WORK IN PROGRESS. PLEASE DO NOT CITE OR CIRCULATE WITHOUT THE AUTHORS' PERMISSION

A COMPARISON OF DECISION MODELLING TECHNIQUES IN ECONOMIC EVALUATION: A CASE STUDY OF SCREENING FOR ATRIAL FIBRILLATION

Sue Jowett [1]

Pelham Barton [1]

David Fitzmaurice [2]

Stirling Bryan [1]

[1] *Department of Health Economics, University of Birmingham*

[2] *Department of Primary Care and General Practice, University of Birmingham*

Paper presented at the Health Economists' Study Group Meeting,
9-11th January 2008, Dunston Hall Hotel, nr Norwich

Address for correspondence:
Dr Sue Jowett, Department of Health Economics, Public Health Building,
University of Birmingham, Edgbaston, Birmingham, B15 2TT.
Tel: 0121 414 7898; e-mail: s.jowett@bham.ac.uk

ABSTRACT

Aims

An area of uncertainty in decision modelling is the level of structural complexity at which a patient level simulation (as opposed to a Markov cohort simulation) should be the approach of choice. The paper aims to compare the processes and outputs of these model types for screening for atrial fibrillation (AF), and to determine the most appropriate method for this case study.

Methods

Two models were constructed to address the long-term cost-effectiveness of screening for AF: a Markov cohort model and an individual sampling model (ISM). The models were compared using a number of criteria including structure, analytical input and results. The feasibility of, and results from, probabilistic sensitivity analysis (PSA) was also explored.

Results

Key features of the ISM were its inherent flexibility and ability to capture the full complexity of the decision problem by considering all patient attributes. Features were incorporated in the ISM with greater ease, and in some cases were not achievable in the Markov model without the structural complexity becoming impractical. In both models, PSA was feasible. However, the first-order uncertainty component of the ISM, resulted in greater uncertainty around the cost-effectiveness result. Only when the number of individuals included in the ISM runs became very large did the result approach that produced by the Markov model.

Conclusions

In this context, the ISM demonstrated greater flexibility and richness of structure compared with the Markov model. When faced with a decision problem, the choice of model type is critical and the complexity of the decision problem needs to be considered carefully.

Keywords: cost-effectiveness analysis; decision modelling; individual level model; Markov model; screening; atrial fibrillation

INTRODUCTION

Economic evaluation is increasingly recognised as a tool to inform decision makers by presenting information on the cost-effectiveness of health technologies.[1] Trial-based evaluations alone may not provide sufficient or appropriate information for decision-making, and suffer from a number of shortcomings including limited patient follow-up, intermediate measures of clinical effectiveness or the inability to address a particular clinical question.[2] Decision modelling has therefore become a standard tool for health economists, due the ability of models to synthesise the available evidence and estimate costs and effectiveness beyond observed outcomes. A number of issues have to be considered when constructing a decision model, such as the complexity of the clinical area, the amount of flexibility required, the appropriate time frame and the level of analytical input required.

Several authors have published papers attempting to guide health economists towards the right modelling approach for a given decision problem. Karnon (1998) described decision trees, Markov chain models and discrete event simulation (DES) using the choice of adjuvant therapies for early breast cancer to illustrate each type.[3] Barton et al (2004) categorised models into those based on independent individuals and those accounting for interaction between individuals, with decision trees, Markov models and Monte Carlo simulation included in the former category.[4] The use of the term 'individual sampling model' (ISM) was proposed to describe a model where each individual is modelled at a time and individuals are tracked as part of the model structure. Information about the patient such as clinical history and risk factors are carried through the model as patient attributes. Previous papers, including Karnon (1998) talk about DES, and others describe a state-transition model or models using Monte Carlo simulation.[3] The remainder of this paper will describe all models of this kind as ISMs, and concentrate on the choice between Markov models and ISMs for decision problems that require recurrent events to be followed over a long period of time.

Karnon (2003) developed his previous work by presenting both model types for the same decision problem, comparing the models in terms of flexibility and level of complexity of analytical input.[5] He concluded that, as the results were very similar, the benefits in terms of flexibility were outweighed by the time to run and develop the model. However, Barton et al (2004) recommended that the choice of model in this situation be dependent on whether a Markov model required an "excessive number of states". More recent publications on model

choice have recognised that individual level models have a place in health economics and are becoming more prevalent.[6,7] Brennan et al (2006) argued that representing the disease appropriately should precede the choice of model, as adopting inappropriate assumptions in order to build a Markov cohort model may lead to inadequate or incorrect results. The paper also states that when attempting to represent a number of risk factors in this type of model, the number of health states may increase to an excessive level.[6]

This paper compares the processes and outputs of two alternative types of model (Markov cohort and ISM) for the same decision problem, screening for atrial fibrillation (AF). AF is the most common chronic heart arrhythmia and is most prevalent in those aged 65 and over. The condition is associated with an increased risk of ischaemic stroke but is often asymptomatic unless identified during clinical examination. Once detected, usually by electrocardiogram (ECG), treatment options are available including anticoagulation (warfarin) or aspirin. Warfarin is highly effective in reducing the risk of stroke but has associated adverse haemorrhagic complications. Although AF can be detected by means of an ECG, no screening programme currently exists in the UK, and there is no evidence on long-term cost-effectiveness of screening.

METHODS

The decision problem

The primary purpose of the modelling work was to answer a number of policy questions concerning AF screening. Firstly to determine if screening for AF is cost-effective compared with current policy of no screening. Two screening policy options were considered: opportunistic screening with pulse-taking at a routine consultation, and an ECG conducted if the pulse was irregular; and systematic screening where all patients would be invited to an ECG clinic. Secondly, what is the optimum frequency of screening? Other policy questions included the type of ECG (i.e. 12-lead or less) and interpretation of the ECG.

This work was undertaken alongside the Screening for Atrial Fibrillation in the Elderly (SAFE) study, a randomised controlled trial to assess the cost-effectiveness of screening of AF in patients aged 65 and over.[8] In addition to a within-trial analysis, modelling was undertaken to measure the benefits and costs of a regular programme of screening in this patient group for their remaining lifetime. Assessment of the long-term outcomes of AF

screening took into account compliance with and detection by screening, warfarin therapy to reduce ischaemic stroke risk, major clinical events (ischaemic strokes, haemorrhagic events), health-related quality of life, patient survival and the costs associated with screening and long-term events.

The models

The models built to address the decision problem were a Markov model and an ISM. The Markov model was constructed using TreeAge Pro 2005 (version 0.8), analysed as a cohort. The ISM was constructed in Borland Delphi 7. The main model scenario reported in this paper is annual systematic screening with a 12-lead ECG interpreted by a cardiologist. Results are presented for males aged 65 without AF, and the comparator is no screening. Alternative scenarios are discussed later in the paper. The structure of the models was based on those events likely to have the most impact on costs and outcomes. The time horizon was patient lifetime, costs were from a health service perspective, and costs and outcomes were discounted at 3.5%. The model structure for the ISM is presented in Figure 1. The diagrams for the Markov model form the Appendix.

Markov model

The Markov model was built as a cohort model with a monthly time cycle, with separate model runs for men and women. Several events were possible, but only one event could occur during a cycle. A patient had a risk of developing AF, dependent on age and gender and all patients with undiagnosed AF or no AF had a probability of being screened. This probability was dependent on a number of factors including compliance, which was related to age and gender, frequency of screening, and the sensitivity and specificity of the screening method. If a patient with undiagnosed AF was diagnosed by screening, or through clinical examination after a stroke, treatment commenced immediately. However, if the patient had previously suffered a haemorrhagic stroke, then warfarin therapy was not given, due to the high risk of further bleeding.

Major clinical events included ischaemic stroke which could be fatal or non-fatal, and those surviving a stroke of this type could suffer a mild or disabling stroke. If a previous ischaemic stroke event was mild, a subsequent event would be disabling. A patient disabled from a previous stroke was assumed to die from a subsequent stroke. Ischaemic stroke risk was

related to a number of patient characteristics, namely age, gender, treatment, AF status and previous stroke history. Haemorrhagic stroke could be fatal or non-fatal, but all survivors were assumed to be disabled and any further stroke resulted in death. Risk of a haemorrhagic stroke was related to age and treatment. If a patient was on warfarin and suffered a stroke of this kind, then treatment was discontinued. A gastrointestinal (GI) bleed was a temporary non-fatal event lasting one month, with risk related to age and therapy. Treatment was assumed to be discontinued for one month. Finally, a patient could also die from non-stroke causes, with mortality related to age and gender.

The initial health state for a patient in the model was "no AF, well". Additional health states were constructed to represent all aspects of patient history influencing the probability of future events. Health states indicated the status of a patient with regards to presence or absence of AF, if AF had been diagnosed, previous ischaemic or haemorrhagic stroke, including level of disability, and temporary GI bleed status. In total, 31 health states were required in the Markov model to represent all the important aspects of a patient history.

Cost and utility inputs for the model are reported in Table 1. Costs included one-off costs and recurring costs. One-off costs occurred at screening and at acute stroke events, fatal strokes and GI bleeds. Recurring costs were warfarin monitoring and long-term care for disabling stroke. Health states with utility values less than full health were those where a stroke event had occurred (with a lower quality of life for a disabling event) or the patient was on warfarin, which was assumed to marginally lower a patient's quality of life. In addition, a GI bleed was assumed to temporarily reduce quality of life over a one month period.

Individual sampling model (ISM)

The individual sampling model (ISM) considered individual patients passing through the model, their status being changed at various times according to the natural history of AF, clinical events and the effects of a screening programme. The aim of the model was to produce a realistic set of virtual patient histories, from which estimates of population mean costs and effects could be estimated.

The model was constructed to sample the time to each possible event. The event having the earliest time was the one that occurred first and all attributes were updated. Resampling took place again to find the earliest event, with these processes continuing until the death of the individual. Although the model was constructed to run in this way with no fixed time cycle, a monthly time cycle could also be imposed, so that these results could also be compared with the Markov model.

The ISM was constructed to be run for men or women or any combination of the two. In addition, model runs could be undertaken for 65 year olds only, or by sampling from a distribution of a population aged 65 and over. The model was run for the population of choice by simply choosing the relevant option and reinitialising the data. Changes of this nature were facilitated by use of a "front-end" for the model which contained a number of features including allowing the appropriate population and scenarios to be chosen, setting of the sample size and viewing the model status and results.

The model calculated the time to each relevant event namely ischaemic stroke, haemorrhagic stroke, GI bleed, death by other causes, and where relevant, time of the next screening event, recommencement of treatment and age of routine AF diagnosis. When an individual experienced an event, their attributes were updated and then resampling occurred taking into account this new information and any changes this would make to risk estimates.

Structural differences between models

There were several features not included in the Markov model that the ISM could incorporate. A restriction was made on the number of GI bleeds that could be suffered before warfarin was discontinued. This was possible by tracking the number of GI bleeds in the individual attributes, so when the second bleed was experienced, treatment was stopped. In order that an appropriate comparison could be made between the results of the two models, the maximum number of GI bleeds was changed from two to a unrealistically large number. Incorporation of the background detection of AF was also possible in the ISM. A timescale of three years from development of AF was used, with random variation attached to this point estimate. As part of the patient attributes, the age AF developed is held is order that the time to routine detection can be determined. Again, to ensure fair comparison of results, this variable was set to a large enough value to ensure background detection did not occur.

The ISM was constructed to take into account all possible screening strategies. Incorporation of an opportunistic screening was straightforward, with the additional steps simply added to the model syntax. The pulse taking was programmed to take place at some point during the year, thus attempting to represent the random nature of patient attendance at their general practice. Systematic screening was assumed to occur when an individual reached the age of screening, an increment of time added to the start age or time of the previous screening event. Any combination of ECG type and method of ECG interpretation could be run by the model. An array was constructed containing data on diagnostic accuracy, cost of the test and interpretation and any additional costs of confirmation of diagnosis, with each row representing a particular combination.

Running the models

The Markov model was run as a cohort simulation producing expected values of costs and QALYs. The ISM was run for 500,000 patients in order that the mean cost and QALY values were as close as possible to a "true" mean value, for appropriate comparison with the Markov model. The ISM was run as a "time to next event" model, but additional runs were also undertaken with a monthly time cycle imposed. The mean values (ISM) and point estimates (Markov) were compared. In addition, the total number of major events were also counted for comparison, and for this the ISM was run 10 million times, again to attempt to achieve values close to the "true" number of events.

Probabilistic sensitivity analysis (PSA) was also undertaken on both models in order to compare processes, including running time, and output. The Markov model was run as a Monte Carlo simulation for 10,000 replications. The ISM with a monthly time cycle was run with simultaneous first- and second-order simulation, the latter randomly sampling from probability distributions (where available) of input parameters. In the base case, the ISM was run for 10,000 patients and 10,000 replications. However further model runs using 20,000 patient (5000 replications) and 50,000 patients (2000 replications) were conducted to investigate the effect of increasing the number of patients. Cost-effectiveness planes and cost-effectiveness acceptability curves (CEACs) were constructed for all PSA runs.

RESULTS

The final analysis of the two models was undertaken on a 1.10Ghz laptop with a Pentium M processor. In terms of the time required to run the models, the Markov model required less time for base case analysis and PSA than the ISM. The results of the Markov cohort simulation appeared almost immediately compared with a first order simulation of 500,000 for an ISM which took 40 seconds. An increase in the number of individuals in the simulation increased the running time. Model runs for PSA took longer, with two hours required for 10,000 simulations by the Markov model, and six hours required for 10,000 simulations of 10,000 individuals in the ISM.

The Markov model does not consider first-order uncertainty, and presents expected values of costs and QALYs for a cohort. In contrast, the ISM can present quasi-standard errors giving information on the uncertainty around the costs and QALYs produced. Costs and QALYs from the base case analyses are presented in Table 2. Mean costs for all strategies are slightly lower for the Markov model and highest for the ISM with a monthly cycle, although differences are very small. Mean QALYs are marginally higher for the Markov model than the ISM with a monthly cycle. The ISM without a fixed time cycle reports the lowest values, in the region of 0.04 QALYs lower. However, the overall cost-effectiveness results for all model runs show annual screening to be dominant over no screening. Table 3 considers the number of events reported for each model. For all events and for all strategies, the Markov model reports the least number of events, although in relative terms the differences are rather small, approximately 1%. As a general rule, the ISM with a fixed time cycle resulted in the most patient events.

The results of the PSA for the Markov model and the ISM are shown in Figure 2 and Figure 3 respectively. The CE plane from the Markov model shows almost all the data points are in the south-east (SE) quadrant, with little variation in the data pairs therefore confined to a limited area on the plane. The CEAC bears out the distribution of points on the plane, with an almost 100% probability of screening being cost-effective. The CE plane for the ISM also shows a concentration of points in the SE quadrant, however there are data pairs in all quadrants and the spread of points in rather wide. The CEAC demonstrates screening has between an 80% and 90% chance of being cost-effective.

Two alternative PSA runs were carried out on the amended ISM, for 20,000 individuals for 5,000 replications and 50,000 individuals for 2,000 replications. The CE planes and 95% confidence ellipses for all three runs are shown in Figure 4. Running the model for 20,000 individuals decreases the spread of points on the plane, with greater concentration in the SE quadrant. Further increasing the number of individuals to 50,000 again concentrates the cost-QALY pairs in the SE quadrant, and the scatter of the points on the plane becomes tighter. Translating these planes to CEACs would result in the curves approaching a 100% probability of screening being cost-effective, thus mirroring the results of the Markov model.

DISCUSSION

In simple terms, it is evident that the results produced by the two models when built with the same structure and assumptions are actually very similar, and the final recommendation on cost-effectiveness would essentially be the same. The case study presented by Karnon (2003) also demonstrated very similar results for the two model types.[5] This led him to argue that any increase in flexibility using a patient level simulation was outweighed by a shorter running time and lower level of analytical input, therefore the Markov model should be chosen. However, there are other matters to be considered when deciding the most appropriate model type. The fact that some of the features easily incorporated into the ISM were problematic for the Markov model, raises the point that the level of detail required to adequately represent the decision problem needs to be taken into consideration.

Prior event history is held in the ISM as attributes, in contrast to the Markov model which requires health states to record any important characteristics having an influence on the probability of events occurring. For example, the tracking of ischaemic stroke history is required, not only due to effect on the risk of further stroke, but also to incorporate the assumption applied to this type of stroke regarding severity and the number of strokes allowed before death. If this assumption was altered to more than two strokes before a fatal stroke, the ISM would only require minor modifications to the syntax, however the number of health states in the Markov model would increase multiplicatively. For this reason, no restriction was applied to the number of GI bleeds suffered before discontinuation of warfarin therapy, although easily incorporated into the ISM. Due to the Markovian assumption, the Markov model requires the incorporation of a number of health states to keep track of patient history. This is because the model is 'memory free' therefore transitions

to health states are not related to patient history and a cohort moving from one state to another must be homogenous. In order to take account of all patient characteristics, a large number of health states are required to ensure every combination of attributes is accounted for. The flexibility of the ISM means that the incorporation of clinical events and assumptions can be undertaken by the inclusion of additional syntax. In contrast, the incorporation of these sorts of features is less straightforward in a Markov model, and may require the addition of further health states, thus increasing the size of the model structure.

In order to determine the cost-effectiveness of screening for AF, alternative screening strategies (systematic, opportunistic), different populations (65 year olds, general population of 65 and over), and choice of screening method (ECG type and interpretation) required exploration. The ISM demonstrated its superior flexibility compared with the Markov model which would have required additional branches and alteration of variable values. Flexibility was also demonstrated by the ISM by its ability to include routine detection of AF as the age of AF development was known.

One of the further key differences between a Markov model and an ISM is that in the latter model the times of events can be recorded to full computer accuracy rather than simply being regarded as taking place in a given cycle. A Markov model is restricted to a particular sequence of events, therefore consideration is required with regards the most appropriate order of events.

Base-case model results were very similar, however the output produced by the PSA for both models differed, with the Markov model giving a more 'exact' result than the ISM. The combination of number of individuals and simulations when conducting a PSA using the ISM has a considerable impact on the results. As the sample size increased, the spread of points on the cost-effectiveness plane became more concentrated and the probability of cost-effectiveness of screening increased. By running the PSA for an increasingly large number of individuals, the result of the PSA becomes more 'exact' and more closely resembles the output from a Markov PSA. However, running the model with a sample size closely resembling that of the patient group of interest may be more relevant as the inherent variability in this population will be characterised. For example, running the model for 10,000 individuals may represent the number of people eligible for screening in a Strategic Health Authority.

Karnon (2003) considered analytic input to be an important factor, in terms of the level of complexity in building the model and the total time required for construction, validation and model runs.[5] Griffin et al (2006) argued that patient level simulations were 'computationally expensive' thus having an impact on the suitability of running a PSA.[9] In this case, the Markov model required less time for running, particularly for the PSA, however the difference in running time was a matter of hours. It may be more important to have a flexible model which can have a rich structure and address all the required issues than have a less satisfactory model which takes several hours less to run. The time required to generate and assimilate the data to populate a model is months or years, therefore the extra time required to run an ISM is rather inconsequential when compared with the time required to generate the data in the first place. This is in agreement with the response to Griffin et al by Caro (2007) who stated that the solution regarding PSA should not be to build a more simple model as 'untenable' assumptions may have to be incorporated.[10]

A further concern about building an ISM is the level of expertise required, as the models are often built in software requiring programming skills. The ISM presented here was built in Borland Delphi 7, and the skills required to build the model were gained over time with guidance from an experienced mathematical modeller. Therefore it is true that expertise is required, either from an experienced modeller or through appropriate training. However, this is also true for other aspects of health technology assessment, where, for example, the expertise of statisticians and systematic reviewers is required. The important point to be made is that the model should be built to appropriately represent the decision problem in terms of patient population and clinical pathways, and suitable expertise sought to achieve this rather than producing an oversimplified model. One solution could be to run the Markov model with Monte Carlo simulation, thereby building the model in a more familiar package and considering first order uncertainty. However, this type of model is still constrained by health states, even if tracker variables are used to follow patient events. Therefore one would gain little extra benefit by building the model in this way if a decision problem requiring a rich structure is faced.

Strengths and limitations

The key strength of this work is that an empirical comparison of two types of model for the same decision problem has been undertaken. In particular, the work has demonstrated the use of a PSA alongside an ISM. In addition, by building two different models for the same decision problem, the validity of the models could be tested, by comparing their outputs. This type of validity check is rarely conducted, because almost always only one model is build to address a research question.

There are limitations of the work undertaken here. External validity checks, to determine the predictive power of the model, were not undertaken. The most appropriate way to undertake these checks would be by comparing model output, particularly event rates, with 'actual observations'. These observations are likely to take the form of published incidence and prevalence data. However, one would have to ensure these real world data represented a population similar to that used in the model.

The model comparison only considered a Markov model and an ISM, however other model types are available. In addition to decision tree models, there are also models that account for interaction between individuals: system dynamics models and discrete event simulation. Therefore the work presented here cannot be generalised to other model types, and further considerations have to be made, particularly when dealing with interactions. These models are beyond the scope of this paper. Finally, only AF was considered, therefore the results may not be generalisable to other disease areas.

Recommendations for future work

The only other published work to date presenting a direct comparison of Markov and individual level models is that by Karnon (2003), as previously discussed. In addition, it is evident that there is a gap in the literature in the presentation of ISMs with no fixed time cycle and no interaction between individuals. Therefore to strengthen the evidence base for the ISM, more models of this type need to be built and further comparisons with Markov models are required. Most importantly, evaluations of this type are required in other disease areas, to determine whether the results of this comparison are generalisable.

CONCLUSIONS

Taking into account analytical input required, flexibility of incorporating further features, richness of structure and the ability to incorporate a PSA, the ISM appears to be the more appropriate type of model for the case study presented. This empirical evidence can also form the basis of general recommendations on the choice between a Markov model and an ISM. When faced with a decision problem with no interactions where patients are followed over a long period of time, the analyst needs to consider the level of detail required in terms of patient characteristics and history to adequately answer the research question. If there are a relatively small number of possible events, and risk factors do not have to be addressed in detail, then a Markov model will be more than adequate, as the number of health states will not be excessive. However, if there are many factors to be considered leading to a large number of health states in order to follow patient history, then an ISM is recommended. Although ISMs require more analytical input in terms of model building and running the model, the analyst should not be dissuaded from using this type of model when in all other respects it appears to be the most appropriate. However an ISM should not be used when a more detailed model is not required. Not only will there be over-specification of the model, there may also be difficulties in identifying data to populate the model.

Table 1 Model parameters: utility values and costs

Utilities	
Well, no therapy	1.0
Well, warfarin therapy	0.986
Mild ischaemic stroke	0.75
Disabling ischaemic stroke	0.39
Disabling haemorrhagic stroke	0.39
Gastrointestinal bleed	0.88
Dead	0.0
Costs	**£**
Mild ischaemic stroke	6820
Disabling ischaemic stroke	4550
Disabling haemorrhagic stroke	4550
Annual (long term) disabling stroke costs	13240
Fatal stroke	8830
Gastrointestinal bleed	1130
Warfarin treatment (annual)	100
Invitation/reminder (systematic screening)	0.70
Screening ECG clinic visit	16.25
12 lead ECG Interpretation by consultant	2.05

Table 2 Comparison of costs and QALYs for annual systematic screening and no screening

Cohort	Screening	Model type	Mean Cost (£)	Mean QALYs
Males aged 65	**Annual**	**Markov**	6730	10.55
		ISM monthly cycle	6765	10.54
		ISM time to event	6749	10.51
	None	**Markov**	7010	10.49
		ISM monthly cycle	7055	10.49
		ISM time to event	7022	10.45
Females aged 65	**Annual**	**Markov**	8439	12.48
		ISM monthly cycle	8462	12.47
		ISM time to event	8442	12.44
	None	**Markov**	8574	12.43
		ISM monthly cycle	8609	12.42
		ISM time to event	8574	12.39

Table 3 Comparison of clinical event numbers for annual systematic screening versus no screening

Cohort	Screening	Model	Ischaemic strokes (,000s)	Haemorrhagic strokes (,000s)	GI bleeds (,000s)	AF (,000s)
Males aged 65	**Annual**	**Markov**	1,824	802	2,421	3,279
		ISM cycle	1,844	811	2,459	3,299
		ISM no cycle	1,846	805	2,452	3,289
	None	**Markov**	2,444	707	1,975	3,279
		ISM cycle	2,476	718	2,002	3,300
		ISM no cycle	2,464	713	1,995	3,288
Females aged 65	**Annual**	**Markov**	2,053	1,173	3,075	2,699
		ISM cycle	2,069	1,186	3,113	2,714
		ISM no cycle	2,070	1,177	3,106	2,706
	None	**Markov**	2,620	1,060	2,621	2,699
		ISM cycle	2,649	1,072	2,646	2,712
		ISM no cycle	2,639	1,065	2,640	2,704

Figure 1 Individual sampling model

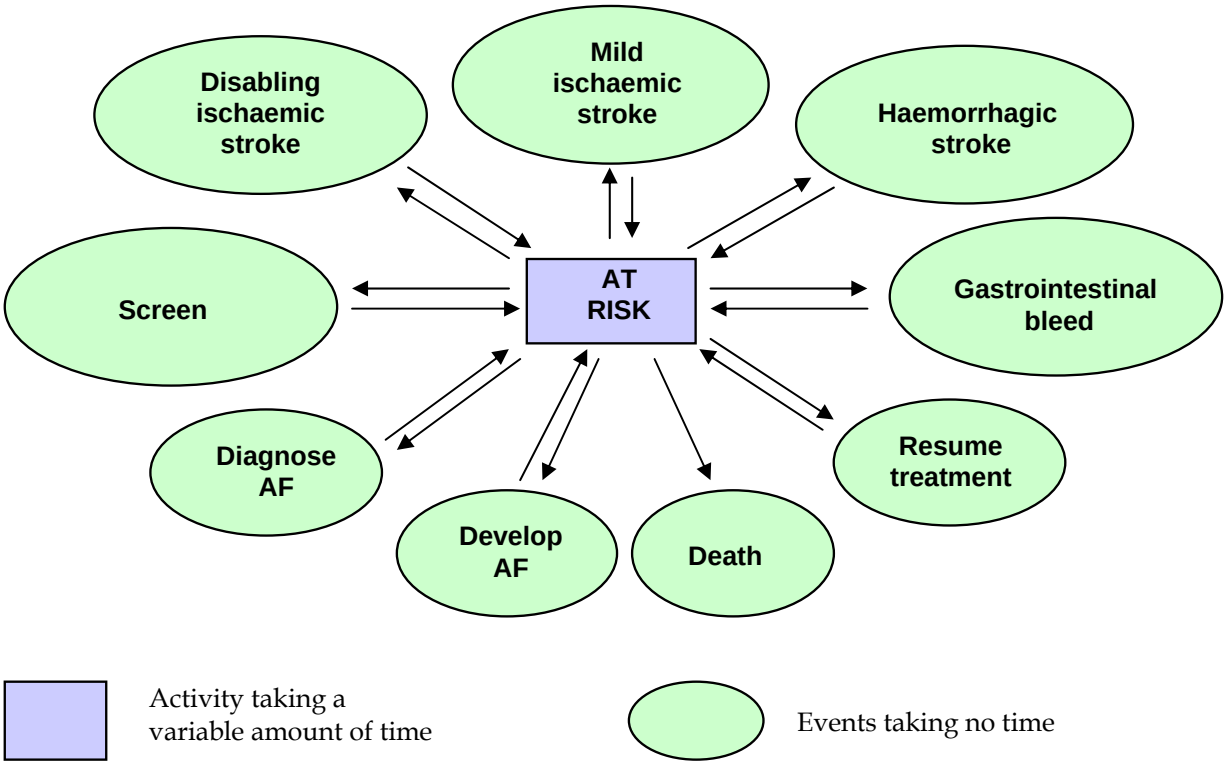
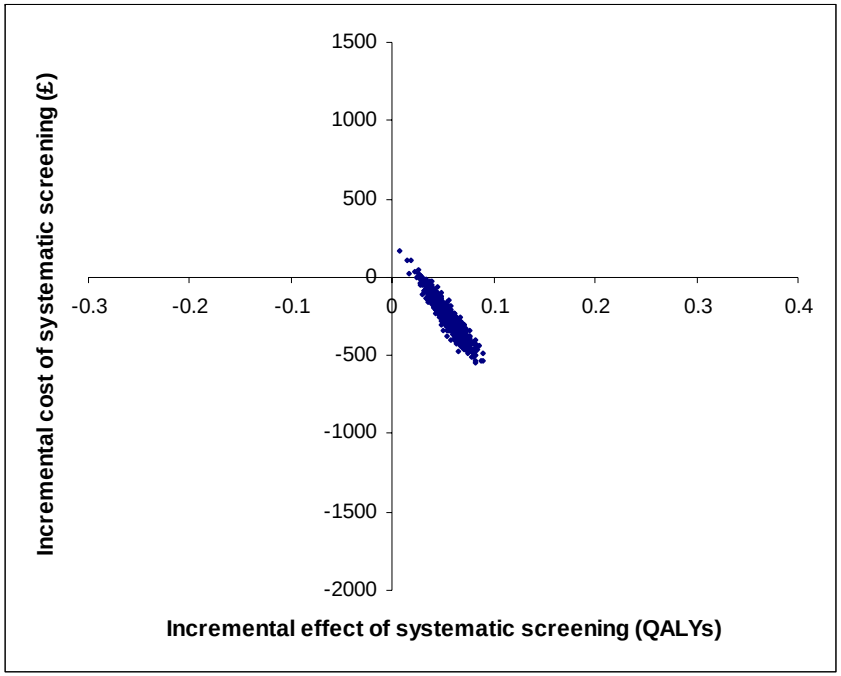


Figure 2 Cost-effectiveness plane and CEAC for the Markov model (Annual systematic screening, men aged 65)



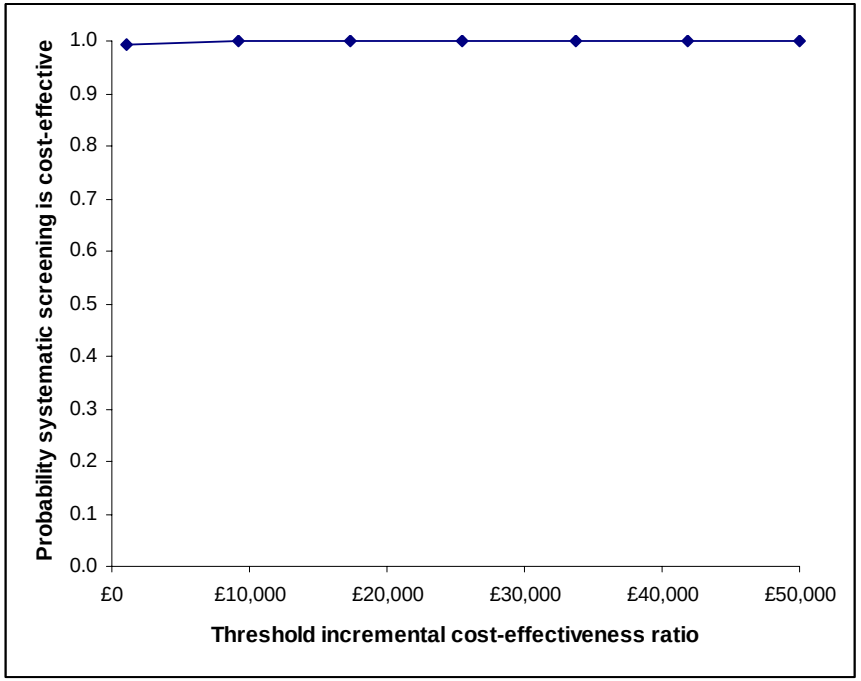
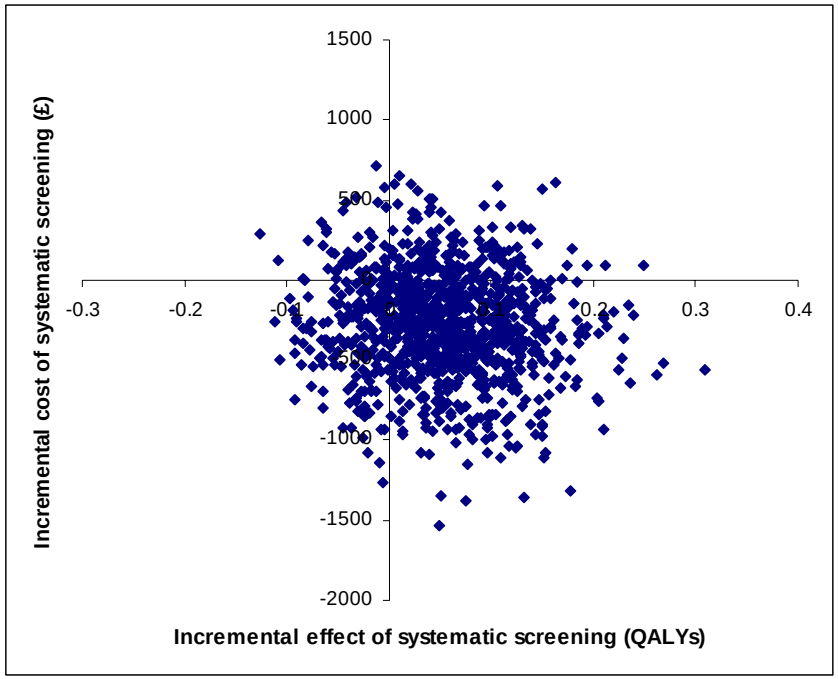


Figure 3 Cost-effectiveness plane and CEAC for the ISM (Annual systematic screening, men aged 65)



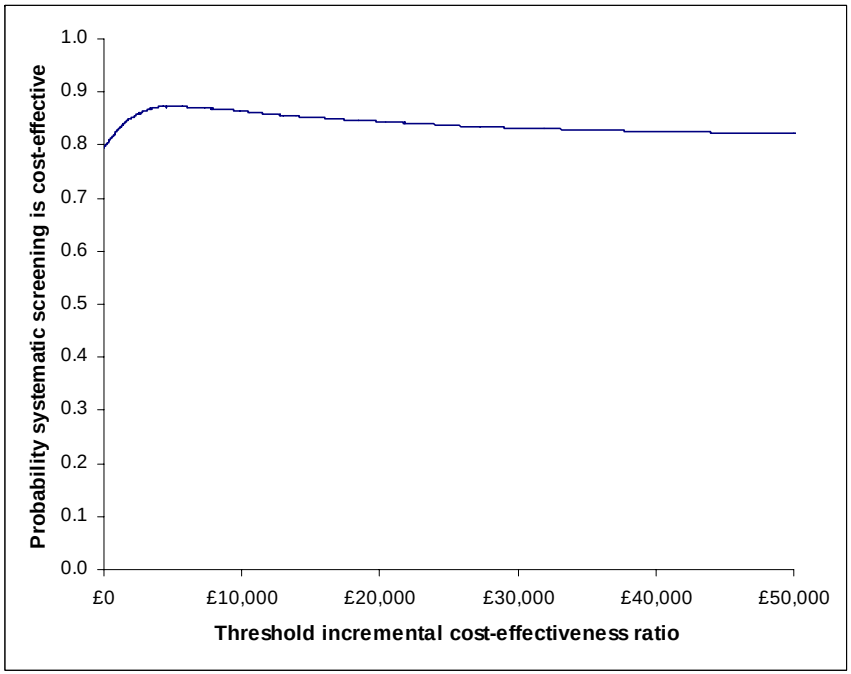
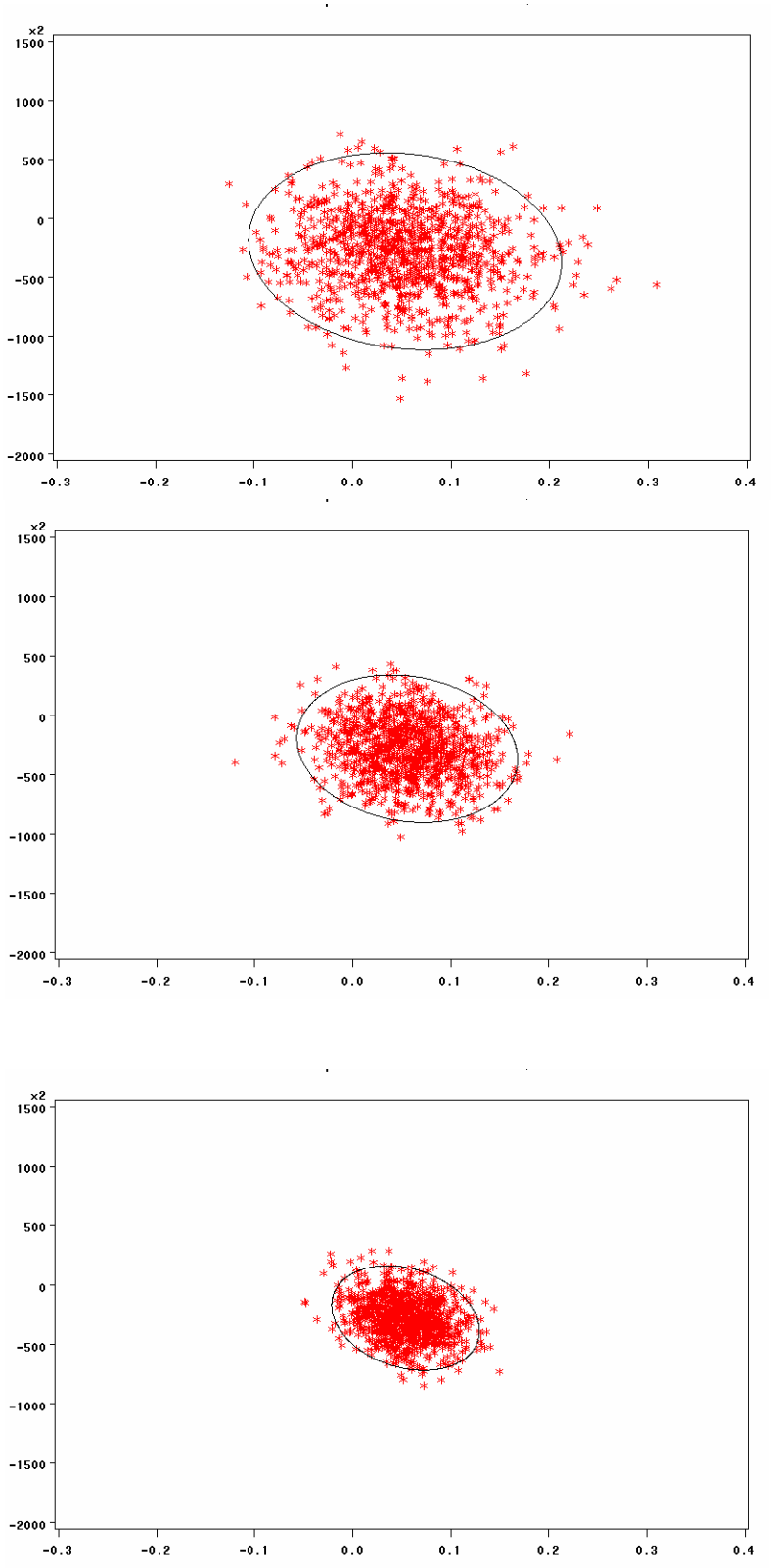


Figure 4 95% Confidence ellipses for the ISM incremental CE planes



Reference List

(1) Bryan S, Williams I, McIver S. Seeing the NICE side of cost-effectiveness analysis: A qualitative investigation of the use of cost-effectiveness analysis in NICE technology appraisals. Health Economics 2007; 16(2):179-193.

(2) Sculpher M, Claxton K, Drummond MF, McCabe C. Whither trial-based economic evaluation for health care decision making? Health Economics 2006; 15:677-687.

(3) Karnon J, Brown J. Selecting a decision model for economic evaluation: a case study and review. Health Care Management Science 1998; 1:133-140.

(4) Barton P, Bryan S, Robinson S. The use of modelling in the economic evaluation of health care. Journal of Health Services Research and Policy 2004; 9:110-118.

(5) Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. Health Economics 2003; 12:837-848.

(6) Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Economics 2006; 15:1295-1310.

(7) Cooper K, Brailsford SC, Davies R. Choice of modelling technique for evaluation health care interventions. Journal of the Operational Research Society 2007; 58:168-176.

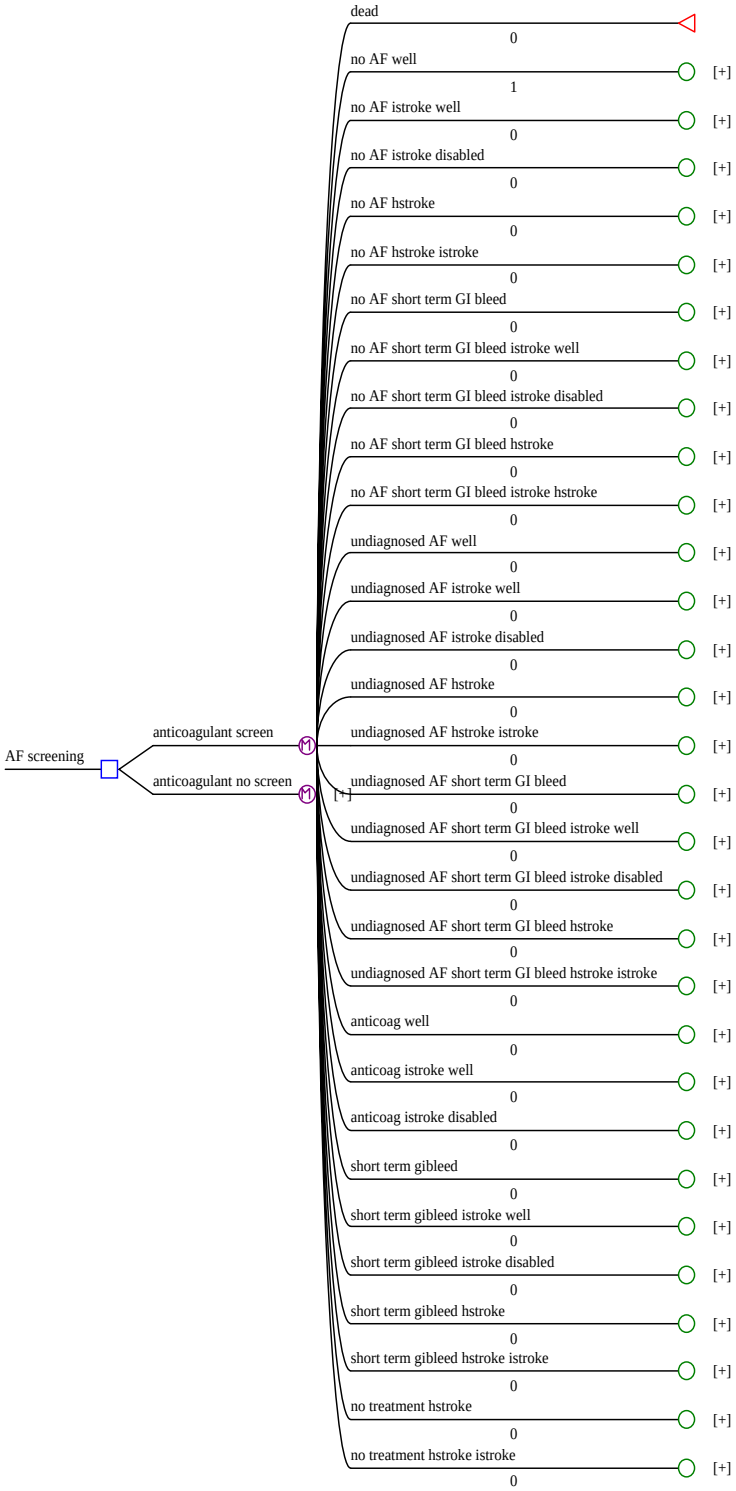
(8) Hobbs FDR, Fitzmaurice DA, Jowett SM, Mant J, Murray ET, Bryan S et al. A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over. The SAFE study. Health Technology Assessment 2005; 9(40):1-90.

(9) Griffin S, Claxton K, Hawkins N, Sculpher M. Probabilistic analysis and computationally expensive models: necessary and required? Value in Health 2006; 9(4):244-252.

(10) Caro JJ. Regarding probabilistic analysis and computationally expensive models: necessary and required? Value in Health 2007; 10(4):317.

Appendix

Markov model health states: screening arm



See next figure

Branches from the ‘No AF well’ health state