing regimen (110mg bid for patients aged ≥80 years) based on a post-hoc subgroup analysis.

Table 10 - Cost-effectiveness results for sub-groups, based on the probabilistic sensitivity analysis.

Sub-group	**Warfarin cost**	**Warfarin QALY**	**Dabigatran 150mg bid Cost**	**Dabigatran 150mg bid QALY**	**ICER (£/QALY)**	**Probability cost-effective at:**	
						£20,000 per QALY	£30,000 per QALY
RE-LY population	£6,480	6.390	£9,850	6.536	£23,082	0.449	0.596
$CHADS_2$ score 2	£7,412	6.283	£10,443	6.433	£20,207	0.475	0.615
$CHADS_2$ score ≥3	£9,912	6.224	£12,646	6.396	£15,895	0.565	0.683
Centre time within therapeutic range ≥65.5%	£6,247	6.517	£9,977	6.605	£42,386	0.137	0.309
Centre time within therapeutic range <65.5%	£6,617	6.261	£9,656	6.410	£20,396	0.469	0.636
Patients' time within therapeutic range ≥66.8%	£6,302	6.401	£9,850	6.536	£26,281	0.393	0.511
Patients' time within therapeutic range <66.8%	£6,694	6.360	£9,850	6.536	£17,932	0.519	0.643
Creatinine clearance <30-50mL/min	£7,991	6.310	£10,788	6.460	£18,647	0.501	0.631
Prior stroke or transient ischaemic attack	£10,004	6.217	£12,787	6.378	£17,286	0.525	0.649
Vitamin-K antagonist naive	£6,437	6.396	£9,792	6.545	£22,517	0.446	0.587
Age ≥75 years	£4,612	4.275	£7,362	4.429	£17,857	0.498	0.635

Final two columns show proportion of simulations in which dabigatran 150mg bid is cost-effective versus warfarin

The incremental cost-effectiveness ratio (ICER) for low-dose dabigatran versus warfarin is £43,074 per QALY gained; while that of high-dose dabigatran is £23,082 per QALY gained (Table 10, P84). Dabigatran 110mg bid is dominated as a strategy by dabigatran 150mg bid, as it is associated with a worse health outcome (-0.052 QALYs) and higher cost (+£679).

Age-adjusted dosing

The use of dabigatran 110mg bid from the age of 80 years is dominated by the 150mg bid dose under both possible modelling methodologies. In the models based on the post-hoc subgroup analysis, and using full RE-LY data, respectively, the use of the lower dose accrues 0.005 and 0.017 less QALYs and costs £62 and £234 more over a lifetime. Compared to warfarin, the ICERs when low dose dabigatran is used in the over 80s are £24,340 and £27,940 per QALY gained, for the two methodologies, respectively.

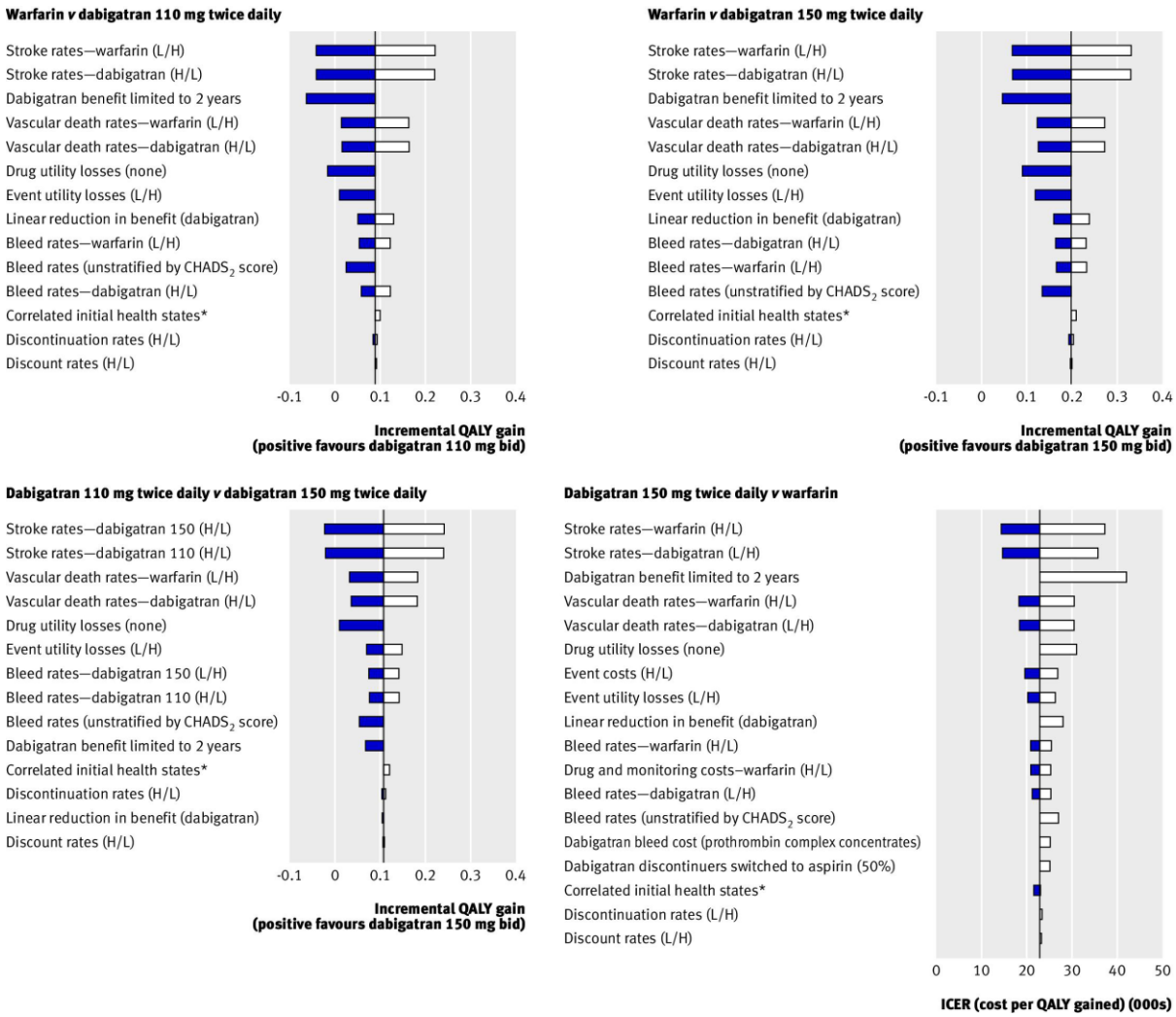


Figure 6

Sensitivity analysis

The tornado plot (Figure 6, P85) indicates the sensitivity of incremental net benefits to stroke rates and the duration of effect of dabigatran. Dabigatran 150mg bid was cost-effective at the lower threshold of £20,000 per QALY when we assumed decreases (or increases) in the rates of stroke or vascular death in patients receiving dabigatran (or warfarin); or increases in either clinical event costs or utility losses.

Compared to warfarin, the ICER for dabigatran 110mg bid exceeded £32,415 per QALY in all sensitivity analyses.

The probabilistic sensitivity analysis (Figure 7) indicates that warfarin has the highest probability of being cost-effective at thresholds of £24,400 or lower. Dabigatran 150mg bid is the most probable cost-effective option at thresholds above that value. Considering a pair-wise comparison between warfarin and dabigatran 150mg bid, warfarin is the most cost-effective treatment at thresholds of £22,800 and below.

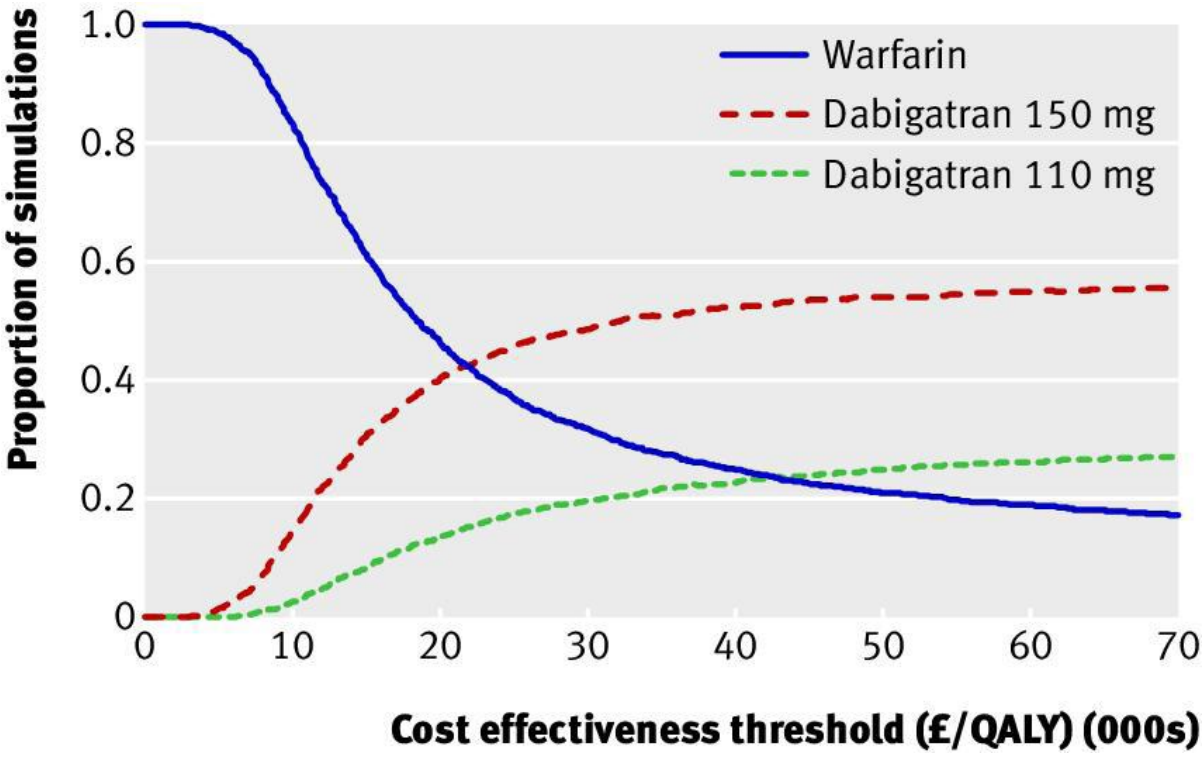


Figure 7

Sub-group analyses

Among the sub-groups analysed, the mean incremental net health benefit consistently favoured both doses of dabigatran over warfarin; and dabigatran 150mg bid over 110mg bid (Figure 8).

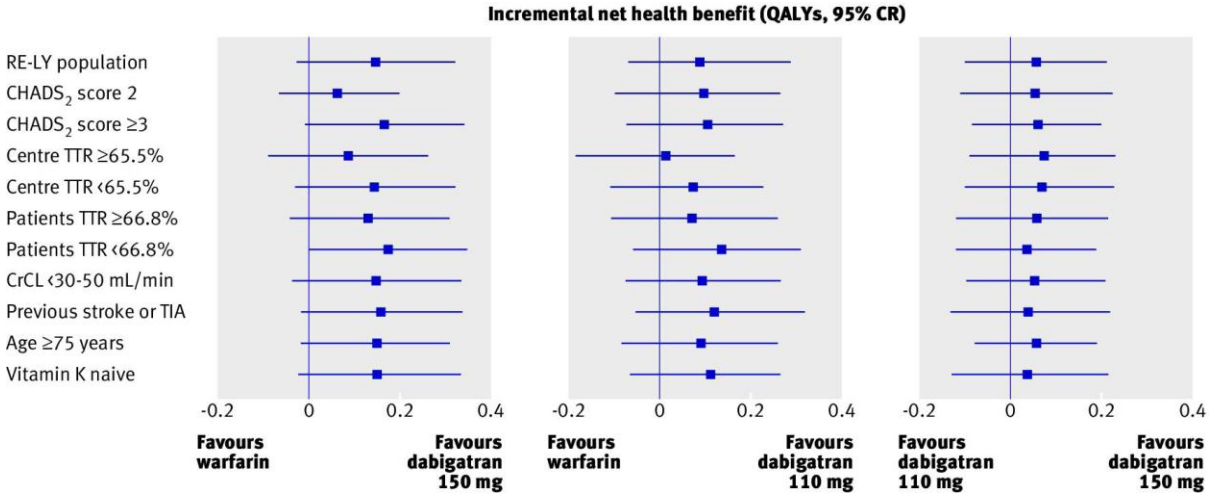


Figure 8

Dabigatran 150mg bid was within the £30,000 per QALY cost-effectiveness threshold for all patient sub-groups other than in centres with mean INR time within therapeutic range ≥65.5% (Table 10, P84). Dabigatran 150mg bid was most cost-effective in patients at high risk of stroke ($CHADS_2$ score ≥3), but even here the probability of being cost-effective is only 68%.

Dabigatran 110mg bid, when used for all ages or restricted to patients ≥80 years, was dominated by the higher dose in all sub-groups.

Discussion

Our quantitative benefit-harm analysis suggests that dabigatran is associated with positive net health benefits when compared with warfarin. High-dose dabigatran was the most clinically effective option. Greatest benefits were evident when compared with patients in whom INR control is poorest (patient time within therapeutic range <66.8%); and least benefits in centres that achieve good INR control (centre time within therapeutic range ≥65.5%). We were unable to identify a patient sub-group in which the lower dose of dabigatran – when used for all ages, or restricted to patients ≥80 years - was superior to the higher dose. The benefits of reduced bleeding rates with the lower dose are offset by reduced efficacy in stroke prevention. These finding are in accordance with the results of the RE-LY study[92], and related sub-group analyses[100,119-121], and lends support to the FDA's rationale for not licensing the 110mg dose[94].

The economic analysis indicated that for the overall RE-LY study population, dabigatran 150mg bid is potentially a cost-effective alternative to warfarin, at £23,082 per QALY gained. However, its probability of being cost-effective at a threshold of £20,000 per QALY is only 45%. This uncertainty is driven largely by stroke rates and, to a lesser extent, vascular death rates and costs of managing strokes. NICE's criteria for decision-making state that "above a most plausible ICER of £20,000 per QALY gained, judgements about the acceptability of the technology as an effective use of NHS resources will specifically take account the degree of certainty around the ICER. NICE will be more cautious about recommending a technology when it is less certain about the ICERs presented"[25]. Dabigatran 110mg bid is not a cost-effective option, and the age-adjusted dosing regimen is dominated in all scenarios by the 150mg dose.

High-dose dabigatran is more cost-effective in patients at a greater risk of stroke (baseline $CHADS_2$ score ≥3). However, at centres which achieve good INR control (centre time within therapeutic range ≥65.6%), dabigatran 150mg bid is no longer cost-effective, at £42,386 per QALY gained. While the mean time within INR range in the UK of 72% in the RE-LY study[119] may be higher than routine practice[108], so too might adherence to dabigatran, which requires twice daily dosing compared with warfarin's once.

Comparison with other studies

We are unaware of any quantitative benefit-harm analyses of dabigatran in atrial fibrillation. However, two economic evaluations of dabigatran in non-valvular atrial fibrillation have

been published. Both used Markov models to estimate lifetime cost-effectiveness based on the RE-LY trial.

The US study, which adopted the costing perspective of a health insurer[122], yielded a quality-adjusted life expectancy of 10.28 with warfarin, 10.70 with low-dose dabigatran, and 10.84 with high-dose dabigatran. These are considerably higher than our estimates, primarily because of patients' starting age which, at 65 years, is 6.1 years younger than in our analysis based on the RE-LY population. Nevertheless, despite this differences, similar results were obtained with respect to dabigatran 150mg bid being associated with positive incremental net health benefits across a range of risks for stroke and intracranial haemorrhage, compared with dabigatran 110mg bid and warfarin. A similar economic outcome also resulted, with the ICER falling just below the cost-effective threshold but with a high level of uncertainty, driven mostly by drug costs and stroke rates.

The Canadian study[123], sponsored by the manufacturer of dabigatran and based on RE-LY patient-level data (though not listed as a pre-specified analysis)[98], assessed its cost-effectiveness according to the same age-adjusted dosing schedule as approved in Europe. In contrast to the US and our study, however, dabigatran was deemed to be cost-effective compared with warfarin, at Can$10,440 per QALY gained. Differences relate largely to costs, which were proportionally much greater in their analysis for the management of events and long term care. Considering a patient, taking dabigatran, who has an acute stroke and five years of follow-up costs; in our analysis, the costs of stroke is about 5-times higher than the cost of medication whilst in the Canadian study they are more than 15-times higher.

Strengths and weaknesses

Our analysis benefited from applying a discrete event simulation methodology, which is the method of choice for conditions where there are no obvious discrete disease states into which patients can be classified, a necessary assumption for a Markov model[27]. It allows for a much larger number of potential health states to be modelled and removes the need to define the additional structural parameters necessary for a Markov model (e.g. cycle length). A discrete event simulation also operates in continuous rather than discrete time, thus more naturally approximating actual patient histories and allowing continuous parameters (e.g. age) to be more appropriately modelled.

Our analysis addresses the concerns raised by NICE in its appraisal of the manufacturer's submission[109]; namely, the inclusion of the age-adjusted dosing regimen, use of reliable estimates of INR monitoring cost, continuation of dyspepsia throughout duration of dabigatran treatment, and treatment-independence of the risks of disability and mortality after stroke. We had no access to data on the quality of life sub-study of RE-LY; and made no attempt at modelling a typical UK atrial fibrillation population who are typically older, with proportionately more females, and a different stroke risk profile than the RE-LY trial population[124]. Patients are also less likely to persist with anticoagulant therapy in routine practice than in a clinical trial setting[124], but we had no additional data for more elaborate modelling.

However, there are a number of caveats. First, the reliance on the RE-LY study as the sole source of clinical data is a potential cause for concern. Although RE-LY is one of the largest trials of atrial fibrillation, this makes it difficult to assess the impact of any possible weaknesses in the design of the RE-LY study (e.g. its open-label design, a significant proportion of patients taking aspirin concomitantly and only about a third of patients with a baseline $CHADS_2$ score ≥ 3). We were further limited by not having access to individual patient data. Our *a priori* decision to base our analysis on the entire RE-LY study population may limit the generalisability of the base-case estimates to a UK context. Sub-groups, defined by centres achieving better INR control and patients within the higher categories of stroke risk, may result in more relevant ICER estimates.

Second, the necessity of bringing together data from a wide variety of sources has the potential to introduce bias into the analysis. For example, relative event rates for aspirin therapy were derived from a separate study, which will have had different patient demographics and different warfarin dosing schedules to RE-LY. There are also issues regarding the extrapolation of a two year trial to a lifetime horizon, and the assumption that utility decrements for events derived from the general population are appropriate for patients with atrial fibrillation. However, approximations such as these are unavoidable in economic modelling.

Third, we did not include the possibility that widespread use of dabigatran might impact on the provision of anticoagulation clinic services, as we considered it unlikely that dabigatran would displace warfarin to such an extent.

Implications for practice and future research

There are advantages of dabigatran over warfarin, the most important being that monitoring will not be required, there is more predictable anticoagulation for a given dose and there are likely to be fewer drug-drug interactions. However, there are disadvantages as well[125]: (a) the lack of monitoring provides little ability to objectively monitor adherence which in the real-world setting is likely to be worse with dabigatran given the need for twice daily dosing and its associated higher incidence of dyspepsia; (b) if the patient has a serious bleed, there are no proven antidotes[126]; (c) there is some uncertainty of dosing in certain clinical settings such as renal failure, elderly and concomitant intake of amiodarone, which may lead to either under- or over-dosing given that there is no ability to monitor with a pharmacodynamic marker; (d) the safety and efficacy of thrombin inhibitors in the longer term (beyond 2 years) are uncertain, though the follow-up study of RE-LY patients should yield valuable information[127].

An important finding from the cost-effectiveness analysis is that dabigatran is not cost-effective when compared to patients whose INR is well controlled, or in centres which achieve good INR control, which includes the UK. Part of the reason why there is such variability in the time within therapeutic range with warfarin is the presence of genetic polymorphisms in the CYP2C9 and VKORC1 genes[108,128]. There are at least 4 randomised trials running globally where genotype-guided prescribing for warfarin, which is predicted to improve the time within therapeutic range, is being tested against current clinical care. Whether dabigatran would be cost effective against genotype-guided prescribing of warfarin is unclear, and needs further evaluation. Furthermore, there are other competitors to dabigatran due to be evaluated for licensing soon such as rivaroxaban and apixaban, which have shown similar clinical effectiveness to warfarin, but have not been tested against dabigatran[129]. Thus, while the arrival of new anticoagulants should be welcomed, their place in the prevention of strokes in patients with atrial fibrillation in comparison to warfarin (perhaps genotype-guided) needs further evaluation. In the end, a stratified approach may represent the best approach to maximise both the clinical and cost-effectiveness of anticoagulation in patients with atrial fibrillation.

Preface to Chapter 4

During the time taken for the construction of the economic model in the previous chapter, two other large trials of novel anticoagulants (the ROCKET-AF trial of rivaroxaban and the ARISTOTLE trial of apixaban) had reported. A decision was made that rather than proceeding directly to the planned population PKPD simulation, an indirect comparison of these trials would be conducted, to obtain estimates of relative treatment efficacy. These could then be fed into the discrete event simulation described in the previous chapter to obtain estimates of comparative effectiveness at a lifetime horizon.

Such an approach would enable us to more accurately reflect the process of drug development in the real world. Since both effectiveness and cost-effectiveness are ultimately relative rather than absolute, in decision making it is important to be able to compare any newly available treatment with all relevant comparators. If we had proceeded directly to our simulations and not included these other drugs, it would have led to an incomplete and potentially inaccurate assessment of their effectiveness and cost-effectiveness. Since not all of the necessary economic data were available to perform a cost-effectiveness analysis at this time (specifically, the costs of rivaroxaban and apixaban were not available) it was decided to perform a comparison based only on effectiveness and net clinical benefit. Any newly available cost data would then be included in the final cost-effectiveness model when all PKPD simulations had been conducted.

Chapter 4

Comparative effectiveness of dabigatran, rivaroxaban, apixaban and warfarin in the management of non-valvular atrial fibrillation

Summary

Introduction: Alternative anticoagulants to warfarin are available for the prevention of thromboembolic stroke in atrial fibrillation, but there is a lack of information on their comparative effectiveness. We evaluate the comparative effectiveness of dabigatran, rivaroxaban, apixaban and warfarin in patients with non-valvular atrial fibrillation.

Methods: A discrete event simulation with a lifetime horizon of analysis and based on an indirect comparison of the RE-LY, ROCKET-AF and ARISTOTLE trial results. Simulations were performed on cohorts of patients with characteristics matching those of the US atrial fibrillation population, with a mean age of 73 years, and mean $CHADS_2$ score of 1.92. Outcomes included the incremental net health benefits, defined in terms of quality-adjusted life-years (QALYs); probability of each treatment being most effective; and clinical event rates.

Results: Over a lifetime, apixaban, dabigatran and rivaroxaban accrued 0.1297 (95% central range [CR] -0.0296 to 0.2644), 0.1055 (95% CR -0.0481 to 0.2477) and 0.0948 (95% CR -0.0523 to 0.2422) more QALYs than warfarin, respectively. They were also associated with 33.6%, 17.0% and 6.7% lower lifetime incidences of stroke or systemic embolism. Apixaban was associated with a 21.3% lower lifetime incidence of major bleeding, with dabigatran and rivaroxaban having 1.3% and 7.0% higher incidences, respectively. This ordering was maintained across patient subgroups and modelling assumptions, with apixaban having a 55% probability of accruing the highest total lifetime QALYs.

Discussion: In the absence of a definitive trial, and acknowledging the limitations of an indirect comparison, the available evidence suggests apixaban to be the most effective anticoagulant.

Introduction

Atrial fibrillation (AF) is estimated to affect 2.5 million people in the United States, results in a fivefold increase in the risk of ischemic stroke[87,130]. Associated costs exceed $7 billion annually[131]. Antithrombotic agents have proven benefits in preventing stroke in patients with AF. Until recently, vitamin K antagonists, such as warfarin, were the only form of oral thromboprophylactic anticoagulation treatment[132]. Although clinically effective and inexpensive, warfarin increases the risk of haemorrhage, interacts with many drugs and, because of the considerable variability in patient response, requires careful monitoring[91,133].

Newer oral anticoagulants, which include the direct thrombin inhibitor dabigatran etexilate (hereafter referred to as dabigatran), and the direct factor Xa inhibitors rivaroxaban and apixaban, have recently been developed. The Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) study evaluated two doses (110mg and 150mg twice daily) of dabigatran as an alternative to warfarin in 18,113 patients with at least one risk factor for stroke[92]. The Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCKET-AF) compared rivaroxaban with warfarin in 14,264 patients at elevated risk of stroke[134]. The Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation (ARISTOTLE) trial compared apixaban with warfarin in 18,201 patients with at least one risk factor for stroke[135].

After median follow up periods of approximately two years, the primary outcome of stroke or systemic embolism showed a potential improvement in the intention-to-treat population with all three alternatives, demonstrating superiority of the licensed 150mg dose of dabigatran (1.11% v 1.71% per year; $p<0.001$) and apixaban (1.27% v 1.60% per year; $p=0.01$) and non-inferiority of rivaroxaban (2.1% v 2.4% per year; $p=0.12$), compared with warfarin. Rates of major bleeding were not significantly different between dabigatran 150mg and warfarin or between rivaroxaban and warfarin, but apixaban was associated with a lower risk of major bleeding (2.13% v 3.09% per year; $p<0.001$).

Both dabigatran and rivaroxaban have been approved by the US Food and Drug Administration and a decision on apixaban is pending. Whilst their efficacies in relation to warfarin have been demonstrated, their comparative effectiveness remains unknown. With no prospect of a head-to-head trial, we describe an adjusted, indirect comparison to help guide treatment selection. The analysis assesses the trade-offs in thrombotic and bleeding

risks for all four anticoagulants, and incorporates a preference-based, patient-centred outcome, the quality-adjusted life-year (QALY), to combine health-related quality of life with survival. Our analysis acknowledges differences in trial designs and populations, and their potential impacts on estimated comparative effectiveness, by adopting a probabilistic approach for a range of plausible scenario analyses.

Methods

Comparative effectiveness was assessed using an indirect analysis that extrapolated benefits and harms to a lifetime horizon, consistent with AF being a lifelong condition requiring indefinite treatment.

The analysis is based on a discrete event simulation model which we have described previously (chapter 3), and which allows for explicit incorporation of both structural and parameter uncertainty[27]. The model simulates the clinical events and outcomes experienced by individual patients. The risks of their occurrence are determined from patients' characteristics which are updated according to time and event history. Comparative effectiveness was determined from incremental net health benefits, measured as the differences between treatments in QALYs, and from modelled clinical event rates[96] (chapter 4).

Model population

In the base-case analysis, patients' baseline characteristics, which were assumed to be uncorrelated, were representative of the stroke risk profile of the US atrial fibrillation population[136]. Patients had a mean age of 73.0 years, with 38.8%, 36.8%, 18.0%, 6.4% having $CHADS_2$ (Congestive heart failure, Hypertension, Age ≥75, Diabetes mellitus, prior Stroke/transient ischemic attack) scores of 1, 2, 3 and ≥4 respectively[136].

For each treatment, identical cohorts of 100,000 patients were generated. Each patient was given a simulated set of characteristics consisting of the presence or absence (at the start of the simulation) of the following: hypertension, diabetes mellitus, congestive heart failure, prior stroke, prior transient ischemic attack, prior myocardial infarction and prior intracranial haemorrhage, drawn from binomial distributions based on the probability of having each condition at baseline (table 11, P98).

Interventions

The analysis considered a dose of 5mg twice daily of apixaban, and the licensed doses of dabigatran 150mg twice daily, rivaroxaban 20mg once daily, and dose-adjusted warfarin.

Table 11 - Patients' baseline characteristics

Baseline characteristics*	**RE-LY**	**ROCKET-AF**	**ARISTOTLE**
Number of patients	18,113	14,264	18,201
Hypertension[92,134,135]	78.9%	90.5%	87.4%
Diabetes[92,134,135]	23.3%	39.9%	25.0%
Heart failure[92,134,135]	32.0%	62.5%	35.4%
Prior stroke[92]	12.5%	34.4%†	11.9%†
Prior transient ischemic attack[92]	9.2%	25.3%†	8.7%†
Prior myocardial infarction[92,134,135]	16.6%	17.3%	14.2%
Prior intracranial haemorrhage[92]	3.9%	10.7%†	3.7%†

*Percentage in initial population.

†These values were imputed from the data available in the RE-LY study and the distribution of $CHADS_2$ scores at the start of the trial, which was known for all three studies, under the assumption that the ratio of strokes to transient ischemic attacks and intracranial haemorrhages would be consistent between trials. Probability of prior stroke or TIA in ROCKET-AF was 55%, and in ARISTOTLE was 19%.

Clinical parameters

Annualised clinical event rates were extracted from the RE-LY, ROCKET-AF and ARISTOTLE trials[92,134,135] identified from a systematic review of the literature[137]. Based on the method of Bucher et al[138], indirect comparisons were adjusted according to the results of their direct comparisons with warfarin. This adjustment accounts for differing baseline risks between trials by assuming a constant relative treatment effect e.g. for two trials comparing A and B, and B and C, with relative risks for a given event of RR_{AB} and RR_{BC} respectively, the indirect, relative effect of C versus A is estimated as:

$$\ln(RR_{AC}) = \ln(RR_{AB}) + \ln(RR_{BC})$$

Event rates for dabigatran, apixaban and rivaroxaban were calculated by multiplying relative treatment effects by warfarin event data, calculated from a meta-analysis of the warfarin arms of the three trials (table 12, P100). Diabetes and hypertension incidence rates were taken from US general population data[139,140], as were age-specific non-vascular mortality data[69], all with the assumption these accurately reflect the atrial fibrillation population.

Whenever the necessary data were not available (e.g. a particular parameter was not reported for a given trial), values were imputed, based on data from the other trials and US population data (tables 11 & 12, P98 & P100). All assumptions were assessed through sensitivity analysis (see structural sensitivity analysis, below).

In order to better reflect the use of oral anticoagulants in routine care, the analysis also includes the trial-derived probabilities of (and reasons for) discontinuation of treatment. Patients who discontinued dabigatran, rivaroxaban or apixaban because of a bleed or who discontinued warfarin for any reason were assumed to be switched to aspirin, whilst other discontinuing patients were switched to warfarin. This assumption was tested in a sensitivity analysis. Relative risks of events for aspirin came from a published meta-analysis of trials comparing warfarin and aspirin[112], and the AVERROES trial comparing apixaban with aspirin in patients deemed unsuitable for vitamin K antagonist therapy[141].

QALYs were discounted at 3% per annum to reflect time preference[142], but no discounting was applied to discrete clinical events.

Table 12 - Clinical event rates

Parameter	Warfarin	Dabigatran	Rivaroxaban	Apixaban	Aspirin
Stroke ($CHADS_2$ score ≤ 1)*	0.00921	0.00536	0.00750†	0.00678	0.01485
Stroke ($CHADS_2$ score 2)*	0.01405	0.00824	0.01255	0.01211	0.02265
Stroke ($CHADS_2$ score 3)*	0.01957	0.01164†	0.01335	0.01133†	0.03157
Stroke ($CHADS_2$ score 4)*	0.03119	0.02154†	0.02442	0.02097†	0.05030
Stroke ($CHADS_2$ score 5)*	0.02899	0.02398†	0.02785	0.02334†	0.04676
Stroke ($CHADS_2$ score 6)*	0.03639	0.03098†	0.03511	0.03015†	0.05869
Systemic embolism*	0.00135	0.00113	0.00031	0.00115	0.00217
Pulmonary embolism*	0.00078	0.00114	0.00091‡	0.00060‡	0.00126
Transient ischemic attack*	0.00839	0.00723	0.00662‡	0.00616‡	0.01354
Myocardial infarction*	0.00763	0.01008	0.00620	0.00666	0.00763
Congestive heart failure*	0.00619	0.00482	0.00488‡	0.00454‡	0.00619
Vascular death (excluding stroke and systemic and pulmonary embolism)*	0.02281	0.02078	0.02155	0.02118	0.02281
Probability of death from stroke or systemic embolism	0.25457	0.25457	0.25457	0.25457	0.25457
Probability of death from pulmonary embolism	0.15909	0.15909	0.15909	0.15909	0.15909
Major bleed ($CHADS_2$ score ≤ 1)*	0.02612	0.01981	0.02248†	0.01590	0.01146
Major bleed ($CHADS_2$ score 2)*	0.03175	0.02916	0.03379†	0.02434	0.01393
Major bleed ($CHADS_2$ score ≥ 3)*	0.04433	0.04674	0.04799†	0.03061	0.01944
Probability that major bleed is intracranial haemorrhage	0.23361	0.10234	0.14947	0.14068	0.23361
Minor bleed*	0.16560	0.15020	0.17140	0.11783	0.07263
Diabetes*	0.0141	0.0141	0.0141	0.0141	0.0141
Hypertension*	0.0323	0.0323	0.0323	0.0323	0.0323
Probability of discontinuation (year 1)*	0.14466	0.22048	0.14481	0.14232	N/A
Probability of discontinuation (year 2 onwards)*	0.06760	0.06695	0.07224	0.04975	N/A

*Presented as rates per 100 person years.

†Where stratified event rates were not available, unknown stratified risks were imputed based on the assumption that the relative risks of events for patients with different $CHADS_2$ scores would be independent of treatment.

‡Imputed, based on the relative risks of different events from the RE-LY study, on the assumption that the relative risks of different thromboembolic events would be independent of treatment.

Table 13 - Health state utilities assigned to treatments and clinical events, and the corresponding duration of acute events.

Parameter	**Value**	**Probabilistic sensitivity analysis distribution**
Atrial fibrillation (age 67)[101]	0.774	1-Gamma(43.06,0.0052)
Stroke/systemic embolism (permanent disutility)[102]	0.235	Normal(0.235,0.0032)
Stroke/systemic embolism (temporary disutility)[101,103]	0.1385	Normal(0.1385,0.01)
Stroke/systemic embolism (temporary disutility, years)[103]	1/12	Uniform(0,0.183)
Myocardial infarction (permanent disutility)[101]	0.0409	Normal(0.0409,0.002)
Myocardial infarction (temporary disutility)[101,103]	0.1247	Normal(0.1247,0.01)
Myocardial infarction (temporary duration, years)[103]	1/12	Uniform(0,0.183)
Intracranial haemorrhage (permanent disutility)[101]	0.0524	Normal(0.0524,0.001)
Pulmonary embolism (temporary disutility)[101,103]	0.1385	Normal(0.1385,0.01)
Pulmonary embolism (temporary duration, years)[103]	1/12	Uniform(0,0.183)
Transient ischemic attack (temporary disutility)[101,103]	0.1032	Normal(0.1032,0.01)
Transient ischemic attack (temporary duration, years)[103]	5/365	Uniform(0,0.027)
Major bleed (temporary disutility)[101,103]	0.1385	Normal(0.1385,0.01)
Major bleed (temporary duration, years)[103]	1/12	Uniform(0,0.183)
Minor bleed (temporary disutility)[103]	0.06	Normal(0.06,0.01)
Minor bleed (temporary duration, years)[103]	5/365	Uniform(0.0.027)
Warfarin disutility[102]	0.013	Gamma(1.3,0.01)
Dabigatran/rivaroxaban/apixaban disutility	0.002	Gamma(0.2,0.01)
Aspirin disutility[102]	0.002	Gamma(0.2,0.01)

Utility estimates

The age-adjusted baseline health utility for a person with atrial fibrillation, together with the utility decrements associated with cardiovascular sequelae (excluding stroke) and haemorrhagic events, were taken from the EQ-5D scores in a US Medical Expenditure Panel Survey of several thousand patients[101,103]. The permanent utility decrement associated with stroke was derived from the European Stroke Prevention Study, based on the proportions of disabling to non-disabling strokes (43% of non-fatal strokes were non-disabling across RE-LY, ROCKET-AF and ARISTOTLE). The analysis incorporated utility losses inherent to warfarin (e.g. as a result of monitoring), and aspirin (assumed to be the same for dabigatran, rivaroxaban and apixaban)[102]. Multiple utility decrements were assumed to be additive and utility values are given in table 13 (P102).

In order to assess the sensitivity of the model to the choice of utility parameters, a sensitivity analysis was conducted, replacing base-case utility values with those from an alternative cost-effectiveness study in atrial fibrillation[122].

Parameter sensitivity analysis

Univariate sensitivity analyses of each parameter in the model were conducted to assess the effect of varying assumptions on the stability of the results. 95% confidence intervals were used as the upper and lower limits for parameters or, where these were not available, plausible percentage ranges.

A probabilistic parameter sensitivity analysis was also conducted as a Monte Carlo simulation of 2,000 sets of parameters sampled from appropriate distributions. This provided estimates of the 95% central ranges (2.5th to 97.5th percentile) for clinical event rates and net health benefits, and the probability of each treatment option resulting in the highest net health benefit.

Structural sensitivity analysis

The model necessitated a large number of assumptions, either for simplification purposes or because the desired data were not available in the necessary format. The robustness of the results in relation to different assumptions was assessed quantitatively.

Univariate analyses considered the different options presented in table 14 (P105). A probabilistic analysis was performed by sampling 10,000 times, at random, from the subspace

of possible structural assumptions, with each assumption being equally likely to be selected. As previously, the outputs are presented as the probabilities with which each treatment option results in the highest net health benefit.

In order to assess whether the choice of a discrete event simulation framework had a significant impact on the results, a secondary analysis was performed, replacing our simulation model with a published Markov model[143].

Scenario analyses

Subgroup analyses were performed to calculate the net health benefits (and associated 95% central ranges) in the following pre-specified populations. Analyses for patients aged 75 or older; patients with a $CHADS_2$ score of 3 or more; patients with the baseline characteristics of those in each of the three studies; and patients who have previously had a stroke or transient ischemic attack were performed by altering the baseline patient characteristics in the model. Separate, indirect comparisons were made for patients with impaired renal function (30-50mL/min creatinine clearance); and patients who were naïve to vitamin K antagonist treatment.

Table 14 - Structural sensitivity analysis

Parameter	Assumption alternatives
Population demographics	• US atrial fibrillation population[136]* • RE-LY population[92] • ROCKET-AF population[134] • ARISTOTLE population[135]
Imputing missing demographic data	• Imputed based on RE-LY data* • Imputed based only on warfarin arm of RE-LY • Imputed based only on dabigatran arm of RE-LY • Extrapolated based on US event rates
Baseline demographics	• Uncorrelated* • Correlated
Warfarin event rates	• Meta-analysis of all three trials* • Taken from RE-LY only • Taken from ROCKET-AF only • Taken from ARISTOTLE only
Stroke stratification	• Full stratification ($CHADS_2$ 1,2,3,4,5,6)* • Partial stratification ($CHADS_2$ ≤2,3,4,5,6) • Partial stratification ($CHADS_2$ 1,2,≥3) • Unstratified
Stroke stratification imputation (rivaroxaban with $CHADS_2$ score of 1)	• Imputed from RE-LY and ARISTOTLE* • Imputed from RE-LY • Imputed from ARISTOTLE • Imputed from US population data
Stroke stratification imputation (dabigatran and apixaban with $CHADS_2$ score ≥3)	• Imputed from ROCKET-AF* • Imputed from US population data
Systemic embolism	• Disaggregated from strokes* • Included as a single parameter with strokes

Transient ischemic attack rates for rivaroxaban and apixaban	• Imputed from relative stroke to transient ischemic attack rates in RE-LY* • Imputed from relative stroke to transient ischemic attack rates in US population
Congestive heart failure rates for rivaroxaban and apixaban	• Imputed from relative stroke to congestive heart failure rates in RE-LY* • Imputed from relative stroke to congestive heart failure rates in US population
Pulmonary embolism rates for rivaroxaban and apixaban	• Imputed from relative stroke to pulmonary embolism rates in RE-LY* • Imputed from relative stroke to pulmonary embolism rates in US population • Pulmonary embolisms excluded from model
Probability of death from stroke	• Included as a separate parameter* • Deaths from strokes included as part of general vascular mortality
Probability of death from pulmonary embolism	• Included as a separate parameter* • Deaths from pulmonary embolisms included as part of general vascular mortality
Major bleed stratification	• Stratified ($CHADS_2$ 1,2,≥3)* • Unstratified
Major bleed stratification imputation (rivaroxaban)	• Imputed from RE-LY and ARISTOTLE* • Imputed from RE-LY • Imputed from ARISTOTLE • Imputed from US population data
Minor bleed rates	• Included in model* • Excluded from model
Utility values	• Values in table 13* • Values derived from Freeman study[122]
Stroke severity	• Assumed to be independent of treatment* • Taken from the individual arms of the trials
Treatment benefit duration	• Lifetime* • Study time horizon • Linear reduction of treatment benefit to 0 over the 10 years following the study time horizon

*Base case assumptions.

Table 15 - Lifetime estimates of event rates, net health benefits, and incremental differences versus comparator, derived from probabilistic sensitivity analysis

Referent	**Mean estimate (95% central range)**	**Mean difference (95% central range)**	**Comparator**
Quality-adjusted life-years (QALYs)			
Warfarin	5.6364 (5.5461, 5.7326)	-0.0948 (-0.2422, 0.0523)	Rivaroxaban
Rivaroxaban	5.7312 (5.6305, 5.8342)	-0.0107 (-0.1644, 0.1443)	Dabigatran
Dabigatran	5.7419 (5.6519, 5.8535)	-0.0242 (-0.1744, 0.1298)	Apixaban
Apixaban	5.7661 (5.6518, 5.8805	0.1297 (-0.0293, 0.2646)	Warfarin
Life years			
Warfarin	9.6375 (9.4982, 9.7367)	-0.0917 (-0.2861, 0.1201)	Rivaroxaban
Rivaroxaban	9.7292 (9.5793, 9.8652)	-0.0336 (-0.2411, 0.1716)	Dabigatran
Dabigatran	9.7628 (9.6043, 9.8929)	-0.0454 (-0.2541, 0.1466)	Apixaban
Apixaban	9.8082 (9.6546, 9.9462)	0.1707 (-0.0311, 0.3617)	Warfarin
Stroke or systemic embolism			
Warfarin	0.3030 (0.2636, 0.3388)	0.0204 (-0.0334, 0.0739)	Rivaroxaban
Rivaroxaban	0.2826 (0.2381, 0.3192)	0.0312 (-0.0288, 0.0827)	Dabigatran
Dabigatran	0.2514 (0.2126, 0.3012)	0.0501 (-0.0014, 0.0986)	Apixaban
Apixaban	0.2013 (0.1688, 0.2543)	-0.1017 (-0.1539, -0.0502)	Warfarin
Transient ischemic attack			
Warfarin	0.1230 (0.0905, 0.1579)	0.0312 (-0.0187, 0.0844)	Rivaroxaban
Rivaroxaban	0.0918 (0.0697, 0.1234)	-0.0056 (-0.0565, 0.0462)	Dabigatran
Dabigatran	0.0974 (0.0694. 0.1279)	0.0201 (-0.0342, 0.0691)	Apixaban
Apixaban	0.0773 (0.0545, 0.1041)	-0.0457 (-0.0933, 0.0082)	Warfarin
Intracranial haemorrhage			
Warfarin	0.0727 (0.0637, 0.0808)	0.0142 (-0.0016, 0.0258)	Rivaroxaban
Rivaroxaban	0.0585 (0.0520, 0.0663)	0.0184 (0.0001, 0.0253)	Dabigatran
Dabigatran	0.0401 (0.0353, 0.0467)	-0.0019 (-0.0153, 0.0144)	Apixaban
Apixaban	0.0420 (0.0333, 0.0470)	-0.0307 (-0.0462, -0.0134)	Warfarin
Major bleed (including intracranial haemorrhage)			

Warfarin	0.3070 (0.2620, 0.3470)	-0.0214 (-0.0761, 0.0360)	Rivaroxaban
Rivaroxaban	0.3284 (0.2828, 0.3743)	0.0174 (-0.0367, 0.0691)	Dabigatran
Dabigatran	0.3110 (0.2741, 0.3632)	0.0694 (0.0135, 0.1148)	Apixaban
Apixaban	0.2416 (0.2052, 0.2778)	-0.0654 (-0.1162, -0.0100)	Warfarin
Non-fatal myocardial infarction			
Warfarin	0.0665 (0.0468, 0.0864)	0.0071 (-0.0132, 0.0293)	Rivaroxaban
Rivaroxaban	0.0594 (0.0429, 0.0798)	-0.02191 (-0.0432, -0.0018)	Dabigatran
Dabigatran	0.0813 (0.0633, 0.0994)	0.0189 (-0.0022, 0.0330)	Apixaban
Apixaban	0.0624 (0.0441, 0.0856)	-0.0043 (-0.0216, 0.0187)	Warfarin

Results

The results of the simulations at 2 years matched the results of each trial. No value deviated by more than 3.2%, a level of variability that would be expected given the stochastic nature of the simulation. At 2 years, apixaban accrued 0.15 more quality-adjusted life-weeks than dabigatran, 0.26 more than rivaroxaban, and 0.78 more than warfarin.

Clinical outcomes and net health benefit

In the base-case analysis, apixaban, dabigatran and rivaroxaban extended life by 2.05, 1.51 and 1.10 months respectively, compared with warfarin (table 15, P107). The corresponding incremental net health benefits were 0.1297 (95% central range [CR] -0.0296 to 0.2644), 0.1055 (95% CR -0.0481 to 0.2477) and 0.0948 (95% CR -0.0523 to 0.2422) QALYs. In pairwise comparisons, using warfarin as the comparator, apixaban, dabigatran and rivaroxaban were associated with a positive incremental net health benefit in 90%, 84% and 82% of simulations, respectively. Using rivaroxaban as a comparator, apixaban and dabigatran were associated with an incremental net health benefit in 71% and 61% of simulations, and finally apixaban was associated with an incremental net health benefit against dabigatran in 65% of simulations.

Lifetime incidences of stroke or systemic embolism were 33.6% lower with apixaban, 17.0% lower with dabigatran and 6.7% lower with rivaroxaban, when compared to warfarin. Lifetime incidences of major haemorrhagic events were 21.3% lower with apixaban, but 7.0% and 1.3% higher with rivaroxaban and dabigatran, respectively. Incidences of myocardial infarction were 10.7% lower with rivaroxaban and 6.5% lower with apixaban, but 22.3% higher with dabigatran.

The relative effects of each treatment on the two constructs of the QALY, health-related quality of life and life-years gained, is illustrated in figure 9 (P110).

Univariate parameter sensitivity analysis results are presented in the form of tornado plots (figure 10, P111).

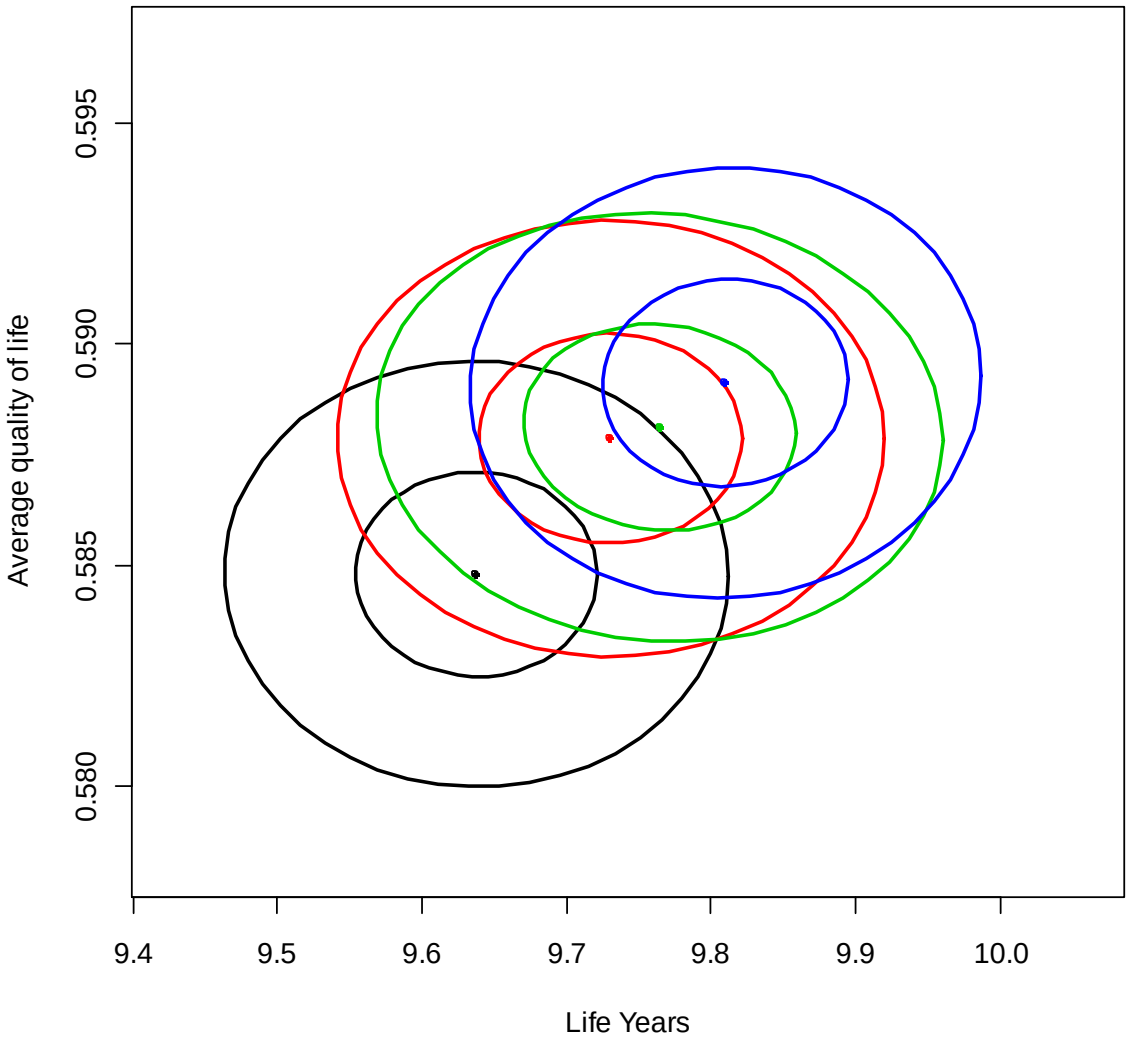
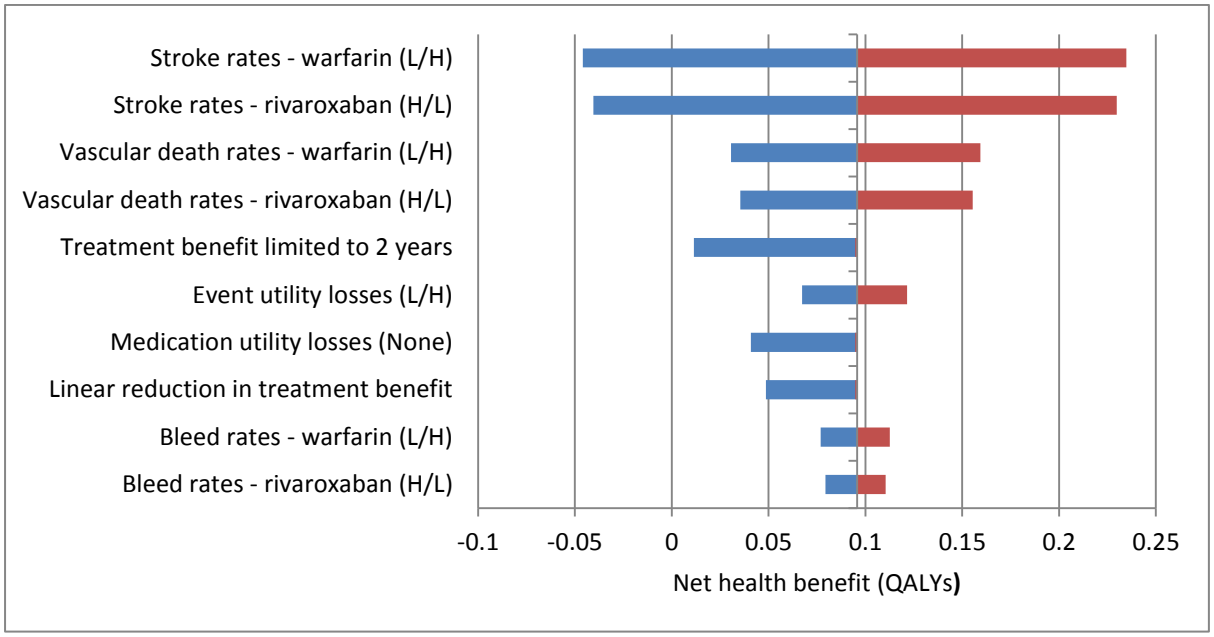


Figure 9

Structural sensitivity analysis

In the probabilistic sensitivity analysis of structural uncertainty, the base-case ordering of QALYs (apixaban, dabigatran, rivaroxaban, warfarin, in descending order) was replicated in 65.1% of the simulations. The alternative ordering (dabigatran, apixaban, rivaroxaban, warfarin) occurred in 17.2% of simulations, and the ordering (apixaban, rivaroxaban, dabigatran, warfarin) occurred in 13.4% of simulations. No other ordering occurred in more than 1% of simulations. Overall, apixaban accrued the highest number of QALYs in 79.9% of the simulations, dabigatran in 18.0% and rivaroxaban in 2.1%, with warfarin never accruing the highest number.

Warfarin vs. rivaroxaban



Warfarin vs. dabigatran

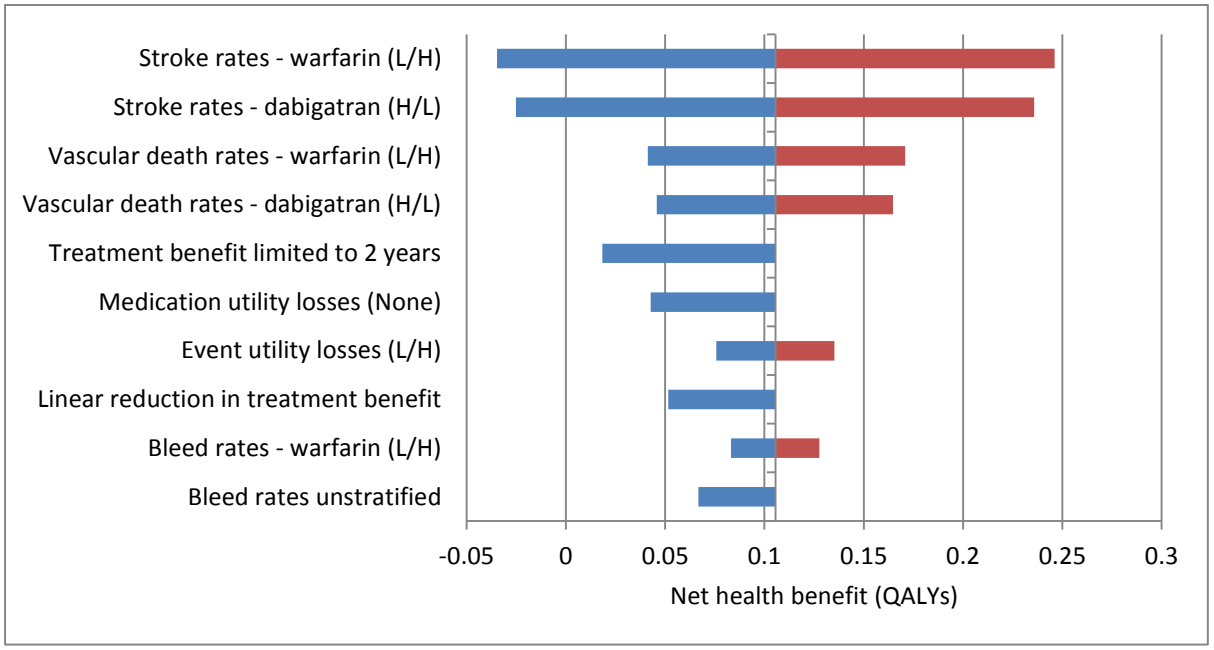


Figure 10

Using a Markov model, instead of a discrete event simulation, led to changes in both event rates and the numbers of QALYs and life-years accrued. However, the same ordering was maintained, with apixaban the most effective with 8.49 QALYs, then dabigatran with 8.35, rivaroxaban with 8.28 and warfarin with 8.08 QALYs.

Subgroup analyses

Among the subgroups analysed, the ordering of medicines according to mean QALYs and probability of being most effective was consistent with the base-case analysis (table 16, P113 & figure 11). Apixaban had the highest probability of being the most effective in patients with impaired renal function, and the lowest in older populations (≥75 years), but these probabilities were over a very narrow range (50.1% to 61.6%).

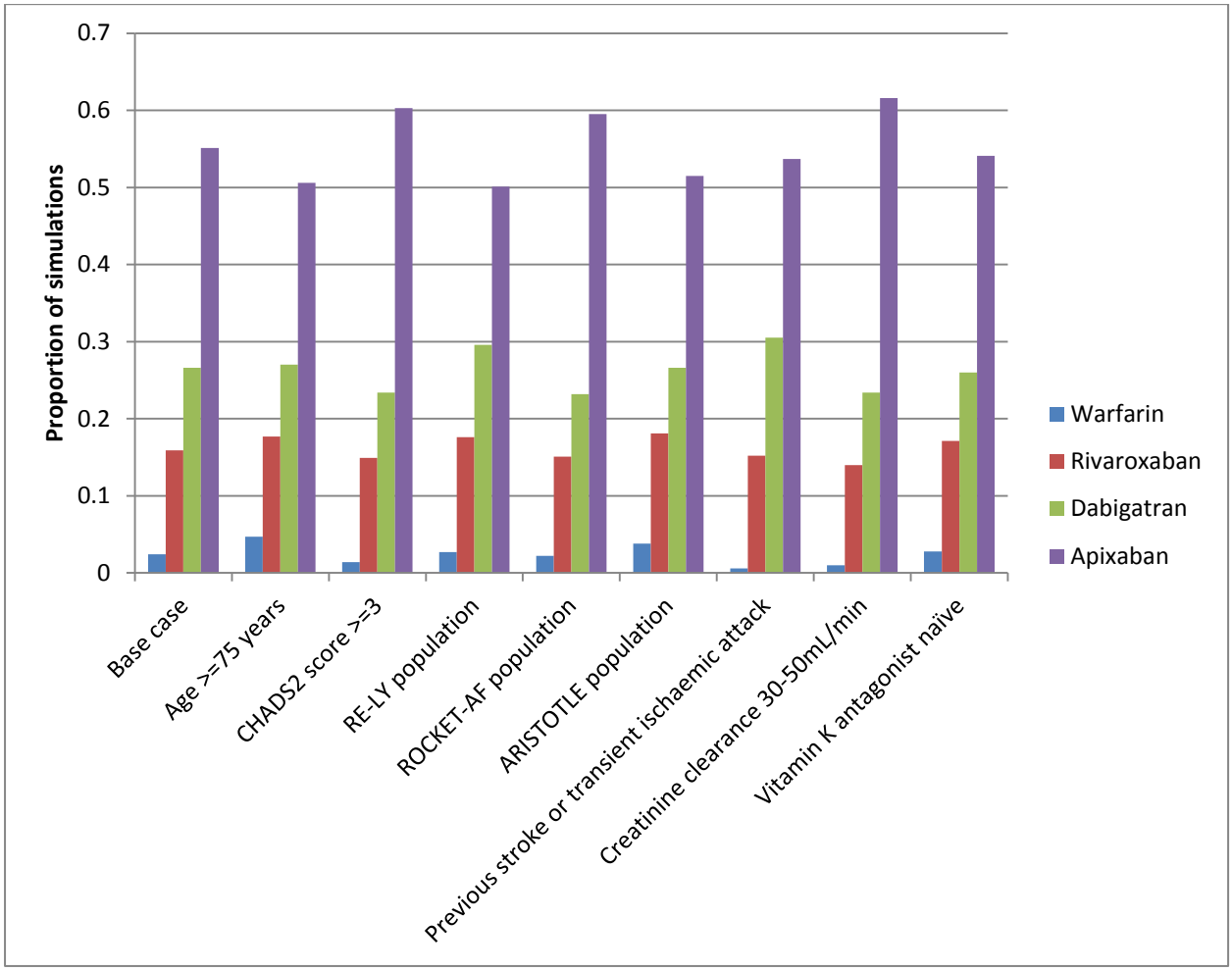


Figure 11

Table 16 - Net benefit results for subgroups, based on probabilistic sensitivity analysis

Subgroup	**Warfarin QALYs (Probability most effective)**	**Rivaroxaban QALYs (Probability most effective)**	**Dabigatran QALYs (Probability most effective)**	**Apixaban QALYs (Probability most effective)**
Base case	5.6374 (0.024)	5.7349 (0.159)	5.7450 (0.266)	5.7667 (0.551)
Age ≥75 years	3.8481 (0.047)	3.9403 (0.177)	3.9479 (0.270)	3.9719 (0.506)
$CHADS_2$ score ≥3	5.4818 (0.014)	5.5899 (0.149)	5.6170 (0.234)	5.6485 (0.603)
RE-LY population	5.6575 (0.027)	5.7460 (0.176)	5.7690 (0.296)	5.7837 (0.501)
ROCKET-AF population	5.5820 (0.022)	5.6662 (0.151)	5.6775 (0.232)	5.7101 (0.595)
ARISTOTLE population	5.6510 (0.038)	5.7472 (0.181)	5.7611 (0.266)	5.7858 (0.515)
Previous stroke or transient ischemic attack	5.4597 (0.006)	5.5448 (0.152)	5.5621 (0.305)	5.5817 (0.537)
Creatinine clearance 30-50mL/min	5.5609 (0.010)	5.6642 (0.140)	5.6773 (0.234)	5.7014 (0.616)
Vitamin K antagonist naïve	5.6409 (0.028)	5.7290 (0.171)	5.7421 (0.260)	5.7659 (0.541)

Discussion

Based on an accepted method of comparative effectiveness research (CER) that preserves the randomisation of treatment allocation, and which results in an adjusted, indirect comparison[138], apixaban appears as the most effective oral anticoagulant, followed by dabigatran, rivaroxaban then warfarin. Differences were driven principally by differential stroke rates and the risks of intracranial haemorrhage, which were lower for all newer agents compared with warfarin. This ordering remained consistent across patient subgroups, though the differences in net health benefits changed, with groups having lower risks of stroke associated with smaller differential QALYs.

There is no subgroup in which the probability of apixaban being the most effective is below 50%, and none where the probability of warfarin being the most effective is above 5%. The sensitivity analyses indicate that the parameters to which the outcome was most sensitive were stroke rates and vascular death rates.

We are aware of two other adjusted,[144,145] and one unadjusted[146] indirect treatment comparisons. Lip et al.[145] concluded that there were no discernible differences among treatments, while the analysis by Mantha et al.[144], despite being based on the same clinical trial data, indicated that apixaban was equally effective to dabigatran 150mg, more effective than rivaroxaban, and associated with less major bleeding than both. The crude estimates of net clinical benefit calculated by Banerjee et al.[146] for a Danish population are subject to bias, as the odds ratios derived from the trials were not adjusted. None of the analyses modelled patients representative of the US atrial fibrillation population, used a preference-based, patient-centred outcome measure, used an appropriate time horizon of analysis or considered alternative scenarios of analysis.

Our analysis, by contrast, adopts a lifetime horizon, uses QALYs to synthesise the differential impacts of benefits and harms on health, and considers both structural and parameter uncertainty. QALYs may reveal differences among treatments which might be less apparent when considering individual clinical events. Adjusted, indirect treatment comparisons are accepted by healthcare decision-makers across several jurisdictions as the method of choice in situations where data from head-to-head trials are unavailable[147]. This Bayesian approach results in a meaningful outcome for prescribers, that is, the probability of a treatment being the best option. Judgements based on confidence intervals and hypothesis tests, based on the Frequentist notion of assuming the null hypothesis until sufficient evidence indicates

otherwise, can be criticised as inappropriate in this context, as decisions regarding treatment alternatives cannot be deferred as such evidence is unlikely to become available. Moreover, a lack of statistical significance does not necessarily imply equivalence in outcome[148].

There are potential caveats to our CER methodology, however. First, there are many important differences across trials in terms of their design (e.g. RE-LY being open-label and the use of sham INR testing only in ROCKET-AF), patient populations (e.g. higher risk of stroke, greater experience of previous stroke or TIA, and less time in INR range in ROCKET-AF) and reporting (e.g. difference in the definitions of some clinical events). These have been discussed extensively elsewhere[139,147,149] and, collectively, potentially undermine the assumption necessary for indirect comparison methodology, that any such differences do not affect the comparative effectiveness of the treatments being assessed. Although there is no method of testing the validity of this assumption, there is no prospect of a head-to-head comparison among newer anticoagulants and prescribers will in any case make qualitative judgements or naïve, unadjusted comparisons of competing treatment options. Our analysis is a more valid approach than simply comparing individual trials or trial arms in an unadjusted way[150].

Second, the modelled extrapolation to a lifetime horizon of analysis is necessary to reflect differential impacts of treatments on health and survival that extend beyond the protocol-defined trial follow-up period[151]. The analysis used a discrete event simulation method, given there are no obvious discrete disease states into which patients can be classified[27]. However, this required an assumption that risk equations derived from 2-year data apply beyond that time. Relaxing this assumption by analysis at 2-years resulted in the same rank ordering of net health benefits, as did the use of an alternative, Markov model structure.

Despite these limitations there was a high level of consistency in the ranking of treatment effectiveness across all the different simulations performed. Had the results been dependent on any specific modelling assumption, then this would be apparent in the sensitivity analysis. We can reasonably conclude that any biases in our analysis, if any are present, are inherent in the data and therefore impossible to correct under any modelling framework.

It is important to note that the analysis did not consider additional factors that impinge on treatment choice. These include cost-effectiveness[96], patient convenience, preference (or aversion) to individual treatments[152], the relative forgiveness of treatments to missed doses[153], the lack of antidotes to over-anticoagulation with the newer oral anticoagulants[154], the merits

or otherwise of patients being monitored regularly when prescribed warfarin, and longer term and rarer adverse events that might only become apparent with more extensive experience of use in routine practice[155-157]. Furthermore, the newer agents also interact with other drugs and there are specific safety considerations in certain vulnerable populations (e.g. the very elderly and those with severe renal impairment). There may also be sub-group(s) of patients, which were not explored in the present analysis – such as those with genetic polymorphisms of *CYP2C9* or *VKORC1* – in which the balance of harms and benefits differ significantly from the mean[128].

There is no doubt of the efficacy of the newer oral anticoagulants and the favourable risk-benefit profile when compared to warfarin in the pivotal trials; however there are important differences among the agents and, in the absence of a definitive trial, modelling offers the only practical approach to estimate their comparative effectiveness. Whilst inevitably such conclusions will need to be kept under review as the evidence matures, our analysis currently points to the likely superiority of apixaban over others.

Preface to Chapter 5

Following the previous two chapters, there was now an economic model suitable for the extrapolation of short-term data to a lifetime horizon, and an indirect comparison of all treatment alternatives where trial data was available. This meant all the necessary pre-requisites were in place for the full population pharmacokinetic-pharmacodynamic-pharmacoeconomic model to be constructed. This followed broadly three phases, with the first the PKPD based clinical trial simulations of both genotype-guided and clinically dosed warfarin. Since the population PKPD model available only produces an intermediate endpoint as an output, rather than relevant clinical outcomes, a meta-analysis was then conducted to estimate the link between the simulated output and the necessary event rates. Finally, these simulated event data were used to populate the discrete event simulation (chapter 3), with full costs and outcomes attached to each health state and event.

Chapter 5

Cost-effectiveness of pharmacogenetic guided warfarin therapy versus alternative anticoagulation in atrial fibrillation

Summary

Introduction: Warfarin dosing regimens informed by pharmacogenetic information have been suggested as an alternative to standard clinical dosing algorithms for patients with nonvalvular atrial fibrillation. They are anticipated to increase time within therapeutic range (TTR) and thus improve clinical outcomes. We compare the effectiveness and cost-effectiveness of genotype-guided warfarin with both standard clinically dosed warfarin and three new oral anticoagulants (dabigatran, rivaroxaban and apixaban).

Methods: A clinical trial simulation based on a PKPD model of S-warfarin was used to predict differences in TTR between genotype guided and clinically dosed warfarin. A meta-analysis of trials linking TTR with outcomes was conducted to obtain relative risks of different clinical events. Finally, an economic analysis (from the perspective of the UK National Health Service) was conducted, based on a discrete event simulation model to extrapolate event risks to a lifetime horizon. The patient population modelled was representative of the AF population in the UK and the main outcomes produced were quality-adjusted life-years and incremental cost-effectiveness ratios among the various treatment options.

Results: Neither dabigatran nor rivaroxaban were cost-effective options, with rivaroxaban dominated by apixaban and dabigatran extendedly dominated. Apixaban and genotype guided warfarin had positive incremental net benefits of 0.1298 and 0.0031 QALYs versus clinically dosed warfarin, respectively, with apixaban having an ICER of £20,671 per QALY gained versus genotype guided warfarin therapy, which in turn had an ICER of £13,226 per QALY gained versus clinically dosed warfarin.

Discussion: Clinical trial simulations based on pharmacological models offer a new way to obtain estimates of cost-effectiveness in circumstances where trial data are not available. Based on our simulations, apixaban appears to be the most cost-effective treatment, but this may only be the case in centres were INR is likely to be poorly controlled.

Introduction

Warfarin has been the mainstay of oral thromboprophylactic treatment for patients with atrial fibrillation[90]. However, due to certain limitations associated with warfarin therapy, in particular the variability in patient response and the need for frequent monitoring and dose adjustment to ensure optimal anticoagulation, there has recently been considerable interest in the development of alternative anticoagulants[91].

Dabigatran, rivaroxaban and apixaban are three such novel oral anticoagulants. Each have shown, in the RE-LY, ROCKET-AF and ARISTOTLE trials, respectively[92,134,135], to be at least non-inferior to warfarin with regards to stroke prophylaxis. There is evidence for reductions in intracranial haemorrhages with all three, though apixaban is the only one associated with a reduction in the risk of major bleeds and to significantly reduce all-cause mortality when compared with warfarin.

While these studies compared the novel anticoagulant to standard, dose-adjusted warfarin, there have been a number of approaches developed for increasing the effectiveness of warfarin, including self-monitoring of INR[158], and dosing based on pharmacogenetics[159]. The latter uses information on *CYP2C9* and *VKORC1* genotype for more personalised dosing algorithms, with the aim to achieving the correct maintenance dose more quickly than standard clinical dosing algorithms, with evidence the use of genotype data increases average time in therapeutic range[160]. There is thus the potential for new anticoagulants, which may appear to be cost-effective compared to standard warfarin therapy[161], not to be so against genotype-guided dosing.

There are challenges, however, in assessing both the clinical and cost-effectiveness of warfarin dosing algorithms. First, the large number of potential algorithms, both clinical and pharmacogenetic, means conducting trials that cover even a significant subset of these is impractical. Second, the potential differences in both benefits and costs are sufficiently small that the sample sizes necessary to identify significant differences would be prohibitively expensive. Third, and in contrast to the newer oral anticoagulants, clinical trials of warfarin pharmacogenetics have only been powered to detect differences in intermediate endpoints, such as time within INR range, and not clinical outcomes[162].

One way of dealing these difficulties is through clinical trial simulations, based on pharmacometric analysis of the dose-exposure-response relationship of warfarin. This

approach may be used to simulate the clinical outcomes of a large number of possible dosing algorithms[163], and may be extended to incorporate economic information and cost-effectiveness as an endpoint[46] (chapter 2).

Against this background, we present a novel application of pharmacokinetic-pharmacodynamic based clinical trial simulation to determine the cost-effectiveness of warfarin therapy. We simulated the costs and outcomes of a variety of both clinical and pharmacogenetic algorithms, and compared the cost-effectiveness of genotype- and clinical guided warfarin dosing algorithms, and the newer anticoagulants. The model was validated externally by comparing the simulated results with those from a trial comparing genotype guided and standard warfarin therapy[164].

Methods

The model consisted of three distinct stages. First, we conducted a clinical trial simulation of both genetic and clinical dosing algorithms for warfarin, based on a pharmacokinetic-pharmacodynamic (PKPD) model which produces an output of time in various INR ranges. A systematic review and meta-analysis of studies that reported on the association between INR values and clinical events provided the necessary link between time in INR range and the risks of these events. In the final stage, we used a discrete event simulation to extrapolate event rates to a lifetime horizon, and to facilitate comparison, via indirect methods, with dabigatran, rivaroxaban and apixaban.

The outputs of this simulation were net health benefits, measured as the differences between treatments in quality-adjusted life-years (QALYs), a preference based measure of quality of life, and a cost-effectiveness analysis, resulting in an incremental cost per QALY gained.

PKPD simulation

The simulation was based on a published population model of the pharmacokinetics and pharmacodynamics of S-warfarin[165], from which we obtained *CYP2C9* and *VKORC1* genotype prevalence. This model consists of a linear, single compartment PK model and a KPD model with two, three-state transit compartment chains, with the last states of each chain representing the effect sites. The chains are used to model the time delay between peak drug concentration and peak pharmacodynamic response. The distribution of baseline INR values was taken from a pharmacogenetic cohort study[166]. Daily INR estimates were simulated for 5,000 hypothetical patients and combined, using the method of Rosendaal et al[167], to give a proportion of time below, in and above therapeutic INR range (2.0-3.0) over three months.

In the base case analysis, and in line with common UK practice[168], patients were initiated with a 10mg dose of warfarin on days 1 and 2, then a 5mg dose on day 3. Maintenance doses were predicted using two algorithms developed by the International Warfarin Pharmacogenetics Consortium (IWPC); one based on clinical and demographic variables, and the other also on genetic information[128]. The predicted maintenance doses were estimated from:

Pharmacogenetic algorithm:

√(Weekly Dose) = 5.6044 – 0.2614*(age in decades) + 0.0087*(height) + 0.0128*(weight) – 0.8677*(*VKORC1 AG*) – 1.6974*(*VKORC1 AA*) – 0.4854*(*VKORC1* unknown) – 0.5211*(*CYP2C9 *1*2*) – 0.9357*(*CYP2C9 *1*3*) – 1.0616*(*CYP2C9 *2*2*) – 1.9206*(*CYP2C9 *2*3*) – 2.3312*(*CYP2C9 *3*3*) – 0.2188*(*CYP2C9* unknown) – 0.1092*(Asian) – 0.2760*(Black or African American) – 0.1032*(Missing or Mixed Race) + 1.1816*(Enzyme inducer) – 0.5503*(amiodarone)

Clinical algorithm:

√(Weekly Dose) = 4.0376 – 0.2546*(age in decades) + 0.0118*(height) + 0.0134*(weight) – 0.6752*(Asian) + 0.4060*(Black or African American) – 0.0443*(Missing or Mixed Race) + 1.2799*(Enzyme inducer) – 0.5695*(amiodarone)

Dose adjustments were then made at each scheduled clinic visit using a widely used nomogram, which also schedules the date of the next clinic visit[178]. A variety of other algorithms were also simulated to explore their differences (table 17, P124).

To better reflect the real world use of warfarin, patient non-adherence was incorporated in two ways; by assuming a fixed proportion of doses were missed at random, based on data from the IN-RANGE study[181], and that variability in the timing of dosing could be approximated by a normal distribution with a standard deviation of 2 hours.

Meta-analysis

We updated a published systematic review and meta-analysis of the association between INR and risk of strokes and major bleeds in patients with AF receiving warfarin[182]. MEDLINE was searched for relevant articles published between January 1, 2000 and June 15, 2012, and manual searches were conducted of references in retrieved articles. Search terms were: atrial fibrillation, anticoagulant and warfarin. Studies of patients with atrial fibrillation who were receiving warfarin were included if the target INR range was 2.0-3.0 and if data were presented on strokes or bleeds stratified by time in therapeutic range (TTR).

In the base case analysis, odds ratios for clinical events, stratified by INR within, above (>3.0) or below (<2.0) target range, where multiplied by the output from the PKPD model to generate simulated relative risks of thromboembolic and bleeding events over the first three

Table 17 - Simulated warfarin algorithms

Initiation algorithm	**Predicted maintenance dose**	**Maintenance dose adjustment**
10mg-10mg-5mg (Days 1,2,3)†	IWPC clinical algorithm[128]†	Ansell[178]†
10mg-5mg-5mg (Days 1,2,3)	IWPC pharmacogenetic algorithm[128]*†	Wilson[179]
5mg-5mg-5mg (Days 1,2,3)	Anderson[174]*	Keeling[180]
Fennerty[169]	Gage[175]	
Meckley[170]*	Wadelius[160]*	
Avery[171]*	Zhu[176]*	
Hillman[172]*	Solomon[177]	
Kovacs[173]		
Anderson[174]*		

*These algorithms made use of genetic information.

†Algorithms used in base case analysis.

months of treatment with each algorithm. The impact of subdividing INR ranges to smaller increments was tested in a sensitivity analysis.

Discrete event simulation

To extrapolate to a lifetimes horizon, we used a discrete event simulation model we have described previously (chapter 3), which simulates clinical events and outcomes for each patient, and allows explicit incorporation of parameters and structural uncertainty[27]. The risks of events at any given time are determined by a patient's characteristics, which are updated over time and as events occur.

Interventions

We modelled clinical- and genotype-guided, dose-adjusted warfarin, dabigatran etexilate (150mg twice daily), rivaroxaban (20mg once daily) and apixaban (5mg twice daily).

Model population

For the base case analysis, patients' baseline characteristics were assumed to be uncorrelated and follow the average profile of the UK atrial fibrillation population[124]. The modelled population had a mean age of 72.3 years, with 35.0%, 35.1%, 17.0%, 9.3%, 3.2% and 0.5% having $CHADS_2$ (congestive heart failure, hypertension, Age≥75, Diabetes mellitus, prior Stroke/transient ischaemic attack) scores of 1 to 6, respectively[124].

For each treatment, identical cohorts of 100,000 patients were generated. Each patient was given a simulated set of characteristics consisting of the presence or absence (at the start of the simulation) of the following: hypertension, diabetes mellitus, congestive heart failure, prior stroke, prior transient ischemic attack, prior myocardial infarction and prior intracranial haemorrhage, drawn from binomial distributions based on the probability of having each condition at baseline.

Clinical parameters

Annualised clinical event rates for apixaban, dabigatran, rivaroxaban and warfarin therapy were calculated from an indirect comparison of the RE-LY[92], ROCKET-AF[134] and ARISTOTLE[135] studies, as described previously (chapter 4). Events rates for pharmacogenetic-guided warfarin dosing over the first three months of therapy were calculated by multiplying the event rates for clinical algorithm dosed warfarin by the relative

risks from our clinical and pharmacogenetic warfarin simulations and meta-analysis, as described above. Event rates after 3 months were assumed to be the same as clinical algorithm dosed warfarin, as any benefits of pharmacogenetic testing are likely to occur principally during the initiation phase[183]. All clinical parameters are presented in table 18 (P127).

Utilities and costs

Utilities for AF and disutilities associated with the medications and clinical events came from the European Stroke Prevention Study, a US Medical Expenditure Panel Survey and published data, as described previously (chapter 3). Multiple utility decrements for a given patient were assumed to be additive.

Costing was performed from the perspective of the UK National Health Service, following the methodology of our evaluation of dabigatran (chapter 3). Event costs were based on NHS reference costs, following a National Institute for Health and Clinical Excellence costing template for long term care[184], and inflated to 2011 GBP (£) values. Drug costs were taken from the British National Formulary[75], however, as apixaban has yet to be licensed for use in AF, we assumed the same percentage reduction in price would occur as happened when dabigatran and rivaroxaban gained their respective licensed extensions. This assumption was tested in a sensitivity analysis. Full costs and utility parameters are given in table 19 (P129).

Costs, life years and QALYs were discounted at 3.5% per annum to present value, but there was no discounting of clinical events[25].

Sensitivity analysis

Univariate sensitivity analyses were performed for each parameter in the discrete event simulation, to assess the stability of the results. Where available, ranges where based on 95% confidence intervals, otherwise we assumed plausible percentage ranges. We also varied the duration of benefit from genetic testing, from one month, to a linear reduction to zero over the course of a year.

A probabilistic sensitivity analysis was also performed, using a Monte Carlo simulation of 2,000 sets of parameters sampled from appropriate distributions. This provided estimates of the 95% central ranges (2.5th to 97.5th percentile) for clinical event rates and net health benefits, and the probabilities of each treatment option resulting in the highest net health

Table 18 - Clinical parameters used in the discrete event simulation

Parameter	**Warfarin**	**Warfarin PGx**	**Dabigatran**	**Rivaroxaban**	**Apixaban**	**Aspirin**
Stroke ($CHADS_2$ score ≤ 1)*	0.00921	0.00921	0.00536	0.00750†	0.00678	0.01485
Stroke ($CHADS_2$ score 2)*	0.01405	0.01406	0.00824	0.01255	0.01211	0.02265
Stroke ($CHADS_2$ score 3)*	0.01957	0.01958	0.01164†	0.01335	0.01133†	0.03157
Stroke ($CHADS_2$ score 4)*	0.03119	0.03120	0.02154†	0.02442	0.02097†	0.05030
Stroke ($CHADS_2$ score 5)*	0.02899	0.02900	0.02398†	0.02785	0.02334†	0.04676
Stroke ($CHADS_2$ score 6)*	0.03639	0.03641	0.03098†	0.03511	0.03015†	0.05869
Systemic embolism*	0.00135	0.00135	0.00113	0.00031	0.00115	0.00217
Pulmonary embolism*	0.00078	0.00078	0.00114	0.00091‡	0.00060‡	0.00126
Transient ischemic attack*	0.00839	0.00839	0.00723	0.00662‡	0.00616‡	0.01354
Myocardial infarction*	0.00763	0.00763	0.01008	0.00620	0.00666	0.00763
Congestive heart failure*	0.00619	0.00619	0.00482	0.00488‡	0.00454‡	0.00619
Vascular death (excluding stroke and systemic and pulmonary embolism)*	0.02281	0.02281	0.02078	0.02155	0.02118	0.02281
Probability of death from stroke or systemic embolism	0.25457	0.25457	0.25457	0.25457	0.25457	0.25457
Probability of death from pulmonary embolism	0.15909	0.15909	0.15909	0.15909	0.15909	0.15909
Major bleed ($CHADS_2$ score ≤ 1)*	0.02612	0.02458	0.01981	0.02248†	0.01590	0.01146
Major bleed ($CHADS_2$ score 2)*	0.03175	0.02988	0.02916	0.03379†	0.02434	0.01393
Major bleed ($CHADS_2$ score ≥ 3)*	0.04433	0.04171	0.04674	0.04799†	0.03061	0.01944
Probability that major bleed is intracranial haemorrhage	0.23361	0.23361	0.10234	0.14947	0.14068	0.23361
Minor bleed*	0.16560	0.15583	0.15020	0.17140	0.11783	0.07263
Diabetes*	0.0141	0.0141	0.0141	0.0141	0.0141	0.0141
Hypertension*	0.0323	0.0323	0.0323	0.0323	0.0323	0.0323
Probability of discontinuation (year 1)*	0.14466	0.14466	0.22048	0.14481	0.14232	N/A

Probability of discontinuation (year 2 onwards)*	0.06760	0.06760	0.06695	0.07224	0.04975	N/A

*Presented as rates per 100 person years.

†Where stratified event rates were not available, unknown stratified risks were imputed based on the assumption that the relative risks of events for patients with different $CHADS_2$ scores would be independent of treatment.

‡Imputed, based on the relative risks of different events from the RE-LY study, on the assumption that the relative risks of different thromboembolic events would be independent of treatment.

Table 19 - Cost, health state utility and discount rate parameters used in the discrete event simulation

Parameter	Value	Probabilistic sensitivity analysis distribution
Baseline characteristics		
Hypertension*	14283/18113	Beta (14283,3830)
Diabetes*	4221/18113	Beta (4221,13892)
Heart failure*	5793/18113	Beta (5793,12320)
Prior stroke*	2273/18113	Beta (2273,15840)
Prior transient ischaemic attack*	1663/18113	Beta (1663,16450)
Prior myocardial infarction*	3005/18113	Beta (3005,15108)
Prior intracranial haemorrhage*	713/18113	Beta (713,17400)
Health state utilities		
Atrial fibrillation (age 67)	0.774	1-Gamma (43.06,0.0052)
Stroke (permanent disutility)†	0.233	Normal (0.233,0.0032)
Stroke (temporary disutility) †	0.1385	Normal (0.1385,0.01)
Stroke (temporary duration, years)†	1/12	Uniform (0,0.183)
Myocardial infarction (permanent disutility)	0.0409	Normal (0.0409,0.002)
Myocardial infarction (temporary disutility)	0.1247	Normal (0.1247,0.01)
Myocardial infarction (temporary duration, years)	1/12	Uniform (0,0.183)
Intracranial haemorrhage (permanent disutility)	0.0524	Normal (0.0524,0.001)
Pulmonary embolism (temporary disutility)	0.1385	Normal (0.1385,0.01)
Pulmonary embolism (temporary duration, years)	1/12	Uniform (0,0.183)
Transient ischemic attack (temporary disutility)	0.1032	Normal (0.1032,0.01)
Transient ischemic attack (temporary duration, years)	5/365	Uniform (0,0.027)
Major bleed (temporary disutility)	0.1385	Normal (0.1385,0.01)
Major bleed (temporary duration, years)	1/12	Uniform (0,0.183)
Minor bleed (temporary disutility)	0.06	Normal (0.06,0.01)
Minor bleed (temporary duration, years)	5/365	Uniform (0.0.027)
Warfarin disutility	0.013	Gamma (1.3,0.01)
Dabigatran/rivaroxaban/apixaban disutility	0.002	Gamma (0.2,0.01)
Aspirin disutility	0.002	Gamma (0.2,0.01)

Costs		
Stroke – year 1†	£11,228.93	Gamma (103.97,103.97)
Stroke – subsequent years†	£2,962.07	Gamma (54.42,54.42)
Myocardial infarction – year 1	£2,510.40	Gamma (61.32,40.93)
Myocardial infarction – subsequent years	£882.80	Gamma (35.14,25.12)
Pulmonary embolism	£1,511.12	N/A
Transient ischaemic attack	£887.26	N/A
Major bleed	£1,794.12	N/A
Minor bleed	£99.23	N/A
Proton pump inhibitors (1 year)	£197.24	N/A
Warfarin – drugs (1 year)	£41.23	Uniform (32.98,49.48)
Warfarin – monitoring (1 year)	£198.39	Gamma (202.59,0.979)
Dabigatran (1 year)	£919.80	N/A
Rivaroxaban (1 year)	£766.50	N/A
Apixaban (1 year)	£878.35	N/A
Aspirin (1 year)	£7.39	Gamma (73.9,0.1)
Cost of genetic test	£20.00	N/A
Discount rate		
Utilities	3.5%	N/A
Costs	3.5%	N/A

*Proportion in initial population.

†Includes both stroke and systemic emboli.

benefit, and being the most cost-effective at different threshold values. Probabilities of cost-effectiveness are presented through multiple cost-effectiveness acceptability curves with the cost-effectiveness acceptability frontier (the optimal treatment choice at a given threshold) represented by the treatment with the highest probability at any given threshold[185].

To externally validate the simulated INR data, a separate analysis was conducted using INR data from the CoumaGen-II trial, a randomised controlled trial comparing a pooled set of 504 patients dosed with genotype guided algorithms with 1866 patients given standard care[164]. Time in therapeutic range in this trial was 11% higher after 1 month in the pharmacogenetic arm and 12% higher after 3 months.

Scenario analyses

Subgroup analyses were performed to calculate the net health benefits (and associated 95% central ranges) and incremental cost-effectiveness ratios (ICERs) in the following pre-specified populations: patients aged 75 or older; patients with a $CHADS_2$ score ≥ 3; patients who have previously had a stroke or transient ischemic attack; patients with impaired renal function (30-50mL/min creatinine clearance); and patients who were naïve to vitamin K antagonist treatment.

Results

PKPD simulation

In our base case and in comparison with the clinical algorithm, the use of the genetic IWPC algorithm led to an increase in time both within and below INR range, and a decrease in time above range (figure 12, P133). The effect was most pronounced over the first month of simulation, with roughly equal percentages of time above, below and in range for the different algorithms after this, as patients had achieved their maintenance dose.

This pattern was broadly replicated across the different algorithms simulated (table 17, P124), with genetic testing associated with higher times in range, mainly driven by both reaching the target range quicker and not overshooting that range during initiation. Across all algorithms simulated; the use of genotype data during the initiation phase only, for predicting maintenance doses only or for both, led to increases in times in range of 1.34%, 4.76% and 5.71%, respectively, over the first three months.

Meta-analysis

702 abstracts were screened and 153 full articles were retrieved. Of these, seven studies met all the inclusion criteria for updating the Reynolds meta-analysis[181] (table 20, P134). The computed odds for bleeding and thromboembolic events, according to INR ranges, are presented in table 21 (P136). Time in range data from our PKPD simulation are combined with these ratios, to give relative risks of thromboembolic events (RR=1.00047) and bleeds (RR=0.940997) for genotype guided versus clinically dosed warfarin.

Net-health benefits

In the base case analysis, genotype guided-warfarin, rivaroxaban, apixaban and dabigatran extended life by 0.003, 1.11, 2.06 and 1.47 months, respectively, compared with clinical algorithm dosed warfarin (table 22, P137). The corresponding incremental net benefits were 0.0031 (95% central range [CR] -0.1649 to 0.1327), 0.0957 (95% CR -0.0510 to 0.2431), 0.1298 (95% CR -0.0290 to 0.2638) and 0.1065 (95% CR -0.0493 to 0.2489) QALYs.

In pairwise comparisons, using clinical algorithm dosed warfarin as the comparator, genotype guided warfarin, rivaroxaban, apixaban and dabigatran were associated with a positive incremental net health benefit in 57%, 83%, 90% and 85% of the simulations respectively. Using genotype guided warfarin as a comparator, rivaroxaban, apixaban and dabigatran were

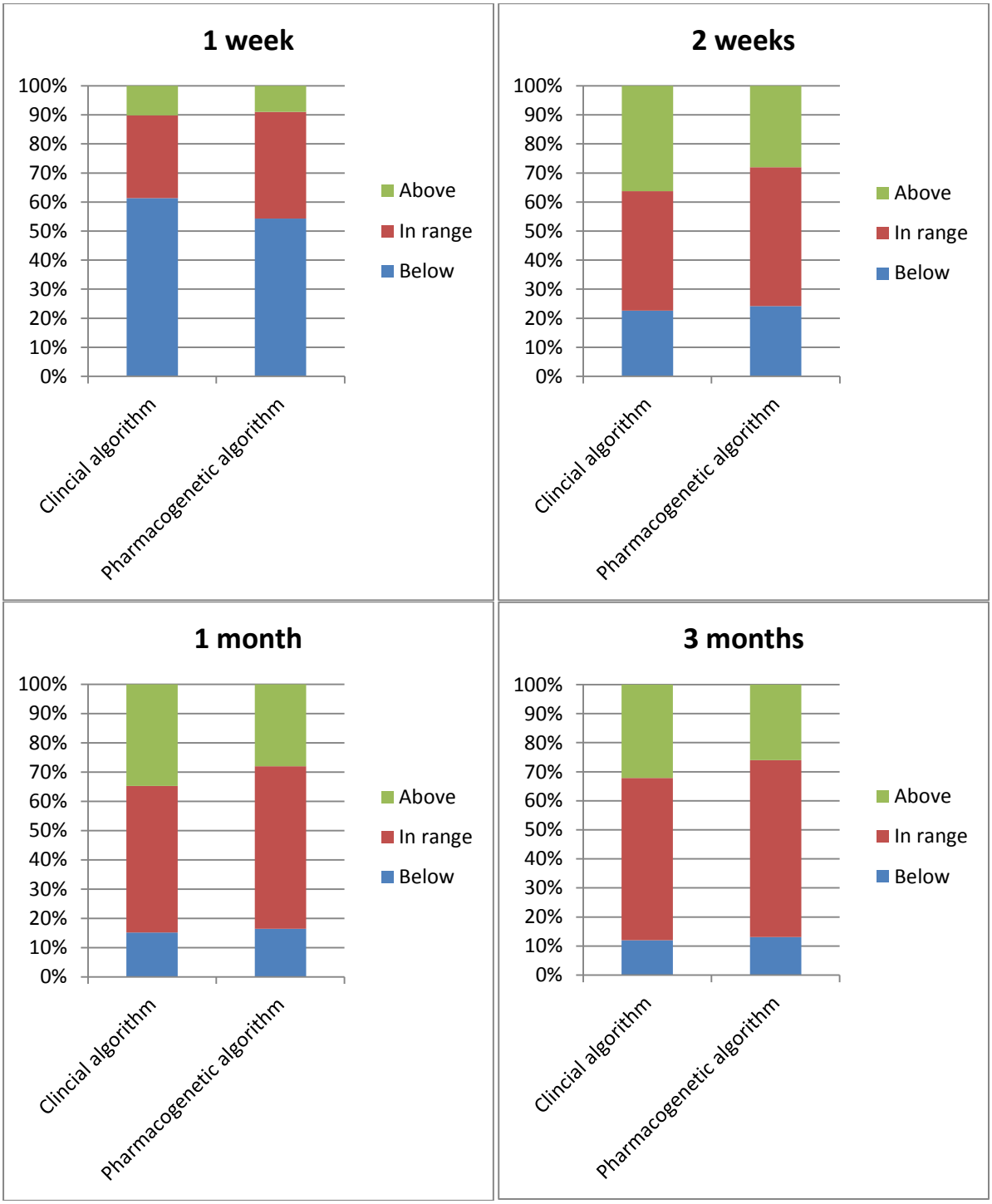

1 week
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Above
In range
Below
Clincial algorithm
Pharmacogenetic algorithm
2 weeks
1 month
3 months

Figure 12

Table 20 - Systematic review papers

Authors	Name	Population	Results
Masaki N, Suzuki M, Matsumara A, et al.	Quality of warfarin control affects the incidence of stroke in elderly patients with atrial fibrillation	120 Japanese AF patients (mean age 77)	Group A: n=57 mean TTR=81%, number of strokes=2, number of bleeds = 2 Group B: n=63, mean TTR=43%, number of strokes =11, number of bleeds = 2
Jones M, McEwan P, Morgan CLl, et al.	Evaluation of the pattern of treatment, level of anticoagulation control, and outcome of treatment with warfarin in patients with non-valvar atrial fibrillation: a record linkage study in a large British population	2223 non-valvar atrial fibrillation patients with no history of heart valve replacement	Odds ratios compared to whole sample: Bleeds In range: 0.945 Below range: 1.008 Above range: 1.097 Thromboembolic events In range: 0.891 Below range: 1.148 Above range: 0.948
Witt DM, Delate T, Clark NP, et al.	Outcomes and predictors of very stable INR control during chronic anticoagulation therapy	6073 patients with INR measured at least every eight weeks over a six month period	Group A: n=2504 TTR=100%, thromboembolic events=10, bleeds = 19 Group B: n=3569 TTR=46.9%, thromboembolic events=26, bleeds =101
Hylek EM, Evans-Molina C, Shea C, et al.	Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation	472 patients followed for one year	Bleed incidence rates per 100 person years INR<2: 4.11 2<=INR<=3: 3.78 3<INR<4: 15.78 INR>=4: 99.26
Matchar DB, Jacobson A, Dolor R, et al.	Effect of home testing of international normalized ratio on clinical events	2922 patients taking warfarin for atrial fibrillation	Group A: n=1463, patient years=4495, TTR=66.2%, strokes=31, major bleeds=147 Group B: n=1452, patient years=4235, TTR=62.4%, strokes=31, major bleeds=143
Singer DE, Chang Y, Fang MC, et	Should patient characteristics influence target	9217 atrial fibrillation patients	Odds ratios compared to INR 2-2.5 reference range Thromboembolic events

al.	anticoagulation intensity for stroke prevention in nonvalvular atrial fibrillation? The ATRIA study		INR<1.5: 7.22 1.5<=INR<2: 2.32 2.5<INR<=3: 1.09 3<INR<=3.5: 0.92 INR>3.5: 1.16 Bleeds INR<1.5: 1.37 1.5<=INR<2: 0.93 2.5<INR<=3: 1.60 3<INR<=3.5: 1.24 INR>3.5: 6.30
Hylek EM, Frison L, Henault LE, et al.	Disparate stroke rates on warfarin among contemporaneous cohorts with atrial fibrillation: potential insights into risk from a comparative analysis of SPORTIF III versus SPORTIF V	3665 patients with atrial fibrillation across the two trials	Percentage time above an INR value of 3 has a hazard ratio of 1.02 as a univariate predictor for stroke/systemic embolism

Table 21 - Meta-analysis of ORs for differing INR levels

	Stroke odds ratio (95% confidence interval)	Bleeds odds ratio (95% confidence interval)
INR < 1.5	4.26 (2.67, 6.81)	1.59 (1.01, 2.51)
1.5 <= INR < 2.0	2.19 (1.85, 2.59)	1.21 (0.78, 1.88)
3.0 < INR < 3.5	1.05 (0.84, 1.31)	2.01 (1.33, 3.04)
3.5 < INR <= 4.0	1.14 (0.93, 1.40)	3.82 (2.57, 5.66)
INR > 4.0	1.26 (0.71, 2.22)	31.76 (22.76, 44.32)

All given as odds ratios compared to a reference INR range of 2-3. In the base case analysis categories were combined to give odds ratios for below, in and above INR range.

Table 22 - Lifetime estimates of event rates, net benefits, and incremental differences versus comparator, derived from probabilistic sensitivity analysis

Referent	Mean (95% central range) estimate	Mean (95% central range) difference	Comparator
Quality-adjusted life-years (QALYs)			
Warfarin (clinical algorithm)	5.7209 (5.6299, 5.8162)	-0.0031 (-0.1649, 0.1327)	Warfarin PGx
Warfarin PGx	5.7240 (5.6352, 5.8204)	-0.0926 (-0.2345, 0.0533)	Rivaroxaban
Rivaroxaban	5.8166 (5.7158, 5.9191)	-0.0108 (-0.1632, 0.1450)	Dabigatran
Dabigatran	5.8274 (5.7362, 5.9382)	-0.0233 (-0.1729, 0.1318)	Apixaban
Apixaban	5.8507 (5.7374, 5.9647)	0.1298 (-0.0290, 0.2638)	Warfarin (clinical)
Life years			
Warfarin (clinical algorithm)	9.7220 (9.5825, 9.8216)	-0.0002(-0.1525, 0.1510)	Warfarin PGx
Warfarin PGx	9.7222 (9.5838, 9.8209)	-0.0926 (-0.2872, 0.1196)	Rivaroxaban
Rivaroxaban	9.8148 (9.6653, 9.9508)	-0.0328 (-0.2406, 0.1723)	Dabigatran
Dabigatran	9.8476 (9.6890, 9.9774)	-0.0457 (-0.2541, 0.1474)	Apixaban
Apixaban	9.8933 (9.7402, 10.0319)	0.1713 (-0.0316, 0.3599)	Warfarin (clinical)
Stroke or systemic embolism			
Warfarin (clinical algorithm)	0.3047 (0.2628, 0.3424)	-0.0004 (-0.0487, 0.0472)	Warfarin PGx
Warfarin PGx	0.3051 (0.2641, 0.3427)	0.0262 (-0.0310, 0.0788)	Rivaroxaban
Rivaroxaban	0.2789 (0.2369, 0.3190)	0.0290 (-0.0273, 0.0811)	Dabigatran
Dabigatran	0.2499 (0.2161, 0.2985)	0.0467 (-0.0033, 0.0974)	Apixaban
Apixaban	0.2032 (0.1678, 0.2567)	-0.1015 (-0.1539, -0.0496)	Warfarin (clinical)
Transient ischaemic attack			
Warfarin (clinical algorithm)	0.1210 (0.0932, 0.1571)	0.0010 (-0.0501, 0.0537)	Warfarin PGx
Warfarin PGx	0.1220 (0.0945, 0.1539)	0.0293 (-0.0208, 0.0825)	Rivaroxaban
Rivaroxaban	0.0927 (0.0689, 0.1226)	-0.0047 (-0.0564, 0.0471)	Dabigatran
Dabigatran	0.0974 (0.0710, 0.1277)	0.0124 (-0.0389, 0.0640)	Apixaban
Apixaban	0.0750 (0.0549, 0.1062)	-0.0260 (-0.0743, 0.0281)	Warfarin (clinical)
Intracranial haemorrhage			
Warfarin (clinical algorithm)	0.0715 (0.0634, 0.0803)	0.0066 (-0.0095, 0.0188)	Warfarin PGx

Warfarin PGx	0.0649 (0.0568, 0.0730)	0.0060 (-0.0100, 0.0191)	Rivaroxaban
Rivaroxaban	0.0589 (0.0528, 0.0657)	0.0176 (-0.0006, 0.0252)	Dabigatran
Dabigatran	0.0413 (0.0344, 0.0468)	0.0002 (-0.0144, 0.0153)	Apixaban
Apixaban	0.0411 (0.0329, 0.0471)	-0.0304 (-0.0466, -0.0137)	Warfarin (clinical)
Major bleed (including intracranial haemorrhage)			
Warfarin (clinical algorithm)	0.3059 (0.2599, 0.3471)	0.0321 (-0.0221, 0.0838)	Warfarin PGx
Warfarin PGx	0.2738 (0.2276, 0.3190)	-0.0544 (-0.1062, 0.0008)	Rivaroxaban
Rivaroxaban	0.3282 (0.2836, 0.3715)	0.0160 (-0.0382, 0.0675)	Dabigatran
Dabigatran	0.3122 (0.2730, 0.3657)	0.0714 (0.0143, 0.1138)	Apixaban
Apixaban	0.2408 (0.2069, 0.2783)	-0.0651 (-0.1168, -0.0091)	Warfarin (clinical)
Non-fatal myocardial infarction			
Warfarin (clinical algorithm)	0.0662 (0.0457, 0.0861)	0.0001 (-0.0210, 0.0229)	Warfarin PGx
Warfarin PGx	0.0661 (0.0449, 0.0863)	0.0071 (-0.0133, 0.0288)	Rivaroxaban
Rivaroxaban	0.0590 (0.0419, 0.0804)	-0.0221 (-0.0436, -0.0013)	Dabigatran
Dabigatran	0.0811 (0.0632, 0.0990)	0.0191 (-0.0026, 0.0327)	Apixaban
Apixaban	0.0620 (0.0447, 0.0861)	-0.0042 (-0.0216, 0.0184)	Warfarin (clinical)

associated with an incremental net health benefit in 74%, 85% and 78% of the simulation; using rivaroxaban as a comparator, apixaban and dabigatran were associated with an incremental net health benefit in 72% and 59% of the simulations; and finally apixaban was associated with an incremental net health benefit against dabigatran in 66% of the simulations.

Lifetime incidence of stroke or systemic embolism were 0.13% higher with genotype guided warfarin, 8.47% lower with rivaroxaban, 33.3% lower with apixaban and 18.0% lower with dabigatran, all compared to clinical algorithm dosed warfarin. Lifetime incidences of major haemorrhagic events were 21.3% lower with apixaban and 10.5% lower with genotype guided warfarin, but 7.3% and 2.1% higher with rivaroxaban and dabigatran, respectively. Incidences of myocardial infarction were 10.9% lower with rivaroxaban and 6.3% lower with apixaban, but 22.5% higher with dabigatran.

Cost-effectiveness

Total discounted lifetime costs for dabigatran, apixaban, rivaroxaban, genotype guided warfarin and clinical algorithm dosed warfarin were £8,426, £8,540, £9,112, £5,921 and £5,880. These were comprised mainly of drug and monitoring costs, which accounted for 43.2%, 45.1%, 39.7%, 22.7% and 22.1% of the overall costs, respectively. The costs of managing strokes or systemic emboli accounted for 41.1%, 35.9%, 45.8%, 57.2% and 56.6% of total costs; the remainder was accounted for by the costs of managing other events.

Rivaroxaban was dominated as a treatment option by both dabigatran and apixaban, meaning it was associated with both higher costs and lower QALYs. Dabigatran was extendedly dominated by apixaban, meaning that for cost-effectiveness thresholds where dabigatran would be selected over warfarin (clinical or genetic-guided dosing), apixaban would be preferred to dabigatran. Finally, the incremental cost-effectiveness ratio (ICER) for genotype guided warfarin versus clinical algorithm dosed warfarin was £13,226 per QALY gained, and the ICER for apixaban versus genotype guided warfarin therapy was £20,671 per QALY gained.

Sensitivity analysis

ICERs were most sensitive to changes in stroke rates, vascular death rates and the duration of treatment benefits. However, none of these changes altered the rank ordering in terms of either net health benefit or cost-effectiveness. The probabilistic sensitivity analysis (figure 13, P140) indicates that apixaban has the highest probability of being cost-effective at thresholds

of £13,782 per QALY gained, or higher. Considering a pairwise comparison of genotype guided warfarin and apixaban, apixaban has a higher probability of being cost-effective above thresholds of £20,600 per QALY gained. Genotype guided warfarin has a consistently higher probability of being cost-effective than clinical algorithm dosed warfarin above thresholds of £6,700 per QALY gained.

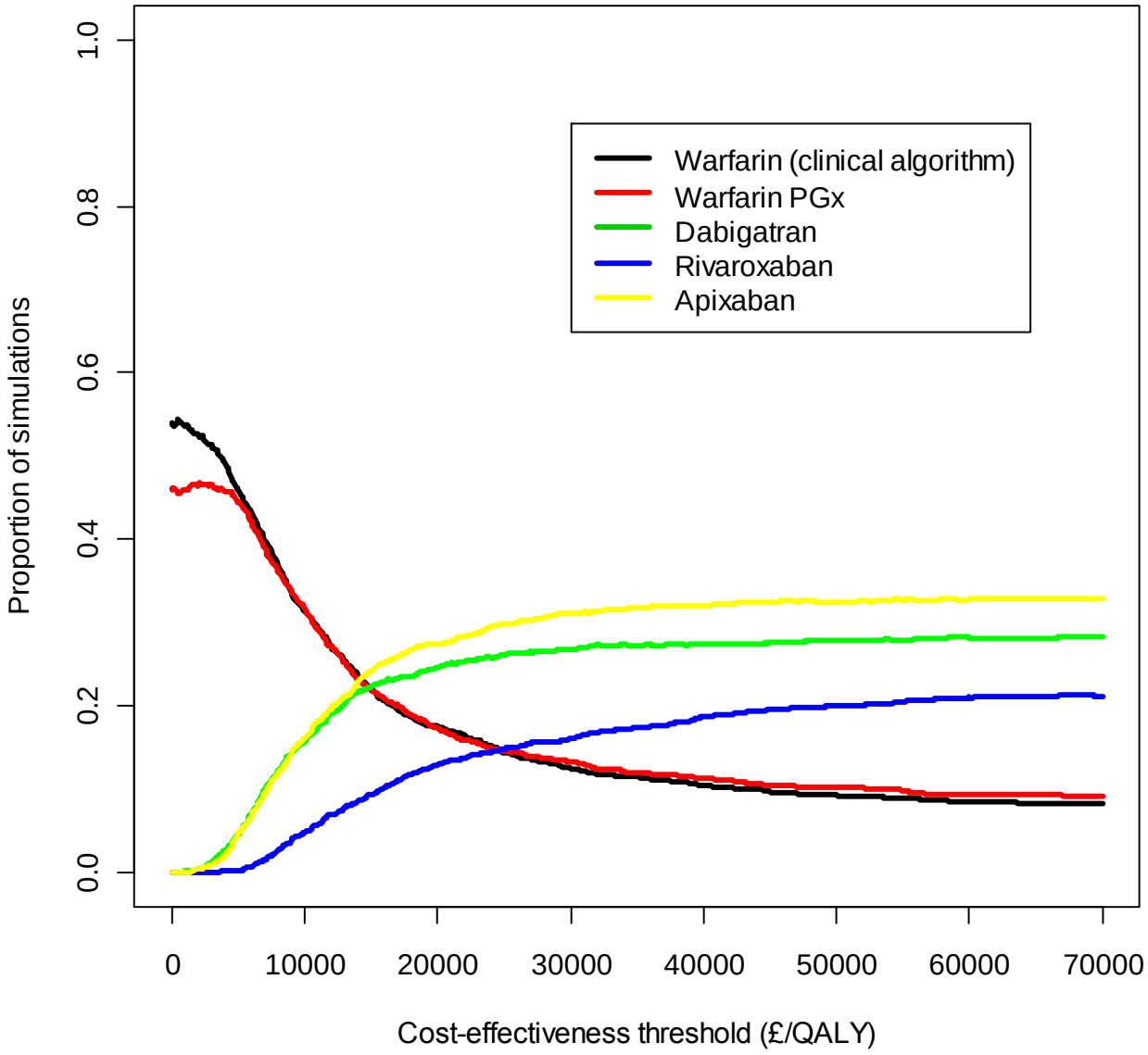


Figure 13

An increase of 25% in the price of apixaban increased the overall cost of the apixaban arm to £9,287, making it no longer cost-effective at £36,964 per QALY, compared with dabigatran. Dabigatran in turn had an ICER of £24,226 per QALY versus genotype guided warfarin. Conversely, a 25% reduction in the price of apixaban (from base-case) decreased overall costs to £7,793, with apixaban now dominating both dabigatran and rivaroxaban, with an ICER of £14,773 per QALY versus genotype guided warfarin. Finally, using the current list price of apixaban increased the cost of the arm to £9,796, with an ICER of £58,782 per QALY versus dabigatran.

Our analysis based on external INR data from the CoumaGen-II trial produced an ICER of £10,946 per QALY for genotype guided warfarin versus standard care, and an ICER of £21,874 per QALY for apixaban versus genotype guided warfarin.

Scenario analyses

Among the subgroups analysed, the mean net health benefits consistently showed the same ordering as the base case analysis (figure 14). Apixaban remained cost-effective compared to other options, and across all scenarios, with genotype guided warfarin also being consistently cost-effective compared with warfarin dosed according to clinical algorithm.

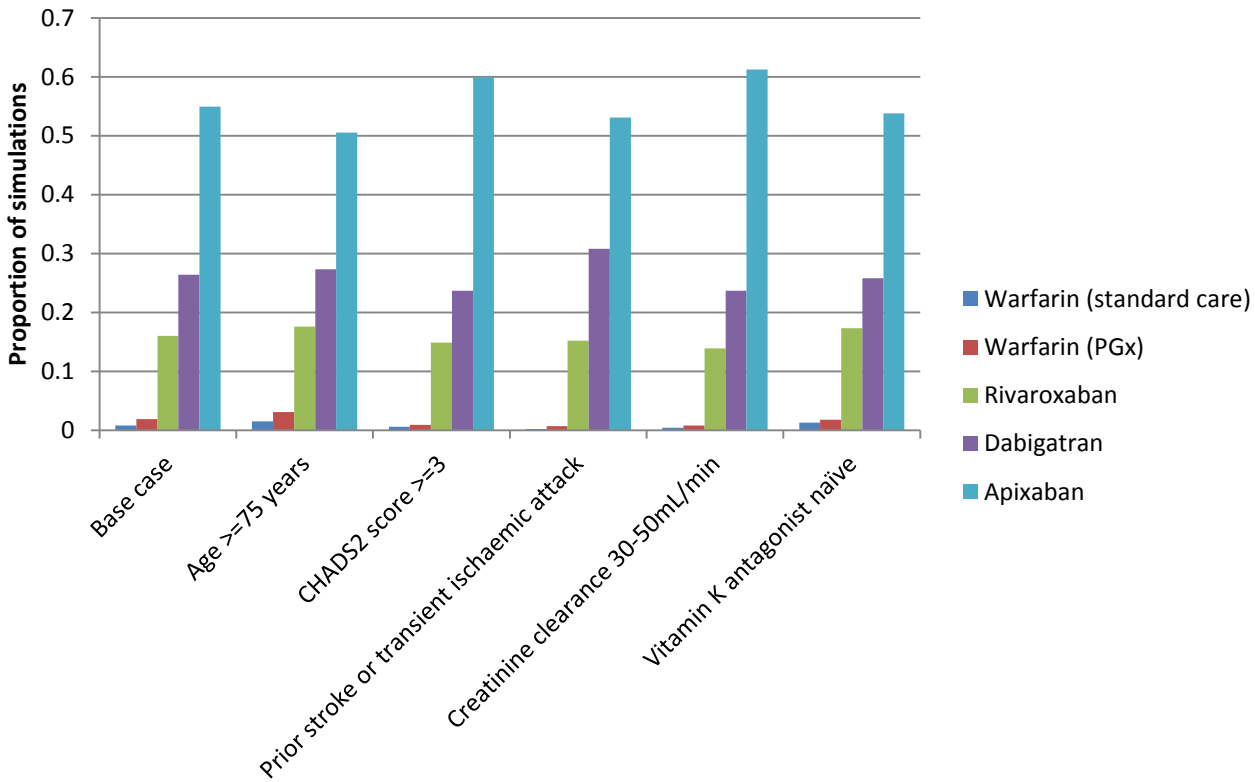


Figure 14

Discussion

Our results suggest that apixaban is the most promising of the novel anticoagulants available, in terms of estimated QALY gains, and that furthermore it has a high probability of being cost-effective against both genotype-guided and clinically-dosed warfarin. Importantly, the differences in costs and QALYs between the two warfarin dosing methods is much smaller than the difference between either and apixaban, meaning that decisions regarding the use of apixaban should not be affected by the type of warfarin dosing used as a comparator. However, it has been previously shown that the cost-effectiveness of new anticoagulants is strongly dependent on INR control in the warfarin arm (chapter 3), meaning apixaban may well ultimately prove to be cost-effective only in situations where warfarin therapy is poorly controlled. Apixaban ceases to cost-effective against dabigatran (at a threshold of £30,000/QALY) once the price of apixaban exceeds £1,032 per annum.

Our results also appear to show genotype guided dosing to be cost-effective compared with clinical dosing. This pattern was consistently repeated across the range of simulations tested, and remained when trial data on INR values was used instead of simulated data. Whilst the duration of treatment benefit from genotyping is short, the prevention of events during this period means this group is at a permanently, albeit modestly, lower risk of future events. The congruence of the results between both trial and simulated data, both in terms of INR values and estimates of cost-effectiveness provides a measure of external validity to the simulated results we have generated. However, this trial still only provides the intermediary endpoint of time in INR ranges, not the full clinical endpoints we would ideally want to assess the accuracy of our simulations.

A number of other studies have attempted to assess the cost-effectiveness of genotype guided warfarin dosing versus clinical dosing. Patrick et al[186] used a Markov model approach and calculated that genotyping would have to increase time in range by 5-9% over the first three months for it to be cost-effective in the United States (our simulated increase was slightly over 5%). Eckman et al[187] calculated an ICER of US$170,000 per QALY, with only a 10% probability of cost-effectiveness at a threshold of $50,000 per QALY. Finally, Meckley et al[188], using a Markov model, calculated a 46% probability of cost-effectiveness at a threshold of US$50,000 per QALY. There is also one study, by You et al[189], which compares genotype guided warfarin with clinically dosed warfarin and dabigatran, with genotype dosing

dominating clinical dosing and dabigatran having an ICER of $13,810 per QALY against genotype dosed warfarin.

However, considerable uncertainty surrounded these analyses owing to the paucity of published data from randomized trials. Our approach, making use of PKPD simulations where data are unavailable, is able to help mitigate this problem and provide estimates with a lower (though still appreciable) amount of uncertainty. A number of larger trials of warfarin pharmacogenetics are now currently recruiting, designed to provide definitive evidence on efficacy[164,190,191]. While these trials will supersede the results of simulations, the modelled-based approach described here may still have a role in determining the optimal algorithm, given that it is implausible to trial each option.

Our analysis has a number of other strengths but also some weaknesses. The model makes use of a discrete event simulation methodology, which is likely to be a more appropriate method than a Markov model due to the higher number of states it allows patients to be classified into, plus its operation in continuous (as opposed to a Markov model's discrete) time, to allow for more realistic modelling of patient histories[27]. This is particular important in the example described as the benefit of genotype-guided warfarin is for only for a limited period of time (approximately 3 months). A Markov model would thus need either to have very short cycles (making parameterisation more difficult) or the additional benefit will accrue over only a very small number of cycles (potentially reducing accuracy).

Our adjusted, indirect comparison is necessary to include all possible treatment options. However, this may introduce bias through differences in trial design, which have been extensively discussed previously[149], a lack of access to individual patient data and the need to extrapolate the available data from trial to lifetime horizons. Nonetheless, these assumptions are virtually unavoidable in any economic evaluation[17], and there is no prospect of any evidence becoming available in the future that would enable us to avoid them.

There are also limitations surrounding the use of simulated data. First, the models used may be misspecified. The PKPD model[165] was derived from a number of data sources and the results obtained from it are difficult to validate externally. Second, both the PKPD and economic models are parameter intensive, increasing the probability that some of the values used are inaccurate. Thirdly, each of the three stages of the methodology (PKPD simulation, meta-analysis and economic model) introduces uncertainties. Whilst these may be quantified

at each individual stage of the process, they may not be synthesised into a single measure of uncertainty for the whole simulation.

While data from phase III trials, if available, should take precedence in evaluating both effectiveness and cost-effectiveness over studies based on simulations, there are situations when PKPD-based simulations can add value. These can be in providing: cost-effectiveness estimates of complex pharmaceutical interventions, early indications of cost-effectiveness before large scale trial data becomes available, and inform the extrapolation of the results of those trials once they have been conducted.

Chapter 6

Discussion

Summary

The approach developed links clinical trial simulations based on pharmacokinetic and pharmacodynamic models with economic modelling techniques. This enables estimates of cost-effectiveness to be derived in situations where sufficient trial data are not available to do so using standard economic models. Crucially, these models are mechanistic, meaning they can be used to extrapolate results to different patient populations and incorporate the effects of protocol deviations and different dosing regimens. The equations underlying these models are linked back to the underlying pharmacology of the system, an advantage in situations where there is a shortage of evidence as it means any necessary modelling assumptions will be grounded in pharmacological theory.

Chapters 2 and 5 both provide evidence of the conceptual feasibility of this approach. In both the PACIFICO model in chapter 2 and the warfarin model in chapter 5 it was possible to construct full prospective PKPDPE simulations of trials that have not yet taken place, and derive values for all the same endpoints as would be expected of a post-trial analysis. By synthesising published data (clinical, utility and cost) all necessary parameter values to populate the economic model can be obtained. However, it is important to remember that neither of these were trials of new drugs, rather of new uses or dosing regimens for existing ones.

Chapter 2 investigated the use of rituximab in a new indication, specifically as an adjuvant to different chemotherapy regimens for follicular lymphoma than those it had been previously used with, and as a maintenance therapy following chemotherapy. This means that prior data was already available on its use with different treatments from which extrapolations could be made. Indeed, the PK model was not from follicular lymphoma patients[60], but from patients with rheumatoid arthritis (there is no evidence PK parameters should differ between these conditions) and the PD model was derived from analyses using different chemotherapy regimens[62]. Likewise, warfarin therapy has been in use for many years, and the analyses in chapter 5 studied different dosing schedules for an already well-studied treatment. This limitation is always likely to be the case for any analysis relying on published data, meaning such analyses are likely to concern new uses of existing medications rather than entirely new treatments. The undertaking of analyses of new drugs will, in most cases, only be possible by the company developing that drug, as they will have access to the necessary phase I/II PKPD data, which will not come into the public domain until much later.

It was also necessary in both case studies (rituximab and warfarin) to produce early estimates of both costs and utilities associated with the new treatment/regimen. This is helped by the existence of databases of both unit cost data for the UK[105] and utility losses associated with various outcome events[101]. These pre-existing data can be used to attach outcome values to each of the clinical events occurring in the models, obviating the need to collect additional prospective data. However, there are difficulties with such an approach, first in that the price of a new medicine (or changes in price as it gains a new indication) are unlikely to become known (publically, though a company undertaking such an analysis will have this information) until after it has received regulatory approval. In chapter 5 it was necessary to make an assumption as to the price of apixaban as no data was available, and since this is likely to be one of the principle incremental costs of the treatment, any inaccuracy in this parameter can have a very large impact on the reliability of the results obtained. Second, whilst health outcomes and utilities are often assumed to be similar across countries (there will be differences based on patient populations and clinical practice differences, but these are hoped to be sufficiently small as to be ignorable), costs will clearly differ greatly, depending on the payment structure of the healthcare system[192]. This means that as soon as economic evidence is included in a decision it becomes necessary to consider multiple analyses for the different countries of interest. Overcoming both of these difficulties will again require information on pricing strategy and intentions only available internally within a drug company.

A number of attempts were made to look at assessing the validity of the modelling approach, by comparing to published data. In chapter 2, where the first two examples are retrospectively modelling previously published trials, these simulated results were directly compared with those from the trial. There were moderate differences in ICERs (£1,355 per QALY and £2,099 per QALY for the two examples) between simulated and trial ICERs, but both approaches led to the same decisions regarding cost-effectiveness. In general, this is evidence of the fairly logical point that simulations which show a treatment to be very clearly either cost-effective or not cost-effective are likely to be reliable, but those which lie close to the threshold will contain sufficient uncertainty as to be much less so. In these cases it becomes importantly to quantify that uncertainty through appropriate probabilistic analyses. In chapter 5 simulated results were compared to trials of genotype-guided warfarin therapy. Again, a difference in ICER between simulated and trial-based data (£2,280) was present but not substantial. In this case, the interest is not in whether the simulation can accurately predict trial-based cost-effectiveness results (as this cost-effectiveness is ultimately going to be

determined on the basis of on-going trials not simulations) but rather whether the method is able to appropriately rank differing algorithms in terms of cost-effectiveness as a method of optimisation.

In the long term, however, the validity of such a method will be judged on whether or not it is able to, across a range of different treatments and scenarios, produce sufficiently accurate predictions as to usefully inform decision making. It is this external validity that will ultimately prove more useful than any measures of internal validity that may be constructed. It will not be possible to address this question until the completion of trials for which data were simulated (e.g. the completion of the PACIFICO trial simulated in chapter 2, or the COAG[189] or EU-PACT[166] treaties for warfarin pharmacogenetics), and until such data becomes available it will not be possible to make a fully informed judgement as to the overall validity of the approach.

The method of PKPDPE modelling may have two areas of application. The first is the derivation of earlier estimates of cost-effectiveness than are currently available, an approach, as stated above, mostly of interest and use to pharmaceutical companies during the drug development process. Whether such an approach is widely adopted in the future will depend entirely on whether it can be shown to improve decision making and efficiency during that process, in the same way that standard clinical trial simulations were shown to do, with their use leading to an increase in the proportion of trials meeting their pre-specified objectives[36]. The second use is in helping to more accurately extrapolate the results of previously conducted trials. This may take the form of correcting for protocol deviations in a trial (warfarin non-adherence, chapter 5), assessing alternative dosing regimens (genotype-guided warfarin dosing, chapter 5) or the use of a drug in different patient populations or for different indications (rituximab, chapter 2). These uses are all based on the principle of making the maximum use of all available evidence in coming to a decision[26].

Strengths and limitations

The basic strength of the approach described is that it makes use of all available data. As discussed earlier, the four key features of a health economic decision model are data synthesis, explicit modelling of uncertainty, the inclusion of all relevant comparators and appropriate extrapolations. The discrete event simulation of warfarin (chapter 3) is designed for the extrapolation of trial results (real or simulated) to a lifetime horizon; the indirect comparison (chapter 4) brings in all relevant comparators and the PKPDPE simulation

(chapter 5) synthesises data from simulations and published trials together with cost and utility data to produce an ICER which is a function of all inputs. Finally, each stage of the model has a probabilistic element to model the parameter uncertainty present.

The approach is also designed to answer questions where no appropriate trial data will be available. Since data from trials will always supersede simulated data when it becomes available, it is important that simulated results should either refer to situations before such data are available or where they never will be so. The two applications of early cost-effectiveness indications and post-trial extrapolations of the data respectively relate to each of these appropriate areas for simulation. In both of these situations a lack of data makes the use of standard empirical economic analyses extremely difficult (the first through a simple lack of data, the second through the difficulties of extrapolation based on empirical models), a problem that the mechanistic nature of PKPD models can help to overcome.

However, there are a number of limitations associated with this approach. First, is the risk of model misspecification. The PKPD models were extrapolated to slightly different contexts and patient populations than those in which they were built, and it is necessary to assume (without any empirical evidence) that the same exposure response relationships will apply. In this, the situation is much the same as with traditional clinical trial simulations, except as the results of these simulations are then used as inputs to other models, any errors may be magnified over the process, potentially resulting in a greater overall error rate. This limitation is unavoidable in any early stage extrapolation where there is a paucity of data, and the only way to minimise it is to use only well verified (preferably externally) PKPD models, and ensure all uncertainty is explicitly modelled. Whereas in traditional CTS it may be desirable to only include inter-patient variability and ignore parameter uncertainty, this would be incorrect here as the desire is to maintain that uncertainty through to the final result.

Similar concerns also apply to the health economic extrapolations used. Due to the holding of randomised controlled trials as the gold standard for evidence of efficacy, there has been concern about the validity of various model extrapolations. In particular, as one of the key features of a well conducted trial is blinding, to ensure the prejudices of both patients and investigators do not bias the results, there is a fear bias will be reintroduced through the choices made in the modelling process. It is certainly true that different structural assumptions (which it is often difficult to make an objective comparison between) can produce considerable variability in output. The approach described in this thesis may

exacerbate this fear as conducting analyses at an earlier stage means less empirical data will be available, thus forcing more assumptions in the approach. Nevertheless, modelling is accepted as an unavoidable part of economic evaluation as the biases caused by relying purely on trial results (short time horizons and not representative of standard practice) are worse than those introduced by modelling[22]. The quantitative nature of the process means all assumptions can be explicitly stated and evaluated, so any biases should at least be transparent rather than hidden.

It is necessary to synthesise data from a range of studies to perform these analyses, often from trials with very different designs. As an example, the RELY[92], ROCKET-AF[134] and ARISTOTLE[135] studies used in the indirect comparison in chapter 4 are all large trials of new drugs for stroke prophylaxis in atrial fibrillation. Nevertheless, there is considerable variability in their designs[149]. Examples include different patient populations (baseline stoke risk much higher in ROCKET-AF), different blinding (ROCKET-AF and ARISTOTLE used sham INR, RELY was open-labelled), different definitions of outcome events, different qualities of warfarin INR control and many others. These problems exist in any modelling exercise but may be exacerbated by the approach adopted here given the need to synthesise a greater number of studies (PKPD as well as trials). However, despite this problem, evidence-based medicine should still involve the use of the maximum amount of data possible[26], and PKPDPE as an approach still adheres to that principle.

The models constructed are also extremely difficult to validate. Whilst this may be possible for the individual components of the model, by testing PKPD simulations against available reference data and comparing health economic models to available trial data, because it is not possible to fully combine the uncertainty from the two different models there is no way to be certain how given inaccuracies at each stage affect the accuracy of the overall result. External validation of an individual model against other data sets is unlikely to be possible as the very nature of the use of these models in situations where no trial has been conducted means such data sets are unlikely to be available.

This problem is not likely to be resolved until sufficient trials which have been simulated using this method report trial results, and these can be compared to the output of the simulations, a process that will only be delayed by the fact most such analyses are likely to be conducted by private industry, and these may well never be published. Even when such comparisons do become available, the likely result that some simulations match trial data

very well and others are extremely inaccurate will leave the difficult job of identifying exactly which features of a model are predictive of how accurate the simulations are likely to be.

There is also the difficulty of the lack of uniformity in healthcare reimbursement structures around the world. Cost-effectiveness evidence requirements will differ considerably between countries with a central reimbursement body (e.g. NICE in England & Wales, SMC in Scotland or AWMSG in Wales) and ones with a higher predominance of private healthcare provision, and the evidence necessary may well differ even where countries systems are similarly structured. This is in contrast to licensing submissions where, whilst there may be differences between bodies (e.g. FDA, EMA), they are likely to be considerably smaller. Thus, whilst a single CTS may be useful to optimise clinical trial designs for all licensing submissions, a single PKPDPE analysis is much more restricted to the specific country for which it was designed. This leaves either suboptimal evidence for other countries or the need to conduct multiple analyses, obviously increasing both the time and cost.

This leads naturally to one of the major difficulties with the applicability of this approach. The very structured nature of the licensing approval process, with its specific evidential hurdles that have to be crossed at each stage, is very amenable to the use of simulations to predict the outcome at each of these points. The lack of such a uniformly structured process in reimbursement decisions makes the application of the technique more challenging.

Finally, the cross-disciplinary nature of the approach naturally makes its implementation more complex, as a larger number of people from different areas will need to work together to perform the work efficiently. This need for cross co-ordination (e.g. between pharmacological modellers, trial statisticians, trial design team) slowed the process of CTS coming into common usage[36], and the addition of economic evidence as well can only make the process more challenging. This fact does not reduce the conceptual utility of the approach; merely reduce the potential speed at which it may be adopted.

Future work

In addition to the various limitations listed above, there are a number of technical issues with the approach which further work would be necessary to address. First, the accuracy of the predictions generated is clearly constrained by the quality of the PKPD models used as the basis for simulations. These could be developed in one of two ways. The first would be

finding the best possible model fit for the data, ignoring any attempt to structure the model in a way where parameters have an intuitive real world interpretation. The closer the model fit, the more accurate the CTS would be expected to be and thus the more accurate the cost-effectiveness estimates should be.

The second approach is for model structure to be informed by the underlying biology of the system. This links to the concept of physiologically based PKPD models discussed earlier[43]. Here it is necessary to accept, at least in the short term, a less good model fit but it is hoped this is compensated for by a greater ability to interpret the results. For example, if a change in patient population or model parameters leads to a change in cost-effectiveness output, it would be hoped to be able to use the structure of the model to hypothesise the biological reason for this change. Which of these two approaches is preferred depends on the specific question being addressed. If the aim is simply the best possible point estimates of cost-effectiveness in a given situation, the best fitting model would be preferred. If, on the other hand, the interest is in comparative cost-effectiveness between differing scenarios, the second approach may be preferred so as to improve the interpretability of the results. PBPKPD models are still at a relatively early stage of development and acceptance into mainstream analyses, so hopefully future developments should lead to more becoming available for simulation testing.

One weakness of the specific modelling approach used, particularly as one of the key goals is the quantitative modelling of uncertainty, is that it is now possible to carry that full uncertainty through the simulation at an individual level. Whilst both inter-individual heterogeneity and parameter uncertainty can be quantified in the CTS, it is necessary to then summarise the output of these simulations as parametric distributions for use in the probabilistic sensitivity analysis of the economic models. Whilst maintaining the same overall level of uncertainty, information is lost by needing to summarise the output in this way. Specifically, it is not possible to maintain a direct connection between the individual patients in the PKPD simulation of clinical events and the individual patients in the economic model. In this way the simulation is not accurately replicating the trial data it is attempting to replace, as here it is possible to pair up clinical data with both utilities and costs.

Ideally, the whole process would be conducted in a single stage, rather than the two separate stages currently necessary. In the same way that within a health economic evaluation an individual patient simulation can provide more information than a cohort model, due to better

modelling of uncertainty, a fully integrated individual level simulation would be better than serial models. The limitation of doing this at present is essentially computational, as the standard programs in use for PKPD modelling (e.g. NONMEM[78]) are not amenable to economic modelling, and vice versa. The development of appropriate interfaces to fully integrate the two should allowed to produce, in a single simulation, a patient's full simulated clinical history, together with the cost and utility data attached to that.

There is also the potential for the integration of the techniques developed here with value of information (VOI) analysis[47]. This is another tool for the optimisation of trial designs, based on calculating the value of a trial by the benefits associated with a reduction in uncertainty, and comparing this to the cost of conducting the trial. Chapter 2 already shows the compatibility of such an approach with PKPDPE modelling, and in fact the necessary data for VOI is exactly that data PKPDPE simulations produce. The current difficulty with this process is simply one of computational complexity. To conduct a full expected value of perfect parameter information (EVPPI) or expected value of sample information (EVSI) calculation, the level necessary to use in trial optimisation, is likely to require at least a thousand replications of the full PKPDPE analysis for each trial in the analysis, which is already a computationally intensive process due to the numbers of patients simulated per trial. However, improvements in hardware or the development of statistical short cuts in the process should render such analysis practical to perform.

VOI analysis is particularly useful with the spreading implementation of value based pricing[49]. Under this system, rather than manufacturers specifying a fixed price for the drug, which is then either accepted or rejected on the grounds of cost-effectiveness, a drug will be assigned a price so as to set its cost-effectiveness at a pre-specified acceptable level. Whilst it is debated whether there will be any benefit in switching to such a system[49], there is no doubt it will change the nature of the decision making process. Specifically, rather than the importance lying in reducing decision uncertainty in whether a new drug will be accepted or rejected for reimbursement, it is now necessary to find the most accurate possible point estimate of cost-effectiveness, as this will affect the finalised price. In such circumstances, VOI analysis becomes an extremely powerful tool, as it enables explicit comparisons between trial costs and the benefits of reduced uncertainty in that point estimate. During the drug development process, PKPDPE simulations become a method of deriving early estimations of this value.

Model-based drug development

One of the overarching frameworks within which the work presented here fits is that of model-based drug development (MBDD). This has been defined by the FDA in its critical path document as the development and application of pharmacostatistical models of drug efficacy and safety from clinical and preclinical data to improve drug development decision making[44]. PKPD modelling, the differentiation of the "learning" and "confirming" phases of drug development and the use of clinical trial simulations are all aspects of MBDD.

The underlying rationale is that drug development is a continuous process where additional data is always being acquired[193]. The aim is for data from each new clinical phase to be synthesised with relevant prior information to inform decision making, rather than prior information being neglected which is commonly the place at present[193]. This naturally leads towards a Bayesian approach where prior beliefs and newly collected data can be combined into posterior probabilities. Models allow for explicit quantitative evidence synthesis, and enable the drug development process to be regarded as a model building exercise where models are continuously updated as new information becomes available.

There are six generally accepted aspects of MBDD (figure 15, P155). The first is PKPD and disease models. Disease (or placebo) models chart the expected time course of patient response over time, given no treatment (or placebo)[194]. Features can include natural disease progression over time, cyclical variations and regression to the mean. This last can be particularly relevant in clinical trials where study inclusion criteria may mean patients are included because they are at a particularly severe point at that time, meaning some natural improvement would be expected. PKPD models can be combined with disease progression models to simulate the expected time course of the condition on a given treatment[195].

The second aspect is the meta-analysis of candidate and comparator data. Since the commercial value of a drug depends on the other drugs available in that indication, combining all available data is extremely important. Model-based meta-analyses of dose response and time course data will help to understand the comparative pharmacodynamics of the treatment alternatives, in a way that standard meta-analyses cannot[196,197]. The fact the models used are mechanistic again means they have a direct link to the underlying biological properties of the drug in a way empirical models do not. This is particularly important at these early stages where one of the major goals is to increase the understanding not just of what effects a drug has, but why those effects occur. Technical issues it is necessary to

address include addressing publication bias, addressing the issue of correlation between time points in a longitudinal analysis and synthesising study level and patient level data from different sources[196].

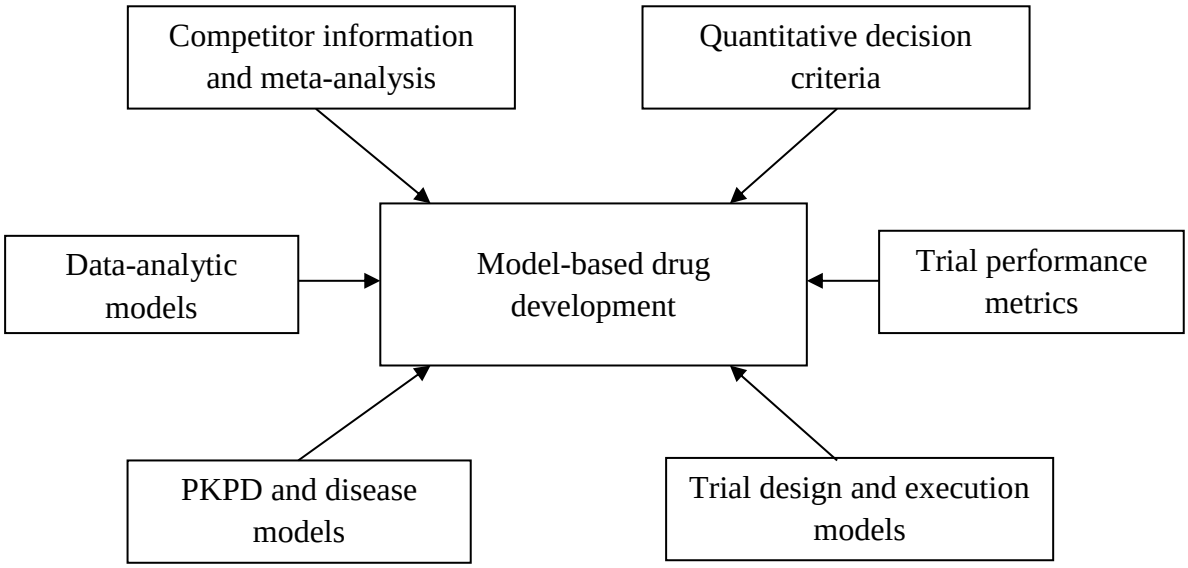


Figure 15

The remaining aspects are trial design and execution models, specifying the trial protocol and expected deviations from that; data-analytic models defining the statistical analysis that will be undertaken following collection of the trial data; and quantitative decision criteria and trial performance metrics to specify decisions made during the trial process (e.g. stop/go decisions) and the specific criteria that would lead to each decision being made at a given time[193]. Putting all these together gives a full quantitative approach to trial design and analysis, with explicitly defined decision rules at each stage of the process, meaning that decisions are evidence based and the reasoning behind those decisions transparent.

The use of MBDD by companies is becoming increasingly common[198]. It is hoped that the integrated approach it provides, together with the necessary integration across different areas of the company, will be able to produce a greater volume of information at a lower cost[199]. One particular area of potential development is improvements in adaptive trial designs. An adaptive, as opposed to a fixed design, is one where rather than a pre-defined protocol that is followed in all circumstances; different paths will be followed depending on the results obtained, though clearly all such possibilities still have to be defined in advance. Examples include interim analyses with defined rules for stopping for failure, stopping for success and

collection more data, or seamless phase II phase III trials, where the trial automatically switch between the two once the necessary information has been obtained (e.g. optimal dose). MBDD can further improve such designs as it gives a larger range of outputs (e.g. pharmacological) which can be used in designing trials. It also produces an explicit framework in which trial design can be structured, with the aim of reaching necessary levels of evidence to gain regulatory approval as the lowest possible cost (financially, patient numbers, time etc).

The principles of MBDD are fully compatible with the modelling framework developed in this thesis. There are already PKPD and trial execution models in the CTS used, and an economic model is a data-analytic method designed to provide quantitative answers to questions that, if well executed, should contain a meta-analysis of all available comparators. Developments across any of the individual fields of MBDD should thus also help to improve the accuracy of the analyses.

The PKPDPE approach described in this thesis suggests the potential integration of a seventh component, economic modelling, within MBDD. At present, such analyses are normally concerned purely with clinical endpoints, and the synthesis of all available efficacy and/or effectiveness evidence[193]. In a framework which is already both data and parameter intensive, and where the decision has already been made that making use of all available evidence is a more important consideration than parsimony, the additional assumptions necessary should not be a major concern. Since the underlying theoretical idea of MBDD is to enable as much evidence as possible to be used to inform decision making, adding an economic component as well should only improve the usefulness of such an approach.

Conclusion

We have demonstrated the feasibility of an integrated pharmacological and economic modelling approach, which can be used both to improve the extrapolation of trial results and obtain earlier indications of cost-effectiveness than are currently available. However, there are difficulties with the practical implementation of this approach, specifically surrounding the validity of extrapolating based on the limited data we will have available. Whether or not such simulations become a regular component of the drug development process depends on the essentially empirical question of whether they improve decision making and ultimately trial success rates within that process. This is not a question that will be possible to answer

until more such analyses have been undertaken and we have the necessary data to fully compare trial and simulated results.

Figure Legends

Fig 1 – Graphical representation of 1, 2 and 3 compartment pharmacokinetic models.

Fig 2 – Pharmacokinetic and pharmacodynamic models combined to give a PKPD model of drug effect over time.

Fig 3 - Markov models used for the health economic analyses: (a) represents the three-state model used in the maintenance and first-line analyses and (b) represents the five-state model used for the PACIFICO simulation. FL = follicular lymphoma.

Fig 4 - Results of the probabilistic sensitivity analysis. Upper panels relate to maintenance rituximab therapy; lower panels relate to first-line use. Broken curves are trial results and solid curves are simulated results. Panels on the left-hand side are cost-effectiveness acceptability curves illustrating the probability of the rituximab-containing regimen being cost effective at different cost-effectiveness thresholds. Panels on the right-hand side illustrate the congruence between simulation and trial-based evaluations, given as the proportions of simulations where simulated and trial results fall on different sides of given cost-effectiveness thresholds.

Fig 5 - Results of the simulation of the PACIFICO trial. Presented are the proportions of patients in each health state according to time for patients receiving (a) rituximab, cyclophosphamide, vincristine and prednisolone (R-CVP) and (b) rituximab, fludarabine and cyclophosphamide (R-FC). The red line represents patients in the progression-free survival on-protocol (maintenance or induction) state; the green line represents progression-free survival off protocol; the blue line denotes progressed follicular lymphoma; and the black line denotes patients who have died. The bottom graph (c) is the cost-effectiveness acceptability curve giving the probability of R-FC being cost effective (compared with R-CVP) at different cost-effectiveness thresholds.

Fig 6 - Tornado plot of univariate sensitivity analyses. First three panels relate to benefit-harm analyses; lower right panel relates to economic comparison of dabigatran 150 mg twice daily and dose adjusted warfarin. L=lower end of 95% CI for parameter set; H=higher end of 95% CI for parameter set (see table 7). bid=twice daily; $CHADS_2$= Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, previous Stroke/transient ischaemic attack; ICER=incremental cost-effectiveness ratio; QALY=quality-adjusted life-year. *Maximum

deviation from all correlation structures tested, which occurred when all patients with hypertension were assumed to have diabetes and all patients with previous myocardial infarction were assumed to also have previous stroke

Fig 7 – Cost-effectiveness acceptability curve for base case analysis. QALY=quality-adjusted life year.

Fig 8 - Results of probabilistic sensitivity analysis on efficacy and safety end points, expressed as incremental QALYs. Values are means and 95% central ranges from 2000 simulations. $CHADS_2$= Congestive heart failure, Hypertension, Age≥75 years, Diabetes mellitus, previous Stroke/transient ischaemic attack; CrCl=creatinine clearance; QALY=quality-adjusted life-year; RE-LY= Randomized Evaluation of Long-Term Anticoagulation Therapy; TIA=transient ischaemic attack; TTR=time in therapeutic range.

Fig 9 - Results of the probabilistic parameter sensitivity analysis indicating the relative impact of treatments on the two constructs of the QALY – quality of life, and life expectancy. Point estimates and associated 50% and 95% central range ellipses for life years and quality of life with warfarin (black), rivaroxaban (red), dabigatran (green) and apixaban (blue).

Fig 10 - Tornado plots of univariate sensitivity analysis. Each figure presents the ten parameters which led to the greatest change in overall QALYs. L/H refers to lower and higher limits of parameter estimates.

Fig 11 - Probabilities of treatment being most effective by patient subgroup. Probability of each treatment is the most effective, based on accrual of most lifetime QALYs, for each identified patient subgroup.

Fig 12 - Results of base case simulation - times spent below, in and above INR range at 1 week, 2 weeks, 1 month and 3 months.

Fig 13 - Cost-effectiveness acceptability curves for base case analysis, showing the proportion of simulations where each treatment is the most cost-effective over different willingness to pay thresholds.

Fig 14 - Scenario analyses – shows probability each treatment accrues the largest number of QALYs in each patient subgroup.

Fig 15 - The components of a model-based drug development analysis.

References

[1] Williams AH. Health economics: the cheerful face of the dismal science? Health and economics: proceedings of section F (economics) of the British Association for the Advancement of Science, Bristol, Houndmills, Macmillan, (1987).

[2] Arrow K. Uncertainty and the welfare economics of medical care. American Economic Review, 53(5):941-73 (1963).

[3] Morris S, Devlin N, Parkin D. Economic analysis in health care. Philadelphia, PA, John Wiley & Sons, Inc (2007).

[4] Barnett K, Mercer SW, Norbury M, Watt G, Wykek S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross sectional study. Lancet, 380(9836):37-43 (2012).

[5] Grossman M. On the concept of health capital and the demand for health. Journal of Political Economy, 80(2):223-25 (1972).

[6] Ng YK. Welfare economics: introduction and development of basic concepts. London, Macmillan (1983).

[7] Sugden R, Williams AH. The principles of practical cost-benefit analysis. Oxford, Oxford University Press (1979).

[8] Culver AJ. The normative economics of health care finance and provision. Oxford Review of Economic Policy, 5:34-8 (1989).

[9] Pritchard C. Trend in economic evaluation. OHE Briefing 36, London, Office of Health Economics (1998).

[10] Sheiner L. Learning versus conforming in clinical drug development. Clin Pharmacol Ther, 61(3):275-91 (1997).

[11] Rosenberger WF, Lachin JM. Randomization in clinical trials: theory and practice. Philadelphia, PA, John Wiley & Sons, Inc (2002).

[12] Splanwinski J, Kunniar J. Clinical trials: active control vs placebo - what is ethical? Sci Eng Ethics, 10(1):73-9 (2004).

[13] Snapinn S. Noninferiority trials. Curr Control Trials Cardiovasc Med, 1(1):19-21 (2000).

[14] Palmer S, Torgerson DJ. Economic notes: definitions of efficiency. BMJ, 318(7191):1136 (1999).

[15] Torrance GW. Measurement of health state utilities for economic appraisal: a review. Journal of Health Economics, 5(1):1-30 (1986).

[16] Williams A. Economics of coronary artery bypass grafting. BMJ, 291:326-9 (1985).

[17] Drummond MF, Sculpher MJ, Torrance GW, O'Brien B, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford, Oxford University Press (2005).

[18] EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. Health Policy, 16(3):199-208 (1990).

[19] Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. Journal Health Economics, 18(3):341-64 (1999).

[20] Claxton K, Sculpher M, McCabe C, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Economics, 14(4):339-47 (2005).

[21] Raiffa H. Decision analysis: introductory lectures on choices under uncertainty. Reading, MA, Addison-Wesley (1968).

[22] Buxton MJ, Drummond MF, Van Hout BA, et al. Modelling in economic evaluation: an unavoidable fact of life. Health Economics, 6(3):217-27 (1997).

[23] Weinstein MC, Fineberg HV. Clinical decision analysis. Philadelphia, WB Saunders Co, 228-65 (1980).

[24] Hjelmgren J, Berggren F, Andersson F. Health economic guidelines - similarities, differences and some implications. Value in Health, 4(3):225-50 (2001).

[25] National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London, NICE (2004).

[26] Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence-based medicine: what it is and what it isn't. BMJ, 312(7023):71-2 (1996).

[27] Caro JJ, Möller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluations? Value Health, 13(8):1056-60 (2010).

[28] van Gestel A, Severens JL, Webers CA, Beckers HJ, Jansonius NM, Schouten JS. Modelling complex treatment strategies: construction and validation of a discrete event simulation model for glaucoma. Value Health, 13(4):358-67 (2010).

[29] Hughes DA, Cowell W, Koncz T, Cramer J. Methods for integrating medication compliance and persistence in pharmacoeconomic evaluations. Value Health, 10(6):498-509 (2007).

[30] Sorensen SV, Kansal AR, Connolly S, et al. Cost-effectiveness of dabigatran etexilate for the prevention of stroke and systemic embolism in atrial fibrillation: a Canadian payer perspective. Thromb Haemost, 105(5):908-19 (2011).

[31] Bombardier C, Maetzel A. Pharmacoeconomic evaluation of new treatments: efficacy versus effectiveness studies? Ann Rheum Dis, 58:82-5 (1999).

[32] Clemens K, Garrison LP, Jones A, Macdonald F. Strategic use of pharmacoeconomic research in early drug development and global pricing. Pharmacoeconomics, 4(5):315-22 (1993).

[33] Arrowsmith J. Trial watch: phase III and submission failures: 2007-2010. Nat Rev Drug Discov, 10(2):87 (2011).

[34] Ananthakrishan R, Gona P. Pharmacological modelling and biostatistical analysis of a new drug. Journal of clinical trials, 2:59-92 (2010).

[35] Hariharan S, Madabushi R. Clinical pharmacology basis of deriving dosing recommendations for dabigatran in patients with severe renal impairment. J Clin Pharmacol, 52(1):119-25 (2012).

[36] Holford N, Ma SC, Ploeger BA. Clinical trial simulations: a review. Clin Pharmacol Ther, 88(2):166-82 (2010).

[37] Holford NHG, Hale M, Ko HC, et al. Simulation in drug development: good practices. http://www.dml.georgetown.edu/edds/SDDGP.html (1999).

[38] Girard P. Clinical trial simulations: a tool for understanding study failures and preventing them. Basic Clin Pharmacol Toxicol, 96(3):228-34 (2005).

[39] Holford NHG, Kimko HC, Monteleone JPR, Peck CC. Simulation of clinical trials. Annu Rev Pharmacol Toxicol, 40:209-34 (2000).

[40] Bonate PL. Clinical trial simulation in drug development. Pharmaceutical research, 17(3):252-6 (2000).

[41] Gauthier M. Clinical trial simulation. Applied Clin Trials, 22-5 (1997).

[42] Brindley PG, Dunn WF. Simulation for clinical research trials: a theoretical outline. J Crit Care, 24(2):164-167 (2009).

[43] Yang RSH, Andersen ME, Dennison JE, Ou YC, Liao KH, Reisfeld B. Physiologically based pharmacokinetic and pharmacodynamic modelling. New York, NY, Ed. Wiley Inc (2004).

[44] US Food and Drug Administration. Innovation/stagnation: challenges and opportunities on the critical path to new medical products. Rockville (MD): US Department of Health and Human Services, FDA (2004).

[45] Poland B, Wada R. Combining drug-disease and economic modelling to inform drug development decisions. Drug Discovery Today, 6(22):1165-70 (2001).

[46] Hughes DA, Walley T. Economic evaluations during early (phase II) drug development: a role for clinical trial simulations? Pharmacoeconomics, 19(11):1069-77 (2001).

[47] Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. Pharmacoeconomics, 24(11):1055-68 (2006).

[48] Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? Health Economics, 15(7):677-87 (2006).

[49] Hughes DA. Value-based pricing: incentive for innovation or zero net benefit? Pharmacoeconomics, 29 (9):731-5 (2011).

[50] Miller P. Role of pharmacoeconomic evaluation in R&D decision making: when, where, how? Pharmacoeconomics, 23 (1):1-12 (2005).

[51] Hartz S, John J. Contribution of economic evaluation to decision making in early phases of product development: a methodological and empirical review. Int J Technol Assess Health Care, 24:465-72 (2008).

[52] Gabrielsson J, Weiner D. Pharmacokinetics and pharmacodynamic data analysis: concepts and applications. 4th ed. Stockholm: Swedish Pharmaceutical Press (2002).

[53] Mager DE, Wyska E, Jusko WJ. Diversity of mechanism-based pharmacodynamic models. Drug Metab Dispos, 31:501-8 (2003).

[54] Jusko WJ, Ko HC. Physiologic indirect response models characterize diverse types of pharmacodynamic effects. Clin Pharmacol Ther, 56:406-19 (1994).

[55] Chan PL, Holford NH. Drug treatment effects on disease progression. Annu Rev Pharmacol Toxicol, 41:625-59 (2001).

[56] Gobburu JV, Lesko LJ. Quantitative disease, drug, and trial models. Annu Rev Pharmacol Toxicol, 49:291-301 (2009).

[57] Roche. Rituximab for the treatment of relapsed follicular lymphoma: single technology appraisal submission to the National Institute for Health and Clinical Excellence. http://www.nice.org.uk/nicemedia/live/11730/38897/38897.pdf (2007).

[58] Liverpool Reviews and Implementation Group. ERG report: rituximab for the first line treatment of stage III-IV follicular non-Hodgkin's lymphoma. http://www.hta.ac.uk/erg/reports/1636.pdf (2006).

[59] International Standard Randomised Controlled Trial Number Register. Purine-alkylator combination in follicular lymphoma immuno-chemotherapy for older patients. http://www.controlled-trials.com/ISRCTN99217456/pacifico (2004).

[60] Ng CM, Bruno R, Combs D, et al. Population pharmacokinetics of rituximab (anti-CD20 monoclonal antibody) in rheumatoid arthritis patients during a phase II clinical trial. J Clin Pharmacol, 45:792-801 (2005).

[61] Sacco JJ, Botten J, Macbeth F, et al. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. PLoS One, 5(1):e8933 (2010).

[62] Ternant D, Hénin E, Cartron G, et al. Development of a drug-disease simulation model for rituximab in follicular non-Hodgkin lymphoma. Br J Clin Pharmacol, 68:561-73 (2009).

[63] van Oers MHJ, Klasa R, Marcus RE, et al. Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial. Blood, 108:295-301 (2006).

[64] Marcus R, Imrie K, Belch A, et al. CVP chemotherapy plus rituximab compared with CVP as first-line treatment for advanced follicular lymphoma. Blood, 105:1417-23 (2005).

[65] Hochster HS, Weller E, Gascoyne RD, et al. Cyclophosphamide and fludarabine (CF) in advanced indolent lymphoma: results from the ECOG/CALGB intergroup E1496 trial [abstract]. J Clin Oncol, Jun 20;25 (18 Suppl.):8004 (2007).

[66] Engauge Digitizer V4.1. http://digitizer.sourceforge.net (2009).

[67] Ray JA, Carr E, Lewis G, et al. An evaluation of the cost-effectiveness of rituximab in combination with chemotherapy for the first-line treatment of follicular non-Hodgkin's lymphoma in the UK. Value Health, 13(4):346-57 (2010).

[68] Joint Formulary Committee of the British Medical Association and the Royal Pharmaceutical Society. British national formulary. 47th ed. London: BMJ Group, Royal Pharmaceutical Society (2004).

[69] Wilmoth JR, Shkolnikov V. The human mortality database. http://www.mortality.org (2006).

[70] Tam CS, Wolf M, Prince HM, et al. Fludarabine, cyclophosphamide, and rituximab for the treatment of patients with chronic lymphocytic leukemia or indolent non-Hodgkin lymphoma. Cancer, 106:2414-20 (2006).

[71] Sacch S, Pozzi S, Marcheselli R, et al. Rituximab in combination with fludarabine and cyclophosphamide in the treatment of patients with recurrent follicular lymphoma. Cancer, 110:121-8 (2007).

[72] Tam CS, Wolf MM, Januszewicz EH, et al. Fludarabine and cyclophosphamide using an attenuated dose schedule is a highly effective regimen for patients with indolent lymphoid malignancies. Cancer, 100:2181-9 (2004).

[73] Hiddemann W, Kneba M, Dreyling M, et al. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. Blood, 106:3725-32 (2005).

[74] Liou SY, Stephens JM, Carpiuc KT, et al. Economic burden of haematological adverse effects in cancer patients: a systematic review. Clin Drug Investig, 27:381-96 (2007).

[75] Joint Formulary Committee of the British Medical Association and the Royal Pharmaceutical Society. British national formulary. 59th ed. London: BMJ Group, Royal Pharmaceutical Society (2010).

[76] Chilcott J, Brennan A, Booth A, et al. The role of modelling in prioritising and planning clinical trials. Health Technol Assess, 7:1-125 (2003).

[77] Office for National Statistics. Cancer incidence and mortality in the United Kingdom, 2004-2006. http://www.ons.gov.uk/ons/rel/cancer-unit/cancer-incidence-and-mortality/united-kingdom-and-constituent-countries--2004-06/index.html (2010).

[78] Beal S, Sheiner LB, Boeckmann A, et al. NONMEM user's guides (1989-2009). Ellicott City, (MD): Icon Development Solutions (2009).

[79] R Development Core Team. R: a language and environment for statistical computing. Version 2.9.2. Vienna: R Foundation for Statistical Computing (2009).

[80] Deconinck E, Miadi-Fargier H, Le Pen C, et al. Cost effectiveness of rituximab maintenance therapy in follicular lymphoma: long-term economic evaluation. Pharmacoeconomics, 28(1):35-46 (2010).

[81] Grupo de Farmacoeconomía del Linfoma Folicular. Rituximab cost analysis for maintenance treatment of patients with follicular lymphoma [in Spanish]. Farm Hosp, 32(1):25-34 (2008).

[82] Groot MT, Lugtenburg PJ, Hornberger J, et al. Cost effectiveness of rituximab (MabThera) in diffuse large B-cell lymphoma in the Netherlands. Eur J Haematol, 74:194-202 (2005).

[83] Hayslip JW, Simpson KN. Cost-effectiveness of extended adjuvant rituximab for US patients aged 65-70 years with follicular lymphoma in second remission. Clinical Lymphoma Myeloma, 8:166-70 (2008).

[84] Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. Health Econ, 5:513-24 (1996).

[85] Hughes DA. From NCE to NICE: the role of pharmacoeconomics. Br J Clin Pharmacol, 70(3):317-9 (2010).

[86] Hughes DA, Walley T. Predicting "real world" effectiveness by integrating adherence with pharmacodynamic modelling. Clin Pharmacol Ther, 74:1-8 (2003).

[87] Lip GYH, Tse, HF. Management of atrial fibrillation. Lancet, 370:604-18 (2007).

[88] Hobbs FD, Fitzmaurice DA, Mant J, et al. A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over. The SAFE study. Health Technol Assess, 9:1-74 (2005).

[89] The Office of Health Economics. Estimating the direct costs of atrial fibrillation to the NHS in the constituent countries of the UK and at SHA level in England. London (UK): Office of Health Economics (2009).

[90] The National Collaborating Centre for Chronic Conditions. Atrial Fibrillation: National clinical guideline for management in primary and secondary care. London (UK): Royal College of Physicians (2006).

[91] Connolly SJ, Pogue J, Eikelboom J, et al. Benefit of oral anticoagulation over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. Circulation, 118:1029-37 (2008).

[92] Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med,361:1139-51 (2009).

[93] Connolly SJ, Ezekowitz MD, Yusuf S, et al. Newly identified events in the RE-LY trial. N Engl J Med, 363(19):1875-6 (2010).

[94] Beasley BN, Unger EF, Temple R. Anticoagulant options - why the FDA approved a higher but not a lower dose of dabigatran. N Engl J Med, 364(19):1788-90 (2011).

[95] European Medicines Agency. Pradaxa product information. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000829/WC500041059.pdf (2011).

[96] Hughes DA, Bayoumi AM, Pirmohamed M. Current assessment of risk-benefit by regulators: is it time to introduce decision analyses? Clin Pharmacol Ther, 82(2):127-30 (2007).

[97] Garrison LP Jr, Towse A, Bresnahan BW. Assessing a structured, quantitative health outcomes approach to drug risk-benefit analysis. Health Aff (Millwood), 26(3):684-95 (2007).

[98] Ezekowitz MD, Connolly S, Parekh A, et al. Rationale and design of RE-LY: randomised evaluation of long-term anticoagulant therapy, warfarin, compared with dabigatran. Am Heart J, 157(5):805-10 (2009).

[99] Ezekowitz MD, Reilly PA, Nehmiz G, et al. Dabigatran with or without concomitant aspirin compared with warfarin alone in patients with nonvalvular atrial fibrillation (PETRO study). Am J Cardiol, 100:1419-26 (2007).

[100] Boehringer Ingelheim. Advisory committee briefing document – dabigatran etexilate. Silver Spring, MD: Food and Drug Administration http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/cardiovascularandrenaldrugsadvisorycommittee/ucm226009.pdf (2010).

[101] Sullivan PW, Ghushchyan V. Preference-based EQ-5D index scores for chronic conditions in the United States. Med Decis Making, 26(4):410-20 (2006).

[102] Gage BF, Cardinalli AB, Owens DK. The effect of stroke and stroke prophylaxis with aspirin or warfarin on quality of life. Arch Intern Med, 156:1829-36 (1996).

[103] Meckley LM, Gudgeon JM, Anderson JL, Williams MS, Veenstra DL. A policy model to evaluate the benefits, risks and costs of warfarin pharmacogenomic testing. Pharmacoeconomics, 28(1):61-74 (2010).

[104] Hemingway H. The effectiveness and cost-effectiveness of biomarkers for the prioritisation of patients awaiting coronary revascularisation: a systematic review and decision model. Health Technol Assess, 14(9):1-151 (2010).

[105] Department of Health. NHS reference costs 2008/09. London, DH http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_111591 (2009).

[106] National Institute of Health and Clinical Excellence. Atrial fibrillation: the management of atrial fibrillation. National costing report. London, NICE http://www.nice.org.uk/nicemedia/pdf/CG036costingreport.pdf (2006).

[107] Leontiadis GI, Sreedharan A, Dorward S, et al. Systematic reviews of the clinical effectiveness and cost-effectiveness of proton pump inhibitors in acute upper gastrointestinal bleeding. Health Technol Assess, 11(51):1-164 (2007).

[108] Jorgensen AL, Al-Zubiedi S, Zhang JE, et al. Genetic and environmental factors determining clinical outcomes and cost of warfarin therapy: a prospective study. Pharmacogenet Genomics, 19(10):800-12 (2009).

[109] National Institute for Health and Clinical Excellence. Atrial fibrillation - dabigatran etexilate: appraisal consultation document. London, NICE http://guidance.nice.org.uk/TA/Wave21/10/Consultation/DraftGuidance (2011).

[110] Eastern Region Public Health Observatory. Modelled estimates and projections of hypertension for local authorities in England. Cambridge: Eastern Region Public Health Observatory. http://www.erpho.org.uk/viewResource.aspx?id=17904 (2009).

[111] Quality and Outcomes Framework. Diabetes in the UK 2010. Diabetes UK. http://www.diabetes.org.uk/Professionals/Publications-reports-and-resources/Reports-statistics-and-case-studies/Reports/Diabetes-prevalence-2010/ (2010).

[112] Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Annals of Internal Medicine, 146(12):857-67 (2007).

[113] Cleland JGF, Findlay I, Jafri S. et al. The Warfarin/Aspirin Study in Heart Failure (WASH): a randomized trial comparing antithrombotic strategies for patients with heart failure. Am Heart J, 148:157-64 (2004).

[114] Drummond A. Cost of aspirin and statins in general practice. Q J Med, 95:23-26 (2002).

[115] Torgerson DJ, Raftery J. Economic notes. Discounting. BMJ, 319(7214):914-5 (1999).

[116] Gage, B.F. Stroke risk stratification in patients taking aspirin. Circulation, 110:2287-92 (2004).

[117] Van Ryn J, Standier J, Liesenfeld K-H, et al. Dabigatran etexilate: a novel, reversible, oral direct thrombin inhibitor: interpretation of coagulation assays and reversal of anticoagulant activity. Thromb Haemost, 103:1116-27 (2010).

[118] Guest JF, Watson HG, Limaye S. Modeling the cost-effectiveness of prothrombin complex concentrate compared with fresh frozen plasma in emergency warfarin reversal in the United Kingdom. Clin Ther, 32(14):2478-93 (2010).

[119] Wallentin L, Yusuf S, Ezekowitz MD, et al. Efficacy and safety of dabigatran compared with warfarin at different levels of international normalised ratio control for stroke prevention in atrial fibrillation: an analysis of the RE-LY trial. Lancet, 376(9745):975-83 (2010).

[120] Diener HC, Connolly SJ, Ezekowitz MD, et al. Dabigatran compared with warfarin in patients with atrial fibrillation and previous transient ischaemic attack or stroke: a subgroup analysis of the RE-LY trial. Lancet Neurol, 9(12):1157-63 (2010).

[121] Ezekowitz MD, Wallentin L, Connolly SJ, et al. Dabigatran and warfarin in vitamin K antagonist naive and experienced cohorts with atrial fibrillation. Circulation, 122(22):2246-53(2010).

[122] Freeman JV, Zhu RP, Owens DK, et al. Cost-effectiveness of dabigatran compared with warfarin for stroke prevention in atrial fibrillation. Ann Intern Med, 154(1):1-11 (2011).

[123] Sorensen SV, Kansal AR, Connolly S. Cost-effectiveness of dabigatran etexilate for the prevention of stroke and systemic embolism in atrial fibrillation: a Canadian payer perspective. Thromb Haemost, 105(5):908-15 (2011).

[124] Gallagher AM, Rietbrock S, Plumb J, van Staa TP. Initiation and persistence of warfarin or aspirin in patients with chronic atrial fibrillation in general practice: do the appropriate patients receive stroke prophylaxis? J Thromb Haemost, 6(9):1500-6 (2008).

[125] Hankey GJ, Eikelboom JW. Dabigatran etexilate: a new oral thrombin inhibitor. Circulation, 123(13):1436-50 (2011).

[126] Rolfe S, Papadopoulos S, Cabral KP. Controversies of anticoagulation reversal in life-threatening bleeds. J Pharm Pract, 23(3):217-25 (2010).

[127] RELY-ABLE long term multi-center extension of dabigatran treatment in patients with atrial fibrillation who completed RE-LY trial. http://www.clinicaltrials.gov/ct2/show/NCT00808067?term=RELY&rank=2 (2011).

[128] International Warfarin Pharmacogenetics Consortium, Klein TE, Altman RB, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. N Engl J Med, 360(8):753-64 (2009).

[129] Ahrens I, Lip GYH, Peter K. What do the RE-LY, AVERROES and ROCKET-AF trials tell us for stroke prevention in atrial fibrillation? Thromb Haemost, 105(4):574-8 (2011).

[130] Naccarelli GV, Varker H, Lin J, Schulman KL. Increasing prevalence of atrial fibrillation and flutter in the United States. Am J Cardiol, 104(11):1534-9 (2009).

[131] Coyne KS, Paramore C, Grandy S, Mercader M, Reynolds MR, Zimetbaum P. Assesing the direct costs of treating nonvalvular atrial fibrillation in the United States. Value Health, 9:348-56 (2006).

[132] Fuster V, Ryden LE, Cannom DS, et al. ACC/AHA/ESC 2006 Guidelines for the Management of Patients with Atrial Fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 Guidelines for the Management of Patients With Atrial Fibrillation). Circulation, 114:e257-e354 (2006).

[133] Gage BF, Fihn SD, White RH. Warfarin therapy for an octogenarian who has atrial fibrillation. Ann Intern Med, 134(6):465-74 (2001).

[134] Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med, 365(10):883-91 (2011).

[135] Granger CB, Alexander JH, McMurray JJV, et al. Apixaban versus warfarin in patients with atrial fibrillation. N Engl J Med, 365(11):981-92 (2011).

[136] Singer DE, Chang Y, Fang MC, et al. The net clinical benefit of warfarin anticoagulation in atrial fibrillation. Ann Intern Med, 151(5):297-305 (2009).

[137] Miller CS, Grandi SM, Shimony A, Filion KB, Eisenberg MJ. Meta-Analysis of Efficacy and Safety of New Oral Anticoagulants (Dabigatran, Rivaroxaban, Apixaban) Versus Warfarin in Patients With Atrial Fibrillation. Am J Cardiol, doi: 10.1016/j.amjcard.2012.03.049 (2012).

[138] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomised controlled trials. J Clin Epidemiol, 50(6):683-921 (1997).

[139] U.S. Department of Health and Human services. National Diabetes Statistics. http://diabetes.niddk.nih.gov/dm/pubs/statistics (2011).

[140] Hajjar I, Morley Kotchen J, Kotchen T.A. Hypertension: trends in prevalence, incidence and control. Ann Rev Public Health, 27:465-90 (2006).

[141] Connolly SJ, Eikelboom J, Joyner C, et al. Apixaban in patients with atrial fibrillation. N Engl J Med, 364(9):806-17 (2011).

[142] Lipscomb J, Weinstein MC, Torrance GW. Time Preference (Chapter 7). In Gold, M. R., J. E. Siegel, L. B. Russell, M. C. Weinstein. Cost-effectiveness in Health and Medicine. New York: Oxford University Press (1996).

[143] Shah SV, Gage BF. Cost-effectiveness of dabigatran for stroke prophylaxis in atrial fibrillation. Circulation, 123(22):2562-70 (2011).

[144] Mantha S, Ansell J. An indirect comparison of dabigatran, rivaroxaban and apixaban for atrial fibrillation. Thromb Haemost, 108(3) doi:10.1160/TH12-02-0093 (2012).

[145] Lip GY, Larsen TB, Skjøth F, Rasmussen LH. Indirect Comparisons of New Oral Anticoagulant Drugs for Efficacy and Safety When Used for Stroke Prevention in Atrial Fibrillation. J Am Coll Cardiol, doi:10.1016/j.jacc.2012.03.019 (2012).

[146] Banerjee A, Lane DA, Torp-Pedersen C, Lip GY. Net clinical benefit of new oral anticoagulants (dabigatran, rivaroxaban, apixaban) versus no treatment in a 'real world' atrial fibrillation population: a modelling analysis based on a nationwide cohort study. Thromb Haemost, 107(3):584-9 (2012).

[147] Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. Value in Health, 14(4):417-28 (2011).

[148] Altman DG, Bland JM. Statistics notes: Absence of evidence is not evidence of absence. BMJ, 311:485 (1995).

[149] O'Dell KM, Igawa D, Hsin J. New oral anticoagulants for atrial fibrillation: a review of clinical trials. Clin Ther, 34(4):894-901 (2012).

[150] Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. BMJ, 326: 472 (2003).

[151] Guzauskas GF, Hughes DA, Bradley SM, Veenstra DL. A risk-benefit assessment of prasugrel, clopidogrel, and genotype-guided therapy in patients undergoing percutaneous coronary intervention. Clin Pharmacol Ther, 91(5):829-37 (2012).

[152] Baczek VL, Chen WT, Kluger J, Coleman CI. Predictors of warfarin use in atrial fibrillation in the United States: a systematic review and meta-analysis. BMC Fam Pract, 13:5 (2012).

[153] Osterberg LG, Urquhart J, Blaschke TF. Understanding forgiveness: minding and mining the gaps between pharmacokinetics and therapeutics. Clin Pharmacol Ther, 88(4):457-9 (2010).

[154] Eerenberg ES, Kamphuisen PW, Sijpkens MK, Meijers JC, Buller HR, Levi M. Reversal of rivaroxaban and dabigatran by prothrombin complex concentrate: a randomized, placebo-controlled, crossover study in healthy subjects. Circulation, 124(14):1573-9 (2011).

[155] Radecki RP. Dabigatran: uncharted waters and potential harms. Ann Intern Med, 157(1):66-8 (2012).

[156] Uchino K, Hernandez AV. Dabigatran association with higher risk of acute coronary events: meta-analysis of noninferiority randomized controlled trials. Arch Intern Med, 172(5):397-402 (2012).

[157] Cairns J. Review: Dabigatran increases MI and reduces mortality compared with warfarin, enoxaparin, or placebo. Ann Intern Med, 156(12):JC6-11 (2012).

[158] Matchar DB, Jacobsen A, Dolor R, et al. Effect of home testing of international normalized ratio on clinical events. N Engl J Med, 363(17):1608-20 (2010).

[159] You JH, Tsui KK, Wong RS, Cheng G. Potential clinical and economic outcomes of CYP2C9 and VKORC1 genotype-guided dosing in patients starting warfarin therapy. Clin Pharmacol Ther, 86(5): 540-7 (2009).

[160] Wadelius, M, Chen LY, Lindh JD, et al. The largest prospective warfarin-treated cohort supports genetic forecasting. Blood, 113(4):784-92 (2009).

[161] McKeage K. Dabigatran etexilate: a pharmacoeconomic review of its use in the prevention of stroke and systemic embolism in patients with atrial fibrillation. Pharmacoeconomics, 30(9):841-55 (2012).

[162] Limdi NA, Veenstra DL. Warfarin pharmacogenetics. Pharmacotherapy, 28(9):1084-97 (2008).

[163] Salinger DH, Shen DD, Thummel K, Wittkowsky AK, Vicini P, Veenstra DL. Pharmacogenomic trial design: use of a PK/PD model to explore warfarin dosing interventions through clinical trial simulation. Pharmacogenet Genomics, 19(12):965-71 (2009).

[164] Anderson JL, Horne BD, Stevens SM, et al. Randomized and clinical effectiveness trial comparing two pharmacogenetic algorithms and standard care for individualizing warfarin dosing (CoumaGen-II). Circulation, 125(16):1997-2005 (2012).

[165] Hamberg AK, Wadelius M, Lindh JD, et al. A pharmacometric model describing the relationship between warfarin dose and INR response with respect to variations in CYP2C9, VKORC1 and age. Clin Pharmacol Ther, 87(6):727-34 (2010).

[166] van Schie RM, Wadelius MI, Kamali F, et al. Genotype-guided dosing of coumarin derivatives: the European pharmacogenetics of anticoagulant therapy (EU-PACT) trial design. Pharmacogenomics, 10(10):1687-95 (2009).

[167] Rosendaal FR, Cannegieter SC, van der Meer FJ, Briet E. A method to determine the optimal intensity of oral anticoagulant therapy. Thromb Haemost, 69:236-9 (1993).

[168] Baglin TP, Keepling DM, Watson HG. Guidelines on oral anticoagulation (warfarin): third edition - 2005 update. British Journal of Haematology, 132(3):277-285 (2006).

[169] Fennerty A, Thomas P, Backhouse G, Bentley DP, Campbell IA, Routledge PA. Flexible induction dose regimen for warfarin and prediction of maintenance dose. Br Med J, 288:1268-70 (1984).

[170] Meckley LM, Wittkowsky AK, Rieder MJ, Rettie AE, Veenstra DL. An analysis of the relative effects of VKORC1 and CYP2C9 variants on anticoagulation relates outcomes in warfarin-treated patients. Thromb Haemost, 100:229-39 (2008).

[171] Avery PJ, Jorgensen A, Hamberg AK, Wadelius M, Pirmohamed M, Kamali F. A proposal for an individualized pharmacogenetics-based warfarin initiation dose regimen for patients commencing anticoagulation therapy. Clin Pharmacol Ther, 90:701-706 (2011).

[172] Hillman MA, Wilke RA, Yale SH, et al. A prospective randomised pilot trial of model-based warfarin dose initiation using CYP2C9 genotype and clinical data. Clin Med Res, 3(3):137-45 (2005).

[173] Kovacs MJ, Rodger M, Anderson DR, et al. Comparison of 10mg and 5mg warfarin initiations nomograms together with low-molecular-weight heparin for outpatient treatment of acute venous thromboembolism: a randomized, double-blind controlled trial. Ann Intern Med, 138:714-19 (2003).

[174] Anderson JL, Horne BD, Stevens SM, et al. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. Circulation, 116:2563-70 (2007).

[175] Gage BF, Eby C, Milligan PE, Barnet GA, Duncan JR, McLeod HL. Use of Pharmacogenetics and Clinical Factors To Predict the Maintenance Dose of Warfarin. Thromb Haemost, 91(1):87-94 (2004).

[176] Zhu Y, Shennan M, Reynolds KK, et al. Estimation of warfarin maintenance dose based on VKORC1 (-1639 G>A) and CYP2C9 genotypes. Clin Chem, 53(7):1199-205 (2007).

[177] Solomon I, Maharshak N, Chechik G, et al. Applying an artificial neural network to warfarin maintenance dose prediction. Isr Med Assoc, 6(12):732-5 (2004).

[178] Ansell J, Hirsh J, Hylek E, et al. Pharmacology and Management of the Vitamin K Antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest. 133(suppl 6):160s-198s (2008).

[179] Wilson SE, Xonstantini L, Crowther MA. Paper-based dosing algorithms for maintenance of warfarin anticoagulation. J Thromb Thrombolysis, 23:195-98 (2007).

[180] Keeling D, Baglin T, Tait C, et al. Guidelines on oral anticoagulation with warfarin - fourth edition. Br J Haematol, 154(3):311-24 (2011).

[181] Platt, AB, Russell Localio A, Brensinger CM, et al. Risk factors for nonadherence to warfarin: results from the IN-RANGE study. Pharmcoepidemiol Drug Saf, 17(9):853-60 (2008).

[182] Reynolds MW, Fahrbach K, Hauch O, et al. Warfarin anticoagulation and outcomes in patients with atrial fibrillation: a systematic review and meta-analysis. Chest, 126(6):1938-45 (2004).

[183] Lenzine P, Wadelius M, Kimmel S, et al. Integration of genetic, clinical, and INR data to refine warfarin dosing. Clin Pharmacol Ther, 87(5):572-78 (2010).

[184] Department of Health. NHS reference costs 2010/11. London, DH http://www.dh.gov.uk/health/2011/11/reference-costs (2009).

[185] Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Economics, 10:779-87 (2001).

[186] Patrick AR, Avorn J, Choudhry NK. Cost-effectiveness of genotype-guided warfarin dosing for patients with atrial fibrillation. Circ Cardiovasc Qual Outcomes, 2(5):429-36 (2009).

[187] Eckman MH, Rosand J, Greenberg SM, Gage BF. Cost-effectiveness of using pharmacogenetic information in warfarin dosing for patients with nonvalvular atrial fibrillation. Ann Intern Med, 150(2):73-83 (2009).

[188] Meckley LM, Gudgeon JM, Anderson JL, Williams MS, Veenstra Dl. A policy model to evaluate the benefits, risks and costs of warfarin pharmacogenetic testing. Pharmacoeconomics, 28(1):61-74 (2010).

[189] You JH, Tsui KK, Wong RS, Cheng G. Cost-effectiveness of dabigatran versus genotype-guided management of warfarin therapy for stroke prevention in patients with atrial fibrillation. PLoS One, 7(6):e39640 (2012).

[190] Joo J, Geller NL, French B, Kimmel SE, Rosenberg Y, Ellenberg JH. Prospective alpha allocation in the Clarification of Optimal Anticoagulation through Genetics (COAG) trial. Clin Trials, 7(5):597-604 (2010).

[191] US National Institute of Health. Warfarin adverse event reduction for adults receiving genetics testing at therapy initiation (WARFARIN) trial. http://clinicaltrials.gov/ct2/show/NCT01305148 (2011).

[192] Sculpher MJ, Drummond MF. Analysis sans frontiers: can we ever make economic evaluations generalisable across jurisdictions? Pharmacoeconomics, 24(11):1087-99 (2006).

[193] Lalonde RL, Kowalski KG, Hutmacher et al. Model-based drug development. Clin Pharmacol Ther, 82(1):21-32 (2007).

[194] Chan PLS, Holford NHG. Drug treatment effects on disease progression, Annu Rev Pharmacol Toxicol, 41:625-59 (2002).

[195] Mould DR. Models for disease progression: new approaches and uses. Clin Pharmacol Ther, 92(1):125-31 (2012).

[196] Demin I, Hamren B, Luttringer O, Pillai G, Jung T. Longitudinal model-based meta-anlalysis in rheumatoid arthritis: an application towards model-based drug development. Clin Pharmacol Ther, 92(3):352-9 (2012).

[197] Mandema JW, Boyd RA, DiCarlo LA. Therapeutic index of anticoagulants for prevention of venous thromboembolism following orthopaedic surgery: a dose-response meta-analysis. Clin Pharmacol Ther, 90(6):8207 (2011).

[198] Lalonde RL, Willke RJ. Comparative efficacy and effectiveness: an opportunity for clinical pharmacology. Clin Pharmacol Ther, 90(6):761-3 (2011).

[199] Grasela TH, Slusser R. Improving productivity with model-based drug development: an enterprise perspective. Clin Pharmacol Ther, 88(2):263-8 (2010).