Characterising structural uncertainty in decision analytic models: review and application of currently available methods

Laura Ginnelly
Centre for Health Economics, University of York

WORK IN PROGRESS: PLEASE DO NOT QUOTE WITHOUT AUTHOR PERMISSION

Acknowledgements:
I would like to acknowledge the advice and technical support of my PhD supervisors Dr Karl Claxton and Professor Mark Sculpher. This work is supported by a 3-year fellowship in evidence synthesis by the National Co-ordinating Centre for Research Capacity Development, Department of Health.

Abstract

Background

An inappropriate model structure can potentially invalidate estimates of cost-effectiveness and estimates of the value of further research. However there are often a number of possible and credible structural assumptions, which can be made. Although it is common practice to acknowledge potential limitations in model structure, there is a lack of clarity about methods to characterize structural uncertainties and their contribution to decision uncertainty.

Methods

This review identifies approaches to handle structural uncertainty available in health technology assessment and other disciplines. I then go on to apply these alternative methods to the assessment of structural uncertainty using a case-study model, screening for age related macular degeneration. This application provides insights into the value and practicality of the alternative methods.

Results

The assessment of structural uncertainty has received little attention in the health economics literature. Despite the fact that, in many circumstances, it is uncertainty over appropriate structure, which is acknowledged as most important by decision-makers. A common method to characterise structural uncertainty is to compute results for each alternative model specification, and to present alternative results as scenario analyses. It is then left to decision maker to assess the credibility of the alternative structures in interpreting the range of results. An alternative method, which has not been widely used, is ‘model averaging’ where alternative models, with different specifications, are built, and their results averaged, using explicit prior distributions often based on expert opinion.

In this case-study model, alternative assumptions regarding model structure have little impact on the decision regarding cost-effectiveness, but have a big impact on the value of additional research estimates.

Background

Decision analytic models are increasingly being used to inform policy questions regarding the optimum allocation of health care expenditures. Although analysts seek to develop models and incorporate data that most accurately present the costs and outcomes associated with a particular disease and intervention, some degree of uncertainty is present in the majority of models.

The dimensions of uncertainty have previously been categorised as parameter, structural and methodological[1]. Within health technology assessment analysts have tended to focus almost entirely on quantifying and assessing the impact of parameter uncertainty. As a consequence methods for dealing with parameter uncertainty are well-developed and common practice[1]. Structural uncertainty, however, has received relatively little attention, although many guidelines of good modelling practice recognize the need to explore structural assumptions[2-4]and the evidence supporting the chosen model structure. Many technicians of 'good practice' advocate some kind of sensitivity analysis using alternative model structures[3]. Exactly what form this sensitivity analysis take is unclear. One option is to compute cost effectiveness estimates for each alternative structural assumption and to examine the results. These alternative structural assumptions can also be averaged using weights or priors[5], however it is uncertain how and from whom these are determined.

This paper reviews alternative methods available to characterise structural uncertainty. The alternative methods are summarised briefly in terms of their applicability to HTA decision models. Suitable methods are then applied to a HTA case-study model, screening for age related macular degeneration. Although the structure of the model developed was consistent with published evidence regarding the natural history of the disease, a number of structural assumptions were questioned after consultation with clinical experts. Alternative structural assumptions, consistent with these alternative clinical opinions, were therefore explored.

Methods

Review of the literature

A systematic search was conducted that looked for papers relating to the assessment, identification or quantification of structural uncertainty. As very little on structural uncertainty has been published in the health economics literature, the searches were not restricted to medical or economics databases. Because of the enormity of the available literature, as a result of widening the database search, searches were restricted to identifying key words in the title. In addition citations from relevant papers were also obtained. Although this search is not necessarily comprehensive, extending the identification of key words to the title and abstract increased the number of hits to more than 80,000, which is an unfeasibly large number of records to shift through. The search strategy used was:

Science Citation Index 1980-2004:

model*)

and (robustness or uncertain* or forecast* or accura* or fit for purpose or predict* or inference)

and (structur* or develop* or construct* or specification or selection)

[title]

Methods of the case-study model

The case-study model looked the use of weekly self-screening, using an Amsler grid, following 1st eye involvement with neovascular AMD. This self-screening strategy is compared to two alternatives: no screen but diagnosis and treatment (with photodynamic therapy) of eligible AMD following self-referral (due to declining visual acuity of one or more lines) to an ophthalmologist (this strategy is consistent with provisional NICE guidance); and a strategy of no screening and no PDT. The model assessed the cost-effectiveness of, and potential value of future research for, these alternative strategies. Full details of this model are reported elsewhere[6].

The structure of the decision model is illustrated in Figure 1. A Markov process[7] is used to model the incidence of 2nd eye neovascular AMD over 10-years and the associated decline in visual acuity following undiagnosed 2nd eye involvement. People can enter the model at two alternative starting visual acuities, 20/40 or 20/80. To reflect uncertainty in the parameters in the model, they were incorporated as

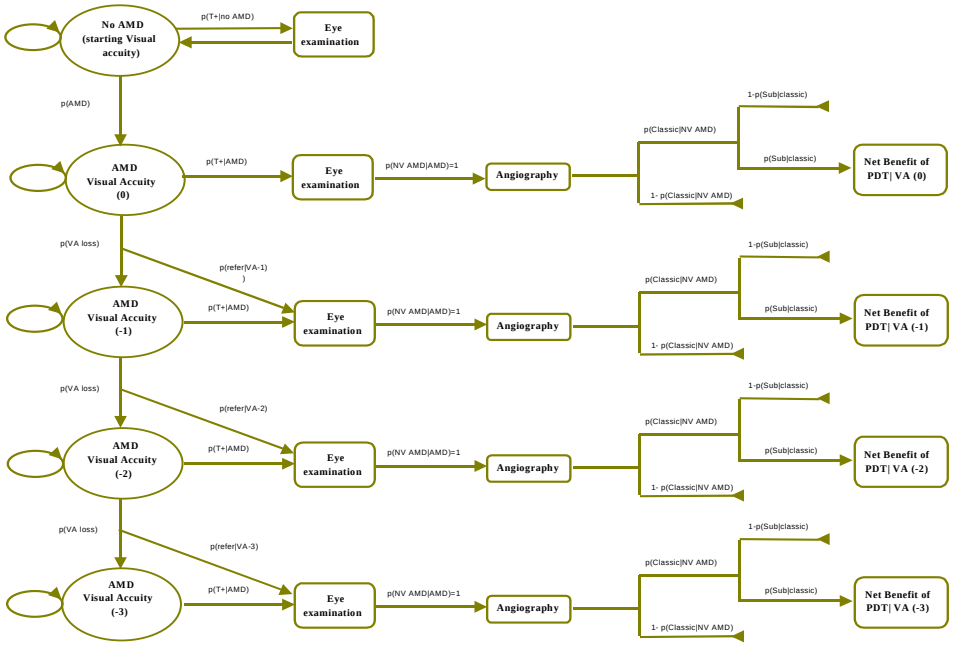
probability distributions[1], full details of which are available elsewhere[6]. Monte Carlo simulation was used to propagate the prior distributions assigned to model inputs and estimate the expected costs and outcomes associated with each alternative therapy.

Alternative structural assumptions

In the base case model the effect of the Amsler grid, in terms of identifying patients with AMD, can occur before AMD develops and at each stage of visual acuity loss. In addition at a loss of 4 or more lines, all patients will have self-referred to the ophthalmologist. However it may be that there is no additional benefit from the Amsler grid after a patient has developed a visual acuity problem (after a loss of one or more lines). It is also possible that all patients will refer to see an ophthalmologist once they have any decline in visual acuity, that is all patients self-refer after a loss of only 1 line. This alternative assumption is plausible given that patients in the model have already had first eye involvement and may therefore be expected to be more vigilant in recognising changes in their vision. The 2 alternative assumptions regarding the effect of the Amsler grid are therefore: the Amsler grid provides no additional benefit, in terms of identifying those patients with AMD after a loss in visual acuity (model 2) and all patients' will self refer after a loss in visual acuity of 1 line (model 3).

There is no evidence to support one way of modelling the screening effect over another and we are therefore uncertain about which is the most appropriate way of modelling. In this application alternative methods were sought to characterise and quantify this structural uncertainty.

Figure 1: AMD model structure



Results

Available methods to characterise structural uncertainty

The review identified 56 papers that looked at methods to characterise and quantify structural uncertainty. Although the searches were limited to key words contained within the title of a paper a number of papers were identified that reviewed current work relating to structural uncertainty and key papers from these reviews were obtained.

As anticipated, very little has been done in the HTA literature to address the issue of structural uncertainty. One method that has been used in HTA, all be it on a limited scale is the use of alternative model scenarios or sensitivity analysis relating to structural assumptions[8, 9]. Although this method can be useful, in as much as it illustrates the impact of structural uncertainties, it can offer little help to decision makers in determining if further research to resolve this uncertainty is worthwhile.

Outside of HTA, in areas such as mathematics, statistics, operational research and environmental modelling, authors have recognised the potential errors arising from model uncertainty and the possibility that model results will be biased when analysts ignore such issues[10-12]. It is in these other disciplines that most of the methods to deal with structural uncertainty have been developed[9].

Many of the methods to assess model performance [11] such as Residual mean squared error (MSE), Finite-prediction-error (FPE) [11] and subjective probabilities[13, 14] are not directly applicable to HTA decision models. These involve the quantification of a statistic that describes goodness of fit or probability of error[5]. In HTA decision modelling, where there are many competing objectives, it is not possible to identify one particular parameter whose performance must be maximised by a fitted model.

Bayesian methods for model averaging appear to be well worked, although not necessarily in the clinical or economic literature. The problem of averaging across models can be viewed in a Bayesian sense as one in which a decision maker needs to make the best possible use of information on a model structure he/she has available[15].

The most widely used method of Bayesian model averaging (BMA) works on the premise that given alternative ways of modelling an intervention effect or other parameter ($\mathrm{M}\kappa$), with Δ as the quantity of interest (such as net benefit), the posterior distribution of Δgiven the data (D) is:

$$\Pr(\Delta|D) = \sum_{K=1}^{K} pr(\Delta|\mathrm{M}\kappa, D)\, pr(\mathrm{M}\kappa|D) = \quad [16]$$

That is Δ is an average of the posterior distributions under each of the models considered, weighted by their posterior model probability[16]. These methods have been applied frequently in the operational research and forecasting literature[17] and more recently when choosing between alternative models to fit survival data[12, 18].

Clearly when applying BMA techniques to HTA decision models there is an issue of determining the posterior distribution of Δ given *D*, when D may not be available or even exist, that is the model may have been developed in the absence of any real information on *D*. The more general techniques of model averaging described in the BMA literature[10], may however be useful. The exact form that this averaging takes is open to interpretation. Previous work looking at the technique has focussed on taking the mean of results of all possible models, weighted or unweighted by the likelihood that particular model specifications are correct. Another possible method is to explicitly address the issue of uncertainty within the model, that is not by calculating separate models and averaging across these, but to characterise this structural uncertainty directly in the model itself.

Three methods that appear to be useful in characterising structural uncertainty in HTA decision model are therefore scenario analysis, parameterising uncertainty directly in models and averaging across potential models using weights or priors. These methods can be applied to HTA decision models as they do not require some measure of prediction performance, which is impossible to calculate in the majority of models where we do not have actual patient level data to assess a models fit against. These 3 methods

are summarised below and then applied to the case study model.

Scenario analysis

Guidelines of good practice in decision analytic modelling, that discuss structural uncertainty, mention the use of scenario analysis to make explicit the effect of structural uncertainty on model results. This analyst is required to compute results for each alternative model specification, and to present alternative results as scenario analyses[8]. The alternative models and their results are then presented to the decision maker. Analysts may choose one particular model as the most likely or conservative and as such alternative scenarios are essentially sensitivity analysis. In this application the 3 alternative models from the case study model were simply ran separately and the results from the models compared.

Weighting the plausibility of the structural assumptions using priors

A possible method, which has not been widely used in the health technology literature, is 'model averaging' where alternative models, with different assumptions, are built, and their results averaged, weighted by the plausibility (prior)[19] of their assumptions usually based on expert opinion. Clearly there is an issue about the validity of using expert opinion to obtain what is essentially another model parameter and this issue has been discussed in depth elsewhere[20]. As described above Bayesian methods can be used to update the posterior of a parameter of interest (Δ) using the dataset to determine performance of a model.

This method is applied to the case study model in its simplest form that is not using any Bayesian methods to update priors or determining a models performance or fit. The alternative models are ran simultaneously using a prior to weight the likelihood of each of the alternative models. Given the lack of evidence regarding the probability that each of the 3 models is the correct model, equal weights have been applied here (models are equally likely to represent the appropriate effect of screening on AMD)

Parameterising the uncertainty explicitly in the model

The impact of any structural uncertainties can be determined in terms of the impact on

the cost-effectiveness results of a model and EVPI. Certain model structures produce a higher EVPI than alternatives, and this is clearly because of additional uncertainties regarding the length of treatment effect have been introduced. Although we can quantify the additional uncertainty and place a valuation on this, in terms of the benefits accrued from further research, we cannot make judgements about the value of conducting further research on those specific uncertain parameters. By explicitly characterising these uncertain parameters in the model we can quantify the increase in decision uncertainty and thus value of research directly, helping to inform future research prioritisation more accurately.

The application of this method in the case study model is undertaken by setting up separate new parameters that represent the parameters we are uncertain about, that is those parameters for which we have ran separate scenarios for. The new parameters can then take values as in the base case model or those from the alternative models (models 2 and 3). The values from the alternative models themselves are still uncertain parameter with distributions. As with the previous methods of handling uncertainty, we are equally likely to believe that each of the structural assumptions is the correct assumption. The probability the base case scenario versus model 2 for the effect of the Amsler grid and base case scenario versus model 3 for probability of self referral, is represented by a uniform distribution between 0 and 1, with values below 0.5 relating to the base case and values above 0.5 relating to the alternative model.

Application of methods to characterise structural uncertainty in case study model

The 3 alternative, relevant, methods were applied to the case-study model of AMD and the results in terms of cost-effectiveness and expected value of perfect information[21] (EVPI) assessed.

Scenario analysis

The results of the 3 models using the alternative structural assumptions are presented in

Tables 1-2 below. Table 1 shows the impact on the cost and effects of the alternative strategies and Table 2 shows the impact on decision uncertainty and EVPI.

Table 1: Costs and outcomes for alternative models

		Mean QALYs	**Mean Costs**	**ICER**
Base case model				
VA = 20/40	Screen and treat	1.2136	£3,651	£12,892
	No screen and treat	0.9836	£2,643	E dominated
	No screen and no treatment	0.9377	£98	-
VA = 20/80	Screen and treat	1.0915	£3,662	£17,757
	No screen and treat	0.9215	£2,644	E dominated
	No screen and no treatment	0.8908	£98	-
No additional effect of the Amsler grid after a loss of visual acuity				
VA = 20/40	Screen and treat	1.2128	£3,650	£12,828
	No screen and treat	0.9816	£2,637	E dominated
	No screen and no treatment	0.9359	£98	-
VA = 20/80	Screen and treat	1.0885	£3,653	£17,855
	No screen and treat	0.9202	£2,640	E dominated
	No screen and no treatment	0.8894	£98	-
All patients' will self refer after a loss in visual acuity				
VA = 20/40	Screen and treat	1.2183	£3,663	£16, 178
	No screen and treat	1.0648	£2,743	E dominated
	No screen and no treatment	0.9980	£99	-
VA = 20/80	Screen and treat	1.0908	£3,652	£23,162
	No screen and treat	0.9737	£2,739	E dominated
	No screen and no treatment	0.9374	£99	-

Table 2: Simulation results (£30,000 threshold value) for alternative models

		Probability cost-effective	**EVPI assuming 10-year lifetime**
Base case model			
VA = 20/40	Screen and treat	0.89	£6,180,538
	No screen and treat	0.02	
	No screen and no treatment	0.09	
VA	Screen and treat	0.73	£15,326,856
	No screen and treat	0.03	

	No screen and no treatment	0.24	
No additional effect of the Amsler grid after a loss of visual acuity			
VA= 20/40	Screen and treat	0.88	£6,686,984
	No screen and treat	0.02	
	No screen and no treatment	0.10	
VA = 20/80	Screen and treat	0.72	£15,865,519
	No screen and treat	0.03	
	No screen and no treatment	0.25	
All patients' will self refer after a loss in visual acuity			
VA= 20/40	Screen and treat	0.66	£30,466,013
	No screen and treat	0.12	
	No screen and no treatment	0.22	
VA = 20/80	Screen and treat	0.52	£40,179,821
	No screen and treat	0.09	
	No screen and no treatment	0.39	

For both starting visual acuities changing the assumption of the additive effect of the Amsler grid had little effect on the costs and QALYs. Screening is still regarded as cost-effective when compared to no treatment. This is because the majority of patients are diagnosed through the self-screen before any loss in visual acuity. Decision uncertainty is however sensitive to structural assumptions, with the probability that screen + treat is cost effective reduced from 0.89 in the base case 20/40 model to 0.66 when assuming all patients will self refer after a loss in visual acuity. The population EVPI for these alternative assumptions increases for both the 20/40 and 20/80 starting visual acuity models, but this is most marked in the model assuming all patients will self refer after a loss in visual acuity, increasing the EVPI from £6 million to £30 million in the 20/40 model and from £15 million to £40 million in the 20/80 model.

The value associated with specific parameters or groups of parameters for all 3 scenarios were also calculated. In the base case model for patients with a starting visual acuity of 20/40, the value of information associated with the expected QALYs with or without PDT is £3.73 million. The other groups of model inputs, such as screening accuracy, have no value of information associated with them. At a starting visual acuity of 20/80

the value of information associated with the expected QALY with or without PDT is £4.67 million and the value associated with the progression of visual acuity is just below £194,000. The other groups of model inputs have no value of information associated with them. A similar pattern emerged from 2 alternative models, that is the greatest value of information was associated with the parameters looking at QALYs with or without PDT. Unlike the base case model the alternative models showed a quantifiable value of information associated with natural history parameters (rate of progression, eligibility for PDH, probability of disease) for 20/80 starting visual acuity. This value was relatively low in model 2 at £246,465, but in model 3 the value associated with this group of parameters constituted a large percentage of the total EVPI, at over £13 million.

Although the illustration of various scenarios representing structural uncertainty can offer the decision maker with an opportunity to decide on which assumption he/she believes and make policy decisions on that basis, clearly what it does not do is provide any explicit framework for quantifying this uncertainty or offer any guidance to decision makers that have no clear preferences over alternative model assumptions. This application did not show any value associated with any of the uncertain parameter, but given that these were ran as separate scenarios, as opposed to modelling the full range of possible values from each scenario in a single model, it is still not clear if resolving the structural uncertainty is worthwhile in terms of the payback from further research.

Weighting the plausibility of the structural assumptions using priors

The choice of prior or weight given to each scenario in the case-study model could not be informed by any expert opinion, as the uncertainties essentially represent different clinical opinion about the likely effect of a screening programme for AMD.

The results from weighting the alternative scenarios (base case model and models 2 and 3) are presented in Tables 3 and 4 below. Table 3 shows the impact on the cost and effects of the alternative strategies and Table 4 shows the impact on decision uncertainty and EVPI.

Table 3: Costs and effects for weighting alternative structural assumptions

		Mean QALYs	Mean Costs	ICER
VA = 20/40	Screen and treat	1.2149	£3655	£13,802
	No screen and treat	1.01	£2674	E dominated
	No screen and no treatment	0.9572	£98	
VA = 20/80	Screen and treat	1.0902	£3656	£19,290
	No screen and treat	0.9385	£2674	E dominated
	No screen and no treatment	0.9058	£99	

Table 4: Simulation results (£30,000 threshold) for weighting alternative structural assumptions

		Probability cost-effective	EVPI assuming 10-year lifetime
VA = 20/40	Screen and treat	0.81	£14,444,511
	No screen and treat	0.05	
	No screen and no treatment	0.14	
VA = 20/80	Screen and treat	0.66	£23,790.732
	No screen and treat	0.05	
	No screen and no treatment	0.29	

Model averaging does not change the decision, that is screening is still regarded as cost-effective when compared to no treatment. Averaging over the 3 possible scenarios shows that we are fairly certain about screening being cost effective in the 20/40 and 20/80 visual acuity models. Decision uncertainty is however still sensitive to structural assumptions, with the probability that screen + treat is cost effective reduced from 0.89 in the base case 20/40 model to 0.81 and reduced from 0.73 in the 20/80 base case model to 0.66, when averaging across the 3 scenarios. It is therefore not surprising that EVPI increases in the 20/40 and 20/80 models as compared with the base case model to over £14 million and nearly £24 million respectively. Partial EVPI analysis gave very similar results to the base case model, in that the greatest value was associated with parameters relating to QALYs with or without PDT.

Like the previous application simple model averaging did help to illustrate the impact of structural uncertainties, in this case they had little impact on the adoption decision. But,

as with the previous method model averaging does not quantify this uncertainty explicitly so that recommendations can me made about further research to inform this uncertainty.

Parameterising the uncertainty directly in the model

The results from the model explicitly characterising the uncertainty regarding the length of effect of the Amsler grid and the probability that patients will self refer due to a loss in visual acuity, are presented in Tables 5 and 6 below. Table 5 shows the impact on the cost and effects of the alternative strategies and Table 6 shows the impact on decision uncertainty and EVPI.

Table 5: Costs and effects for model parameterising structural uncertainty

		Mean QALYs	**Mean Costs**	**ICER**
VA = 20/40	Screen and treat	1.2178	3.661	£14,150
	No screen and treat	1.0219	2,687	E.dominated
	No screen and no treatment	0.9660	98	-
VA = 20/80	Screen and treat	1.0904	£3,659	£19,972
	No screen and treat	0.9451	£2,658	E.dominated
	No screen and no treatment	0.9121	£98	-

Table 6: Simulation results (£30,000 threshold) for model parameterising structural uncertainty

		Probability cost-effective	**EVPI assuming 10-year lifetime**
VA = 20/40	Screen and treat	0.78	£19,748,032
	No screen and treat	0.07	
	No screen and no treatment	0.15	
VA = 20/80	Screen and treat	0.63	£28,042,157
	No screen and treat	0.06	
	No screen and no treatment	0.31	

The results show that by including specific parameters to represent uncertainty about the effect of the Amsler grid and the probabilities of self-referral the ICERs for the 20/40 and 20/80 models are slightly higher than for the base case model (£14,150 compared to

£12,892 and £19,972 compared to £17,757). The adoption decision remains unchanged, that is screening still appears to be cost-effective. No surprisingly the EVPI for the 20/40 and 20/80 models is approximately the average of the EVPI values from the 3 scenarios (base case model and models 2 and 3), at £19,748,032 and £28,042,157 respectively. Most interesting is the value of information associated with particular parameters of groups of parameters. As with the previous 2 methods of characterising structural uncertainty, value of information associated with QALYs with and without PDT is the most significant (in terms of value) set of parameters for the 20/40 and 20/80 starting visual acuities. In addition to the partial EVPI analysis undertaken for the previous 2 methods, partial EVPI was calculated for the 2 'new' parameters, representing the uncertain structural assumptions. Although these 'new' sets of parameters were not associated with any value this result is itself useful. Decision makers when presented with the alternative scenarios may have concluded that additional research needed to be commissioned in order to inform this uncertainty. In this example parameters relating the self-referral and diagnostic accuracy have so little impact on decision uncertainty that it is not worthwhile conducting further research to resolve the issues.

Discussion

Decision-analytic models represent an explicit way to synthesise evidence currently available on the outcomes and costs of alternative health care interventions[22, 23] and are, therefore, a powerful tool for decision making. The results derived from a decision-analytic model will depend on how the model structure has been defined and the data used to populate the model. Recent developments in the analysis of uncertainty in HTA models have typically focused on parameter uncertainty[24, 25] which show the probability of one intervention being more cost-effective than one or more mutually exclusive alternatives, taking account of any uncertainties in the data inputs. Such analyses are, of course, based on the premise that the model has been correctly specified. An inappropriate model structure can potentially invalidate estimates of cost-effectiveness and estimates of the value of further research. An inappropriate model is also of little value to a decision maker. Although it is common practice to acknowledge

potential limitations in model structure, there is a lack of clarity about methods to evaluate structural uncertainties, which is arguably the most important source of uncertainty. Analysts need to address the question as to how model uncertainty will affect forecast accuracy[17].

Many of the methods to look at structural uncertainty, explored in other disciplines such as mathematics and statistics, are not directly applicable to decision modeling. This is primarily because of their focus on prediction performance. This is a criteria not commonly used to assess decision analytic models because of the difficulty of quantifying such a statistic when data to measure goodness of fit is often not available.

Methods, which appear to be relevant in quantifying structural uncertainty in decision analytic models, are, the use of scenarios exploring different structural assumptions, model averaging techniques and the quantification of the structural uncertainty as 'uncertain' parameters directly in the model. These methods to characterise structural uncertainty in the AMD model were explored using 3 alternative structural assumptions: 1) The Amsler grid has an effect even after a loss in visual acuity and patients only self-refer to an ophthalmologist following a decline in visual acuity 2) The Amsler grid provides no additional benefit, in terms of identifying those patients with AMD after a loss in visual acuity 3) All patients' will self refer after any loss in visual acuity.

Although the cost-effectiveness results were not sensitive to structural assumptions EVPI was particularly sensitive to model specification. When looking at the alternative structural assumptions as scenarios the more conservative scenario of assuming all patients self refer after a loss of 1 line, that is there is very little effect of the Amsler grid after a loss in visual acuity, EVPI increases quite substantially compared to the base case analysis (£6 million to £30 million in the 20/40 model). The explicit modeling of these parameters, although of little actual value (in terms of further research) in this example, did illustrate that the uncertain parameters were of so little importance in terms of the adoption decision (screening was the most cost-effective strategy regardless of what values these parameters took) that conducting further research, would not be cost-

effective. Further research relating to this model should therefore focus on the outcomes with or without photodynamic therapy.

Conclusions

Appropriately characterizing structural uncertainty clearly has an impact on a model results and in particular the expected value of research estimated by the model. As a result of increasing the decision uncertainty in each of the alternative models, the value of research increased, in particular for the 20/80 starting visual acuity model. It is therefore it is crucial to determine the most appropriate model structure when using VOI analysis to inform research prioritisation.

References

1. Briggs AH, *Handling uncertainty in cost-effectiveness models.* Pharmacoeconomics, 2000. **17**(5): p. 479-500.
2. Hay, J. and J. Jackson, *Panel 2: methodological issues in conducting pharmacoeconomic evaluations - modeling studies.* Value in Health, 1999. **2**(2): p. 78-81.
3. ISPOR task force, *Principles of good practice for decision analytic modeling in health care evaluation. Draft document,*. 2001.
4. Ramsey, S.D., *Evaluating evidence from a decision analysis.* Journal of the American Board of Family Practice., 1999. **12**(5): p. 395-402.
5. Bates JM and Granger CWJ, *The combination of forecasts.* Operational research quarterly, 1969. **20**: p. 451-468.
6. Claxton K, et al., *A pilot study on the use of decision theory and value of information analysis as part of the NHS health technology assessment programme.* 2004, Health Technology Assessment, Vol 8, No 31.
7. Briggs AH and Sculpher MJ, *An introduction to markov modelling for economic evaluation.* Pharmacoeconomics, 1998. **13**(4): p. 397-409.
8. McCabe, C. and S. Dixon, *Testing the validity of cost-effectiveness models.* Pharmacoeconomics., 2000. **17**(5): p. 501-13.
9. McKay MD, M.J., Upton SC, Wendelberger JR, and Morrison JD, *Model uncertainty in stochastic simulation models. Contributions to the development of a statistical framework for joint mobility modelling*. 1997, Los Alamos National Laboratory: Los Alamos.
10. Chatfield C, *Model uncertainty, data mining and statistical inference (with discussion).* Journal of the Royal Statistical Society, 1995. **158**: p. 419-466.
11. Engle, R.F. and S.J. Brown, *Model Selection for Forecasting.* Applied Mathematics and Computation, 1986. **20**(3-4): p. 313-327.
12. Augustin N, Sauerbrei W, and Schumacher M, *The practical utility of incorporating model selection uncertainty into prognostic models for survival data*. 2004: Glasgow.
13. Bunn, D.W. and E. Kappos, *Synthesis or Selection of Forecasting Models.* European Journal of Operational Research, 1982. **9**(2): p. 173-180.
14. Bunn DW, *Combining forecasts.* European journal of operational research, 1988. **33**: p. 223-229.
15. Clemen RT, *Combining forecasts: a review and annotated bibliography.* International Journal of Forecasting, 1989. **5**: p. 559-583.
16. Hoeting JA, et al., *Bayesian model averaging: a tutorial.* Statistical Science, 1999. **14**(4): p. 382-417.
17. Chatfield C, *Model uncertainty and forecast accuracy.* Journal of forecasting, 1996. **15**: p. 495-508.
18. Volinsky CT, et al., *Bayesian model averaging in proportional hazard models: Assessing the risk of a stroke.* Journal of the Royal Statistical Society, Series C, 1997. **46**(4): p. 433-448.

19. Scott M, *Uncertain models and modelling uncertainty.* 2004, Presentation at Department of Environmantal Modelling. University of Nottingham.
20. Morris PA, *Decision analysis expert use.* Management science, 1974. **20**: p. 1233-1241.
21. Claxton K, *The irrelevance of inference: a decision making approach to the stochastic evaluation of health care technologies.* Journal of Health Economics, 1999. **18**: p. 341-364.
22. Gold, M.R., et al., *Cost-Effectiveness in Health and Medicine*. 1996, New York: Oxford University Press.
23. Drummond, M.F., et al., *Methods for the Economic Evaluation of Health Care Programmes*. 2nd edition ed. 1997, New York: Oxford University Press.
24. Briggs AH and Gray AM, *Handling Uncertainty in Economic Evaluations of Health Care Interventions.* British Medical Journal, 1999. **319**: p. 635-638.
25. Briggs AH, *A Bayesian Approach to Stochastic Cost-effectiveness Analysis.* Health Economics, 1999. **8**: p. 257-261.