What happens to value of information measures as the number of decision options increases?

Pelham Barton
Department of Health Economics
University of Birmingham
B15 2TT
p.m.barton@bham.ac.uk

Background: Probabilistic sensitivity analysis, including value of information analysis, is well-developed as a means of expressing uncertainty when there is a limited number of clearly distinct decision options. However, realistic medical decision making often involves choice between a very wide range of options. The probability that the preferred intervention is cost-effective, and the Expected Value of Perfect Information (EVPI), are both highly sensitive to the choice of allowed options in the model. Analysis based on an essentially arbitrary selection of options cannot reflect the full decision uncertainty.

Aim: To explore what happens to the EVPI as an increasingly large number of decision options is considered.

Methods: A stylised model based on the notion that many men who have prostate cancer die of other causes without the cancer itself becoming apparent. In the model, a cancer may develop and remain asymptomatic and harmless for a long time. If it progresses, it is assumed to be instantly fatal. A screening test is available, as is a treatment which restores normal life expectancy but reduces quality of life and requires lifelong maintenance therapy. The model was to test different possible ages for "once in a lifetime" screening under a variety of assumptions about model parameter uncertainty. Initial model runs only considered screening at ages which are multiples of 16 years, then multiples of 8 years were allowed, then 4 years, etc.

Results: Effects of more refined choice sets on the Cost-Effectiveness Acceptability Frontier (CEAF) and EVPI are shown. The CEAF collapses to zero for threshold Incremental Cost-Effectiveness Ratios at which screening at some age is preferred to no screening, while the EVPI stabilises at a non-zero figure.

Conclusions: The Cost-Effectiveness Acceptability Frontier is of limited use when the options modelled are an essentially arbitrary selection from a very large set of possible options. In contrast, the Expected Value of Perfect Information can be reasonably approximated by a model with a limited choice set. It should, however, be noted that a model with a very restricted choice set may overestimate the total uncertainty measured through the EVPI. This can happen when the limited model does not include an option close to the true optimal choice under current uncertainty that would be found from the "perfect" model including all choices.

Introduction

The Cost-Effectiveness Acceptability Curve was introduced by van Hout and colleagues (1994) in the context of clinical trials in a frequentist paradigm, to cope with the problem that confidence intervals for the Incremental Cost-Effectiveness Ratio (ICER) do not make sense outside the first (or north-east) quadrant of the cost-effectiveness plane.

It is simple and helpful to adapt the CEAC to model-based analysis in a Bayesian paradigm comparing two options. Input parameters to the model are sampled from their joint distribution, and the proportion of replications favouring a particular option can be plotted as a function of the threshold ICER. This proportion is often described as the probability that the option will be cost-effective, under the assumption that all the uncertainty is correctly reflected in the model parameters.

In a Bayesian paradigm, the optimal decision is that indicated by the mean of the output distribution on the cost-effectiveness plane (Claxton, 1999). This is not necessarily the same as the option with the higher probability (see, for example, Fenwick *et al*, 2001). Indeed, it is simple to construct examples where option X dominates option Y in the mean results, but the probability shown for option X at any threshold ICER is arbitrarily low (see Appendix 1).

Fenwick and colleagues (2001) proposed the notion of the Cost-Effectiveness Acceptability Frontier (CEAF) to remove the potentially misleading effect of showing a probability of over 50 percent for an option which is not preferred. At any given threshold ICER, the preferred option on the basis of mean results is determined, and the proportion of replications favouring this option is plotted. This usually involves a discontinuity at the ICER between the two options.

The definitions of the CEAC and CEAF in terms of proportion of model replications favouring a given option extend readily to decision problems involving three or more options. However, just because it is simple to make such a definition does not mean that it is sensible to do so. The probability that a given option is cost-effective is a function of the choice of options available. Again, it is easy to construct examples (comparing options X, Y, and Z, say) in which:

- for each pair of options, the one with a higher probability of cost-effectiveness in a bilateral comparison has a lower probability of cost-effectiveness in the three-way comparison;
- option Z dominates options X and Y in the mean results, but has zero probability of being cost-effective with arbitrarily small (but nonzero) skewness in the output distributions;
- option Z dominates options X and Y in the mean results, but has arbitrarily low probability of being cost-effective, with symmetrical output distributions.

(See Appendix 1 for examples of all of these.)

These phenomena result in part from the fact that multi-way CEACs are based on the idea that one option is optimal, and all the others are effectively treated as equally bad. To overcome this problem, it is necessary to use a measure which incorporates the magnitude of the difference between “correct”

and “wrong” decisions, rather than merely counting probabilities. A measure that seems appropriate for this purpose is the Expected Value of Perfect Information (EVPI).

The EVPI is a function of the threshold ICER, and is the difference between the expected net monetary benefit for the optimal decision before and after resolving the uncertainty.

The purpose of this work is to explore what can happen to the CEAF and EVPI as the number of options in the decision problem increases. This is done in the context of a stylised model for a screening programme for a hypothetical cancer. It is assumed that screening can only be offered once in a person’s lifetime, and the decision question is to find the optimal age for such screening.

The model

The model used in this paper is based on the notion that roughly as many men die with prostate cancer as die of it. In this model, a cancer develops first in a form that is asymptomatic, detectable, and treatable, but becomes immediately fatal if it progresses. A simple screening test is available at a price of £50 with sensitivity and specificity less than 100 percent. Anyone testing positive is assumed to take a confirmatory test costing £1,000. The confirmatory test is taken to be 100 percent sensitive and specific. Anyone with the asymptomatic cancer will be treated surgically at a cost of £50,000. Treatment is assumed to prevent death due to the cancer but require maintenance therapy costing £500 per year and lead to reduced quality of life. Non-cost parameters are assumed to be known to within a range of uncertainty: these are described next.

Table 1. Parameters for time to event

Parameter		mean	s.d.
Age at death from other causes	α	8	1.5
	β	75	5
Minimum age of onset of asymptomatic cancer		30	4
Time from minimum age to actual onset of asymptomatic cancer	α	2	0.2
	β	122	20
Time from onset of asymptomatic cancer to progression of cancer	α	2	0.2
	β	22	4

Variable times to event are taken from Weibull distributions. The formulation used is in terms of a scale parameter α and a shape parameter β, defined so that a variable X follows a Weibull distribution with those parameters if $\left(\frac{X}{\beta}\right)^{\alpha}$ has an exponential distribution with unit mean. In all cases used here, $\alpha > 1$, which means an increasing risk over time. For fixed α, the mean of the distribution is proportional to β. In all cases, the parameters are sampled from independent normal distributions as shown in Table 1. These parameters have been selected so that, in the natural history model,

approximately 5 percent of the population dies from cancer, while just over 7 percent die with the cancer in its asymptomatic form.

Age at death from other causes is "capped" at 127.9 (higher ages are reduced to this figure): this allows a "no screening" option to be coded as screening at age 128.

Sensitivity and specificity of the screening test were set as bivariate normal on the logit scale with a negative correlation between them. Each was given a median of 0.98 and lower 95 percent limit of 0.97, with a correlation coefficient of -0.2 between the underlying normal distributions. For convenience, a similar univariate distribution was used for quality of life post-transplant, with a median of 0.9 and lower 95 percent limit 0.85.

Running the model

It is convenient to describe the running of the model for a particular sampled set of parameters. The ones shown below are the first sampling from the seeded random number generator.

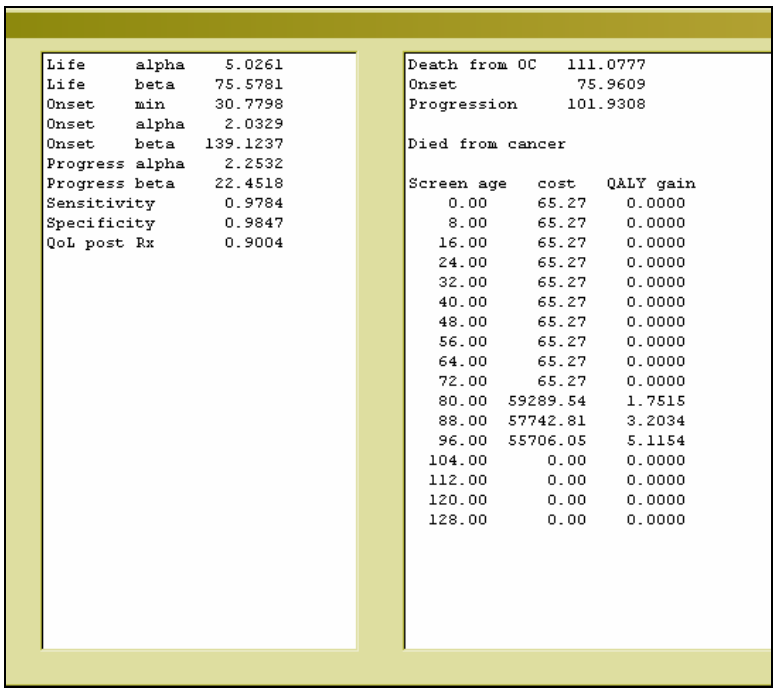
The natural history part of the model is a continuous time individual sampling model (Barton *et al*, 2004; Brennan *et al*, 2006). In the interests of variance reduction, the expected outcomes from screening at any age are calculated exactly for each sampled individual. Each individual is completely described by three ages, representing death from other causes, onset of asymptomatic cancer, and progression to fatal cancer (if untreated). There are three possible cases, which are illustrated in Figures 1 to 3. The sampled values of the non-cost parameters are shown in the left-hand panel of each figure, with the individual's natural history shown at the top of the right-hand panel, with the expected effects of screening below. The detail of these expected effects follows.

The individual shown in Figure 1 would develop asymptomatic cancer just before reaching age 76, and this would progress at age just below 102. If he is screened before age 76, there is just over a 1.5 percent chance of a false positive result requiring a confirmatory test, so the expected cost for screening at such ages is £50 for the original screening test plus just over $0.015 \times £1000 = £15$ for the confirmatory test if needed. Screening below age 76 thus leads to a cost of just over £65 with zero QALY gain in this case.

If the individual in Figure 1 is screened at age 80, there is a nearly 98 percent chance of a true positive result from screening. Without screening he would expect to live nearly 22 years in (apparently) full health, securing $\frac{1-\exp(-21.94\lambda)}{\lambda} = 15.40$ QALYs, where $\lambda = \ln(1.035)$, representing a discount rate of 3.5 percent. If screened and treated, he would live for just over 31 years with quality of life score of just over 0.9, securing $0.9004 \times \frac{1-\exp(-31.08\lambda)}{\lambda} = 17.19$ QALYs. There is thus a QALY gain for successful treatment of 1.79, so an expected QALY gain for screening of 1.75, taking the sensitivity of

the test into account. Similar calculations lead to the expected cost of screening for this individual, where the cost of the original screening test is applied in full, but costs of confirmatory test, surgery, and maintenance treatment (also discounted at 3.5 percent), are multiplied by the sensitivity of the test. Full details of the calculation, and justification for the discounting formula, are shown in Appendix 2.

Figure 1. Example of early death prevented through screening



Life	alpha	5.0261
Life	beta	75.5781
Onset	min	30.7798
Onset	alpha	2.0329
Onset	beta	139.1237
Progress	alpha	2.2532
Progress	beta	22.4518
Sensitivity		0.9784
Specificity		0.9847
QoL post Rx		0.9004

Death from OC	111.0777
Onset	75.9609
Progression	101.9308

Died from cancer

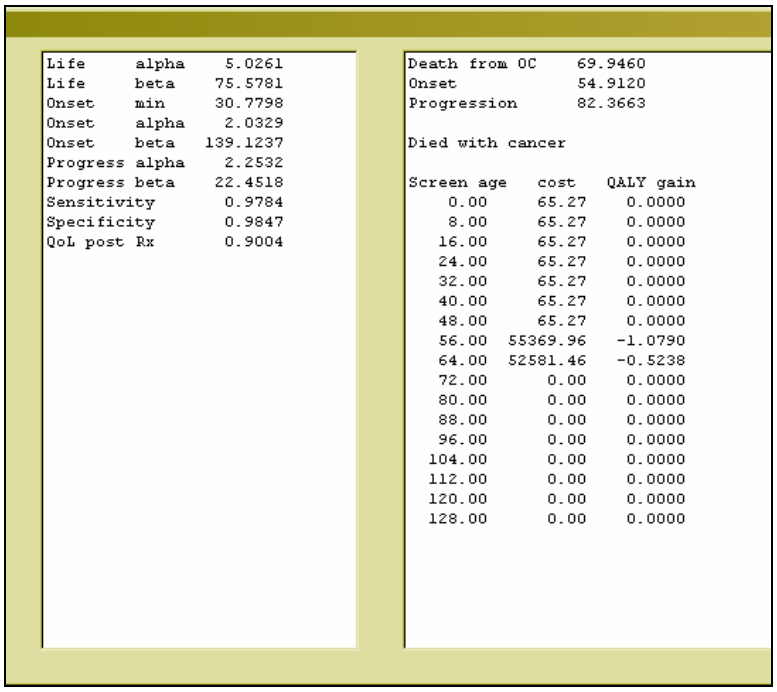
Screen age	cost	QALY gain
0.00	65.27	0.0000
8.00	65.27	0.0000
16.00	65.27	0.0000
24.00	65.27	0.0000
32.00	65.27	0.0000
40.00	65.27	0.0000
48.00	65.27	0.0000
56.00	65.27	0.0000
64.00	65.27	0.0000
72.00	65.27	0.0000
80.00	59289.54	1.7515
88.00	57742.81	3.2034
96.00	55706.05	5.1154
104.00	0.00	0.0000
112.00	0.00	0.0000
120.00	0.00	0.0000
128.00	0.00	0.0000

Screening at age 88 or 96 would lead in this case to slightly reduced costs (because maintenance treatment would be given for less time) and increased QALY gains (because the reduced quality of life post-treatment applies for a shorter time).

If the screening age were set at 102 or higher, the individual shown in Figure 1 would not survive to be screened, so there would be zero cost and zero QALY gain.

Figure 2 shows an individual who, if untreated, would die from other causes with the cancer in the asymptomatic stage. The effect of screening while asymptomatic is thus to incur the cost of treatment, and there is a QALY loss due to reduced quality of life post-treatment. Finally the vast majority of individuals follow the pattern shown in Figure 3, where the cancer never develops even in its asymptomatic form. (In this case, the age of progression of cancer is redundant.)

Figure 2. Example of quality of life loss due to treatment



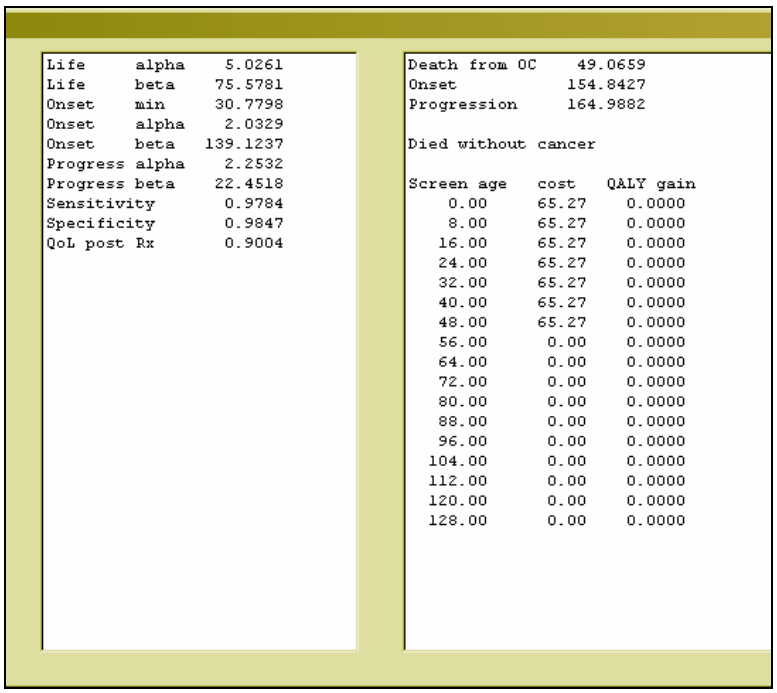
Life	alpha	5.0261
Life	beta	75.5781
Onset	min	30.7798
Onset	alpha	2.0329
Onset	beta	139.1237
Progress	alpha	2.2532
Progress	beta	22.4518
Sensitivity		0.9784
Specificity		0.9847
QoL post Rx		0.9004

Death from OC	69.9460
Onset	54.9120
Progression	82.3663

Died with cancer

Screen age	cost	QALY gain
0.00	65.27	0.0000
8.00	65.27	0.0000
16.00	65.27	0.0000
24.00	65.27	0.0000
32.00	65.27	0.0000
40.00	65.27	0.0000
48.00	65.27	0.0000
56.00	55369.96	-1.0790
64.00	52581.46	-0.5238
72.00	0.00	0.0000
80.00	0.00	0.0000
88.00	0.00	0.0000
96.00	0.00	0.0000
104.00	0.00	0.0000
112.00	0.00	0.0000
120.00	0.00	0.0000
128.00	0.00	0.0000

Figure 3. Example of individual who never develops the cancer



Life	alpha	5.0261
Life	beta	75.5781
Onset	min	30.7798
Onset	alpha	2.0329
Onset	beta	139.1237
Progress	alpha	2.2532
Progress	beta	22.4518
Sensitivity		0.9784
Specificity		0.9847
QoL post Rx		0.9004

Death from OC	49.0659
Onset	154.8427
Progression	164.9882

Died without cancer

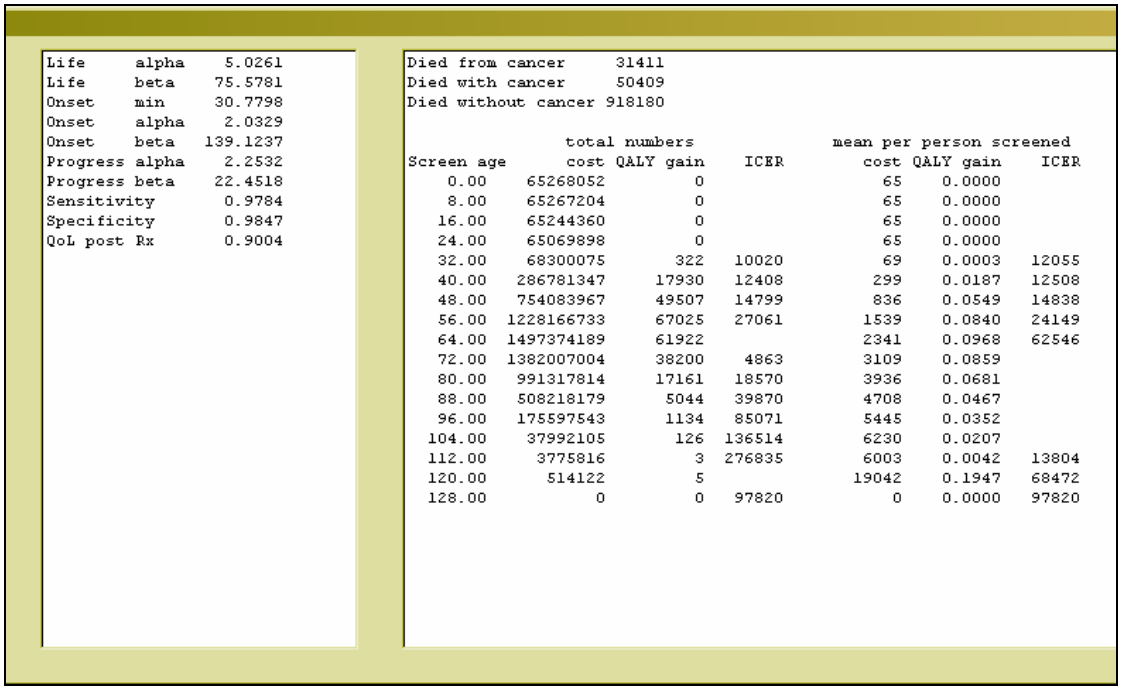
Screen age	cost	QALY gain
0.00	65.27	0.0000
8.00	65.27	0.0000
16.00	65.27	0.0000
24.00	65.27	0.0000
32.00	65.27	0.0000
40.00	65.27	0.0000
48.00	65.27	0.0000
56.00	0.00	0.0000
64.00	0.00	0.0000
72.00	0.00	0.0000
80.00	0.00	0.0000
88.00	0.00	0.0000
96.00	0.00	0.0000
104.00	0.00	0.0000
112.00	0.00	0.0000
120.00	0.00	0.0000
128.00	0.00	0.0000

Population Effects

Figure 4 shows the effects of sampling 1,000,000 individuals using the same set of parameter values as in Figures 1-3. The top of the right-hand panel gives the results for the natural history, with no screening. For these parameters, just over 3 percent of the population die from cancer, while just over 5 percent die with the cancer in its asymptomatic stage. The next figures to note are the total costs and

QALYs for applying screening at any particular age, with apologies for the absurdly large number of significant figures given in the cost column in particular.

Figure 4. Population effects for a single set of parameter values.



Life	alpha	5.0261
Life	beta	75.5781
Onset	min	30.7798
Onset	alpha	2.0329
Onset	beta	139.1237
Progress	alpha	2.2532
Progress	beta	22.4518
Sensitivity		0.9784
Specificity		0.9847
QoL post Rx		0.9004

Died from cancer 31411
Died with cancer 50409
Died without cancer 918180

	total numbers			mean per person screened		
Screen age	cost	QALY gain	ICER	cost	QALY gain	ICER
0.00	65268052	0		65	0.0000	
8.00	65267204	0		65	0.0000	
16.00	65244360	0		65	0.0000	
24.00	65069898	0		65	0.0000	
32.00	68300075	322	10020	69	0.0003	12055
40.00	286781347	17930	12408	299	0.0187	12508
48.00	754083967	49507	14799	836	0.0549	14838
56.00	1228166733	67025	27061	1539	0.0840	24149
64.00	1497374189	61922		2341	0.0968	62546
72.00	1382007004	38200	4863	3109	0.0859	
80.00	991317814	17161	18570	3936	0.0681	
88.00	508218179	5044	39870	4708	0.0467	
96.00	175597543	1134	85071	5445	0.0352	
104.00	37992105	126	136514	6230	0.0207	
112.00	3775816	3	276835	6003	0.0042	13804
120.00	514122	5		19042	0.1947	68472
128.00	0	0	97820	0	0.0000	97820

When the population totals are converted to means per person screened, the denominator varies according to screening age. This is because not all of the original 1,000,000 population survive to be screened. For example, the figures at age 56 represent screening 798,000 people, while those at age 64 represent screening only 640,000 people.

This raises a methodological issue that is completely separate from the main point of this paper, and which is revisited in the discussion. For now, comparisons between different ages will be made both on the basis of total figures and mean per person screened.

ICERs are shown where there is no dominance between options. Note that the most effective option in terms of total population is screening at age 56, which dominates screening at age 64 using that method of calculation, while the most effective based on mean effects per person screened is screening at age 64, although this option is only cost-effective at a threshold ICER in excess of £60,000 per QALY.

In practice, the optimal screening age for any given threshold ICER was calculated using the net monetary benefit approach, using thresholds from 0 to £60,000 per QALY in multiples of £100. As noted above, a “no screening” option is included, coded as screening at age 128.

When this approach was used for the single parameter set shown in Figure 4, the results were as shown in Table 2. Ranges in the body of the table are the range of threshold ICERs (in £/QALY) at which the

given screening age is optimal. For example, at a threshold of £25,000/QALY, the optimal screening age from the choices tried is 48 based on total values, but 56 based on mean values.

Table 2. Results for a single parameter set

Optimal screening age	Using total values	Using mean values
No screening	0 to 15,200	0 to 15,200
48	15,300 to 27,000	15,300 to 24,100
56	27,100 to 60,000	24,200 to 60,000

See text above for full explanation of the table.

Results from the full model

In running the full model, 5000 parameter sets were sampled from the distributions described earlier. For each parameter set, 100,000 individual natural histories were sampled, and the expected effect of screening at various ages calculated. These were then totalled and averaged, so that the optimal strategy for each individual parameter set could be determined, as well as the optimal strategy allowing for uncertainty and hence the CEAF and EVPI could be generated. Different random number generators were used for generating parameters and patient histories.

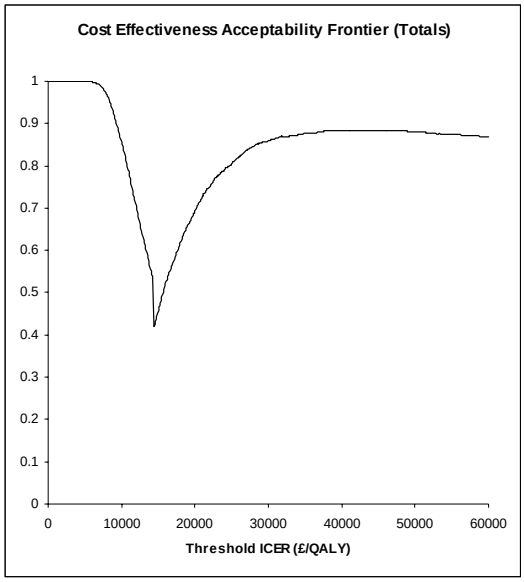
Figure 5 shows the CEAF and EVPI for the model when the only options considered for the screening age were multiples of 16 years. In this case, the only options that are ever optimal (at threshold ICERs up to £60,000/QALY) on current uncertainty are no screening and screening at age 48. The ICER between these options is approximately £14,500/QALY (varying slightly depending on whether calculations were based on totals or means).

Considering first the two CEAFs, in each case at a very low threshold ICER, the no screening option is optimal with certainty, but as the threshold increases, there is a measurable probability that screening at some age will be preferred. When the ICER between no screening and screening at age 48 is reached, there is a discontinuity in the CEAF. (This is quite usual: see, for example, Fenwick *et al*, 2001). As the threshold increases, the probability that screening at age 48 is optimal increases to a maximum, but then starts to decrease. This appears to be because screening at higher ages has a greater probability of being optimal at higher threshold ICERs.

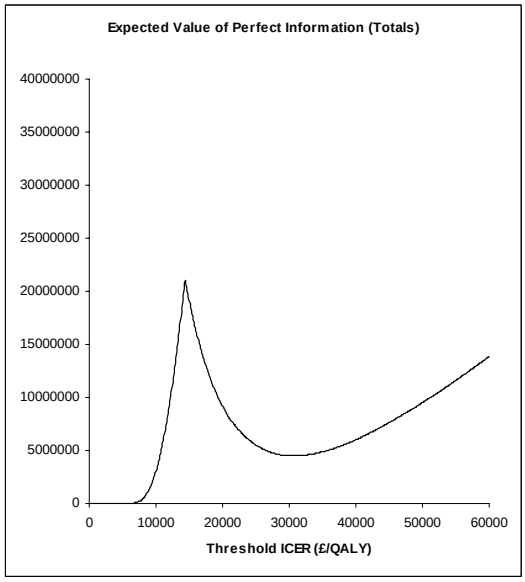
Now consider the two EVPI curves. An unimportant difference between them is that Figure 5(*b*) is based on total values while Figure 5(*d*) is based on means, and hence the values given differ by several orders of magnitude. For the absolute value of EVPI to be meaningful, it should be presented in terms of the size of the population to whom the decision will be applied. This effectively means multiplying by a constant representing the size of the population and time for which the decision will be applicable. For the purpose of the present paper, it is sufficient to note that the various EVPI curves for different choice sets are comparable with each other.

Figure 5 CEAF and EVPI when options for screening age are multiples of 16 years

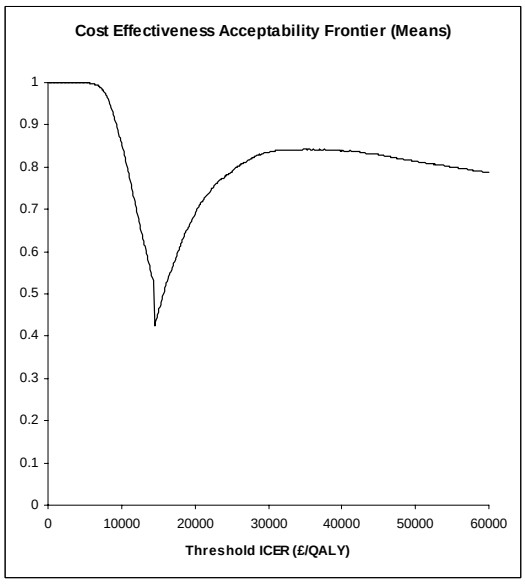
(*a*) CEAF based on total results



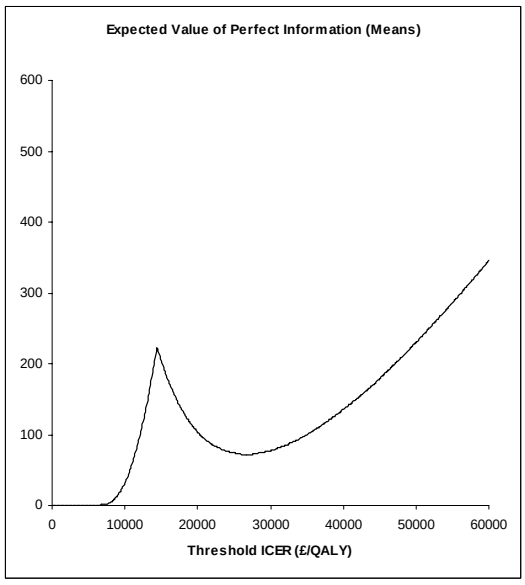
(*b*) EVPI based on total results



(*c*) CEAF based on mean results



(*d*) EVPI based on mean results

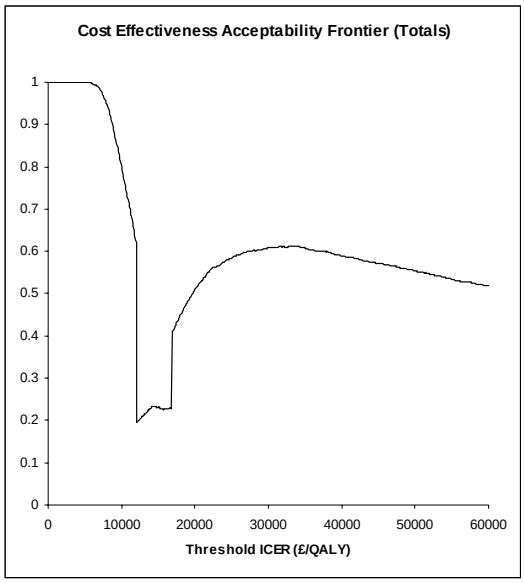


The curves themselves form a shape that is typical of EVPI curves in the case where there is uncertainty as to the most effective option, but not the least costly option. When the threshold ICER is very low, there is no doubt that no screening is the preferred option, so the value of information here is zero. The EVPI rises to a sharp point where the threshold ICER equals the modelled ICER between no screening and screening at age 48, and then decreases as screening at age 48 becomes the more certain option. As the threshold ICER continues to increase, the EVPI decreases to a (smooth) local minimum, and then begins to increase again. This reflects both an increased probability that screening at ages higher than 48 may be cost-effective and an increased value on QALYs possibly lost.

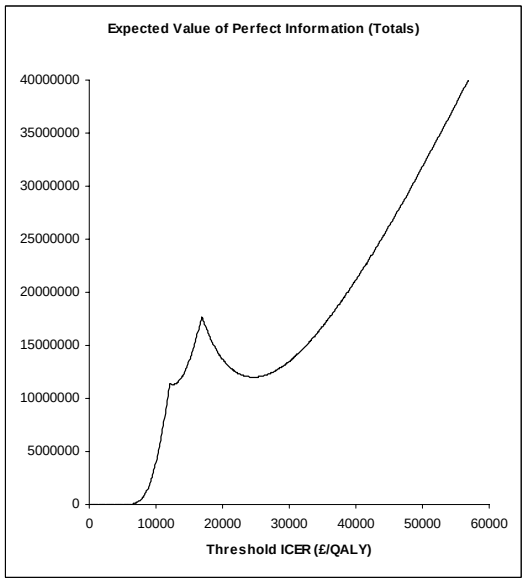
Figure 6 shows the result of allowing screening ages to be any multiple of 8 years. Comparing the general shapes of the curves with those in Figure 5, all curves contain an extra segment reflecting the range of threshold ICERs for which screening at age 40 is now the optimal choice, and the curves based on mean results also show a segment where screening at age 56 is optimal.

Figure 6 CEAF and EVPI when options for screening age are multiples of 8 years

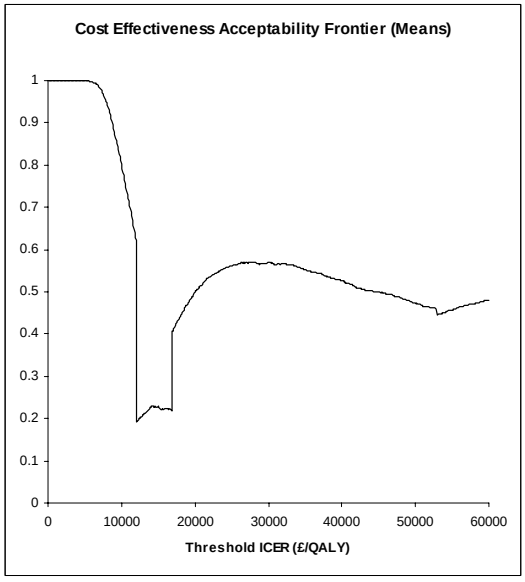
(*a*) CEAF based on total results



(*b*) EVPI based on total results



(*c*) CEAF based on mean results



(*d*) EVPI based on mean results

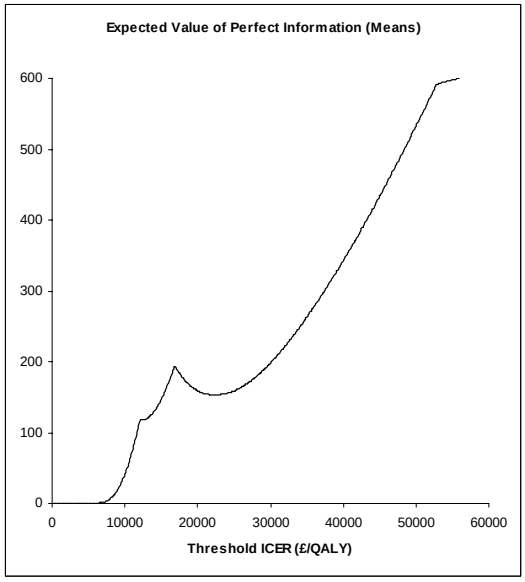
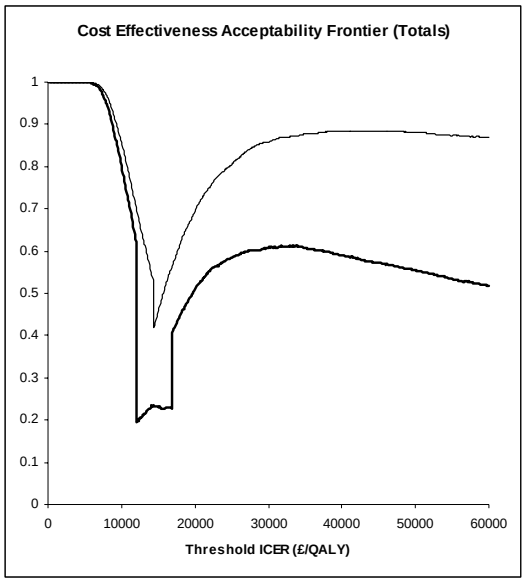


Figure 7 shows the effect of superimposing the various curves in Figures 5 and 6. It is noticeable that, once any uncertainty appears in the curves, the CEAF for the wider option set (thicker curves) is consistently lower than for the narrower option set (thinner curves). The EVPI curves are generally higher for the wider option set than for the narrower. However, this ceases to be true in a region around

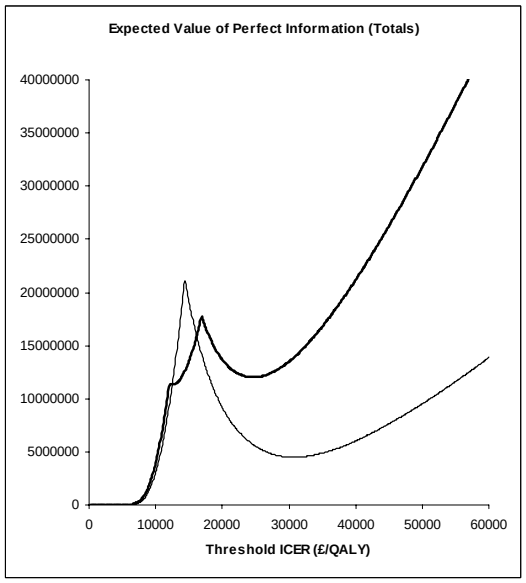
£15,000/QALY. Here the optimal screening age in the wider option set is 40, but this option is not available in the narrower option set.

Figure 7 Comparison of CEAF and EVPI curves for different decision sets

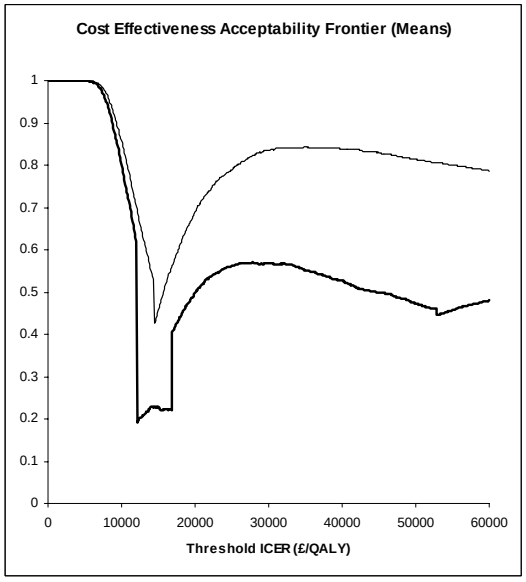
(*a*) CEAF based on total results



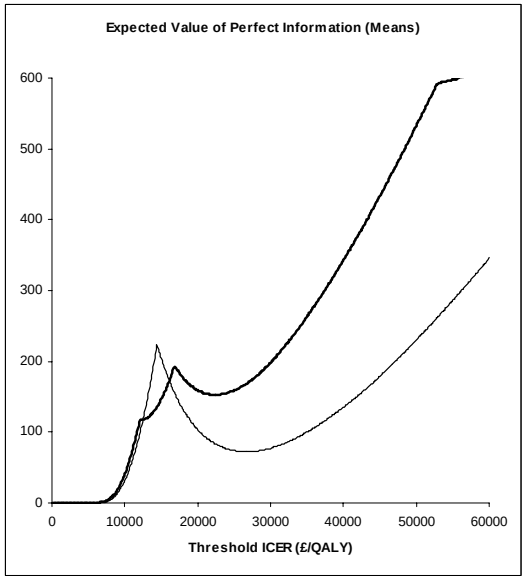
(*b*) EVPI based on total results



(*c*) CEAF based on mean results



(*d*) EVPI based on mean results



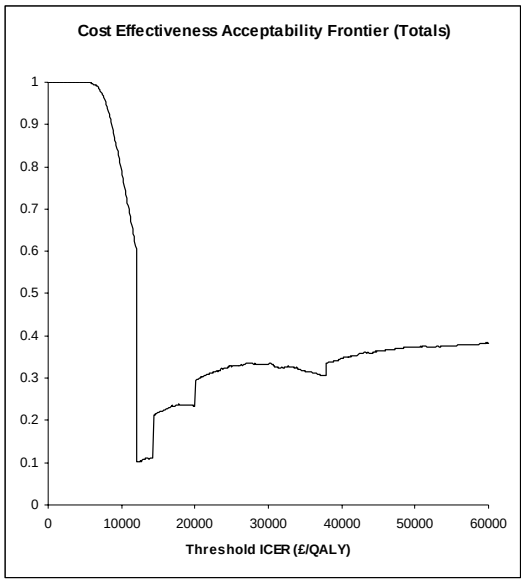
Legend: The thinner curves apply when only multiples of 16 years are acceptable as screening ages, while the thicker curves apply when all multiples of 8 years are acceptable.

It is not difficult to see that, provided that the same option (X, say) remains optimal when the option set is widened, the probability that X is optimal can only decrease while the EVPI can only increase.

Figures 8 to 12 show the CEAF and EVPI curves as the option set is further widened by allowing a greater range of allowable screening ages. These are shown only for calculation base on totals to save space: the general trend in the curves based on means per person screened is the same.

Figure 8 CEAF and EVPI when options for screening age are multiples of 4 years

(*a*) CEAF based on total results



(*b*) EVPI based on total results

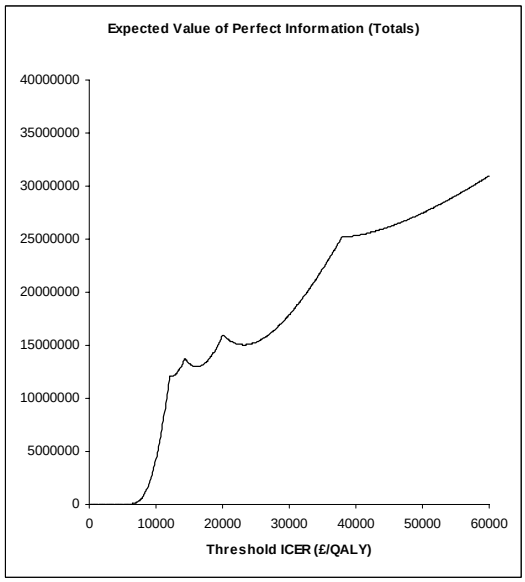
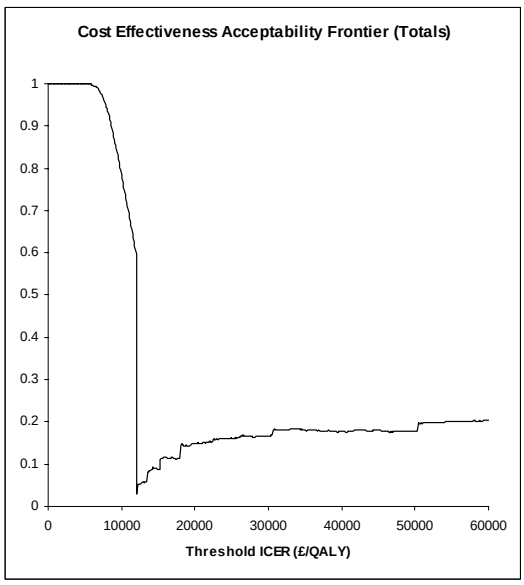


Figure 9 CEAF and EVPI when options for screening age are multiples of 2 years

(*a*) CEAF based on total results



(*b*) EVPI based on total results

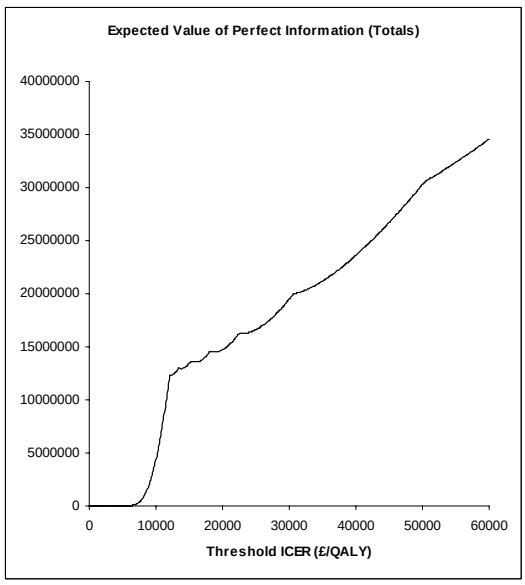
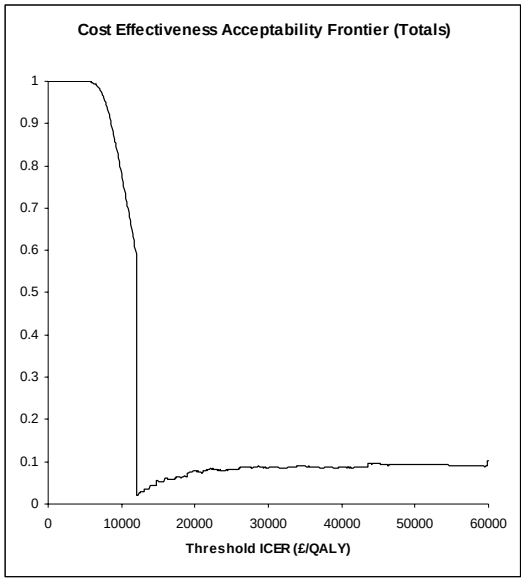


Figure 10 CEAF and EVPI when options for screening age are multiples of 1 year

(*a*) CEAF based on total results (*b*) EVPI based on total results



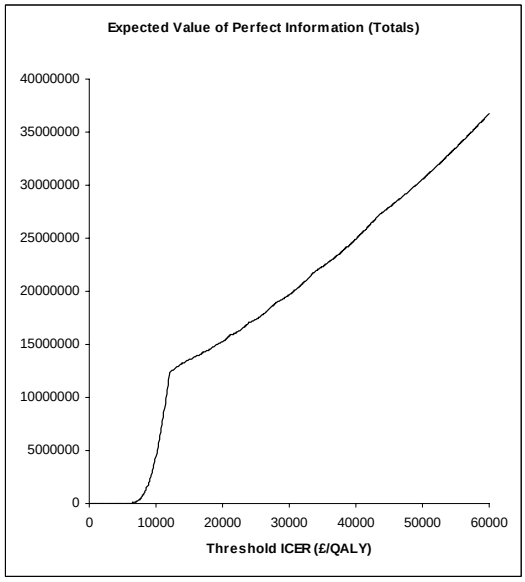
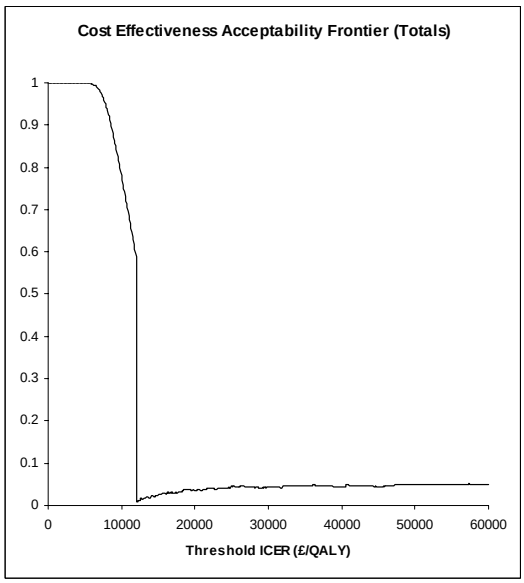


Figure 11 CEAF and EVPI when options for screening age are multiples of 6 months

(*a*) CEAF based on total results (*b*) EVPI based on total results



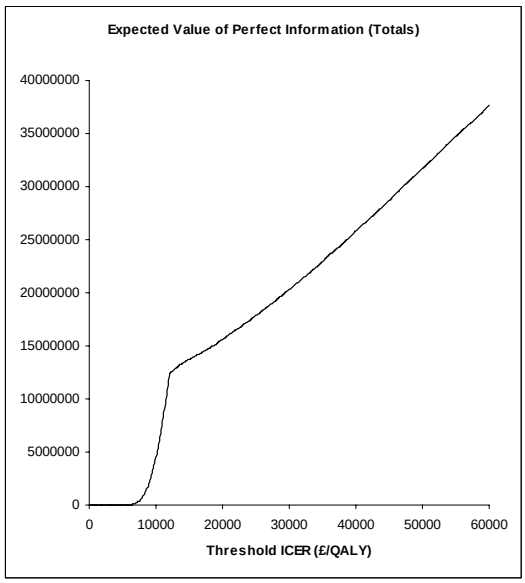
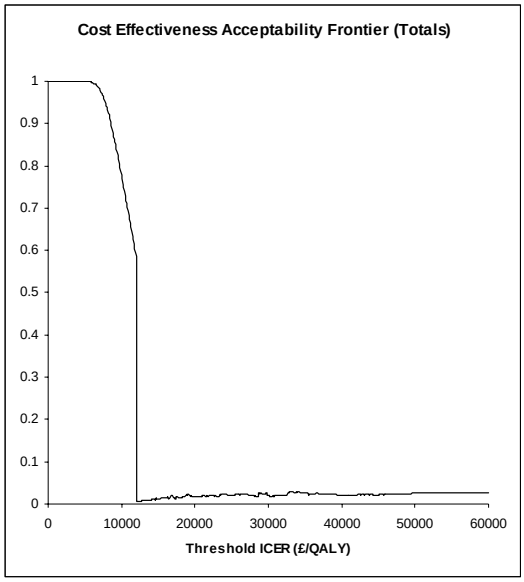
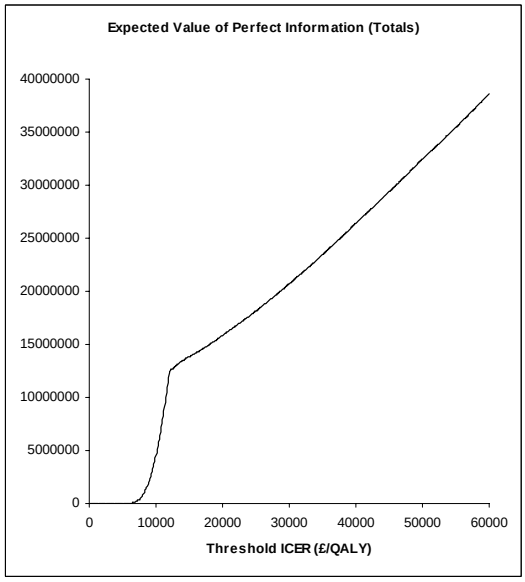


Figure 12 CEAF and EVPI when options for screening age are multiples of 3 months

(*a*) CEAF based on total results (*b*) EVPI based on total results





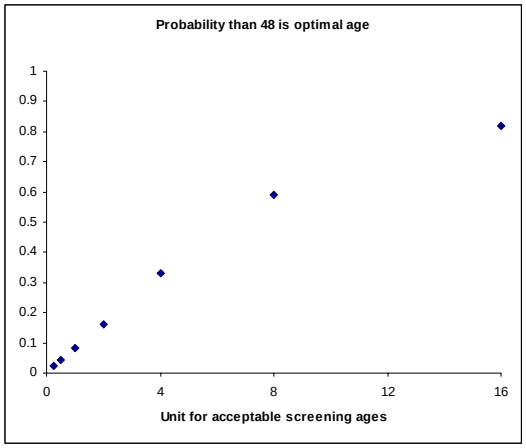
The general trend in the CEAFs is that no screening remains the preferred option, with certainty for very low threshold ICERs, and with probability always greater than 0.5 for threshold ICER up to around £12,000/QALY. Above this threshold ICER, screening is preferred. As the width of the option set increases, the various replications of the model are divided among an increasing number of different optimal options, so the CEAF collapses. In each individual CEAF, the curves are at their lowest just above the point at which screening is preferred to no screening. This is at least partly explained by the fact that there is still a high probability that no screening would be preferred at the relevant threshold ICERs.

Now consider the EVPI curves. As the range of options increases, the number of segments in each curve increases. The sudden decreases in gradient at the points where the preferred option changes become less pronounced; as a result, the part of the EVPI curve where screening at some age is preferred approximates to a smooth curve.

A convenient way of showing the trends in the CEAF and EVPI curve is to plot the heights of these curves against the unit for acceptable screening ages. Figure 13 shows this comparison for a threshold ICER of £25,700/QALY for calculations based on totals. This threshold was chosen as one for which screening at age 48 remains optimal in all the option sets considered. The points plotted appear to lie on a smooth curve in each case, with the CEAF approaching zero and the EVPI approaching a limit of around £19 million as the unit for acceptable screening ages approaches zero.

Figure 13 Trends in CEAF and EVPI curve for threshold ICER £25,700/QALY

(*a*) CEAF based on total results (*b*) EVPI based on total results



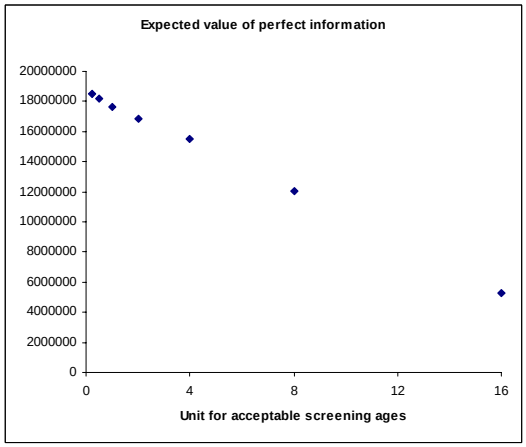
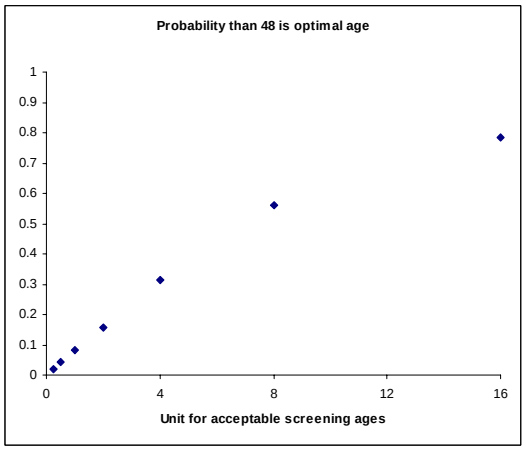


Figure 14 is the equivalent of Figure 13 for calculations based on mean results. In this case, 48 remains the optimal age for screening at a threshold ICER of £24,600/QALY. The essential pattern is similar.

Figure 14 Trends in CEAF and EVPI curve for threshold ICER £24,600/QALY

(*a*) CEAF based on mean results (*b*) EVPI based on mean results



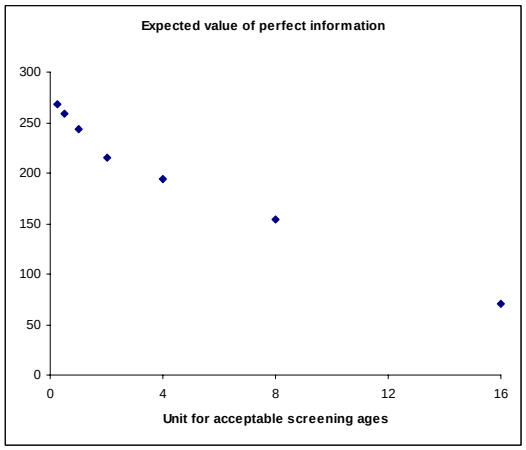
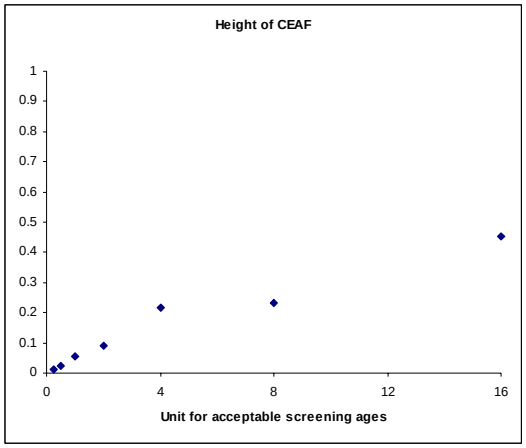


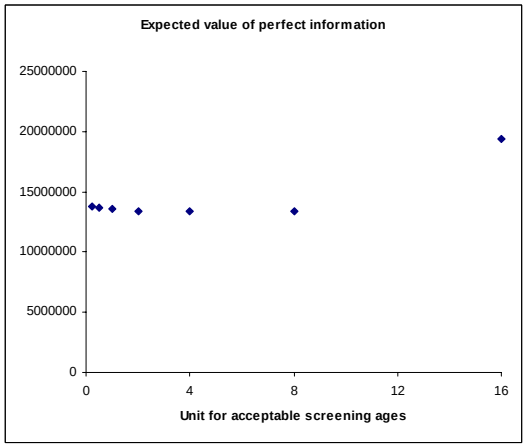
Figure 15 shows equivalent results for a threshold ICER of £14,900/QALY. At this threshold, the optimal age for screening changed every time the decision set was widened. The points lie on a less smooth curve, and the EVPI sometimes decreases as new options become available. However, as the unit for acceptable screening ages approaches zero, the effect of changing the optimal age is reduced, and the curves approach limiting values in a similar way to those shown in Figures 13 and 14.

Figure 15. Trends in CEAF and EVPI curve for threshold ICER £14,900/QALY

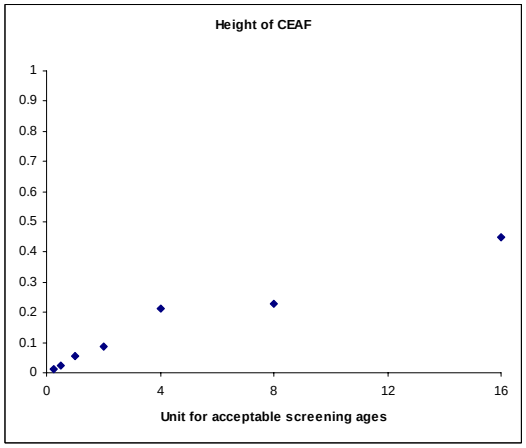
(*a*) CEAF based on total results



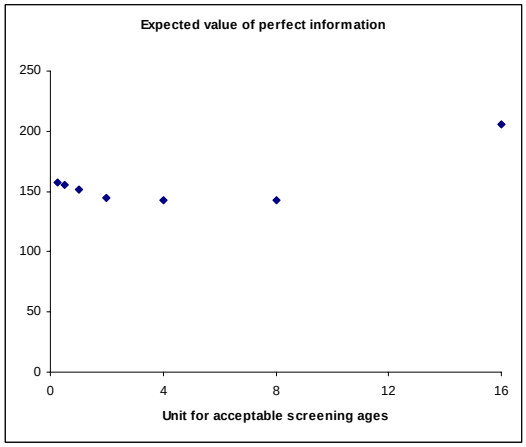
(*b*) EVPI based on total results



(*c*) CEAF based on mean results



(*d*) EVPI based on mean results



Discussion

In this discussion, matters relevant to the original purpose of this investigation are discussed first, followed by other points which have arisen in the course of the work.

The clear result of the work shown is that the Cost-Effectiveness Acceptability Frontier has been shown to be liable to collapse when the number of options considered becomes very large. By contrast, the Expected Value of Perfect Information survives the inclusion of a large number of decision options. Accordingly, the EVPI remains a promising measure of the total uncertainty when a large number of options are possible.

Obvious limitations of the work are that it is based on a single stylised model, which was deliberately structured to be as simple as possible while allowing for the application of a large number of decision options. The number of replications of the model was limited by the computation time available.

The choice of starting with multiples of 16 years as possible screening ages, then including all multiples of 8 years, and so on, was made in order to facilitate the use of a “nested” collection of option

sets. It is of course acknowledged that real-world screening models are likely to use multiples of 5 or 10 years only.

Future plans include extending this work in three ways:

- variations on the model;
- inclusion of strategies with repeated screening at regular intervals;
- calculation of Expected Value of Perfect Parameter Information for some parameter sets within this and possibly other models. It is reasonable to suppose that the absolute EVPPI will converge to a limit in the same way as EVPI. It is not clear how rapidly the ratio of EVPPI to EVPI would converge to its limiting value in that case.

Correct method of comparing different screening ages

An interesting side issue raised by this work is the appropriate method of comparing screening at different ages. Two different methods have been used here, which have essentially regarded the decision as a choice between different population groups to be screened now. The method based on totals is appropriate to a policy of screening all individuals of the appropriate age, but effectively assumes a constant birth rate historically. The method could be adapted by scaling the numbers in each age group to reflect historical variation in birth and survival rates. The method based on means per person screened effectively relates to recruiting a fixed number of individuals to screen, regardless of the variation in the size of the potential target populations.

From the point of view of introducing a new screening programme, there is no need to consider screening two different current age cohorts as mutually exclusive. It would be possible to consider some sort of "catch up" screening programme for those who are already older than the optimal age for once in a lifetime screening, but have missed the opportunity to be screened because the programme did not exist when they were at the optimal age. Such considerations are important, but not the main point of the current work.

An alternative approach to comparing screening at different ages (say, 48 and 56 for consistency with the current paper) is to consider things from the viewpoint of someone who is now 48. Should he be screened now or wait until he is 56 and be screened then? Comparison would then be based on discounting to age 48 for both options, also allowing for the possibility that the individual will not survive to be screened at age 56. This idea could be extended to the full range of screening ages. If applied to repeat screening, discounting should strictly be to the divergence point between strategies. For example, comparing the two options of screening every 2 years or every 4 years starting at age 48. The difference between the two strategies is from age 50 onwards, and so only that part of the patient's potential pathway should be considered, and this discounted to age 50. This is analogous to the methods used in the Birmingham Rheumatoid Arthritis Model (see, for example, Barton *et al*, 2004).

It was the original intention to apply the method described in the previous paragraph in this paper. This was dropped in favour of the simpler method actually used for reasons of convenience. It is thought unlikely that the choice of method would affect the conclusions of this paper.

Which variables to include in the probabilistic analysis?

Another point worth raising is the question of which parameters of the model should be varied in the probabilistic analysis. It is (relatively) uncontroversial that quality of life scores should be included as well as parameters relating to test characteristics, natural history of the condition, and treatment effect. On the question of costs, it is reasonable to distinguish between genuine unit costs and costs which are the average for some course of treatment. The latter could be modelled by considering an average with some uncertainty or by explicitly modelling the extent of resource use (with uncertainty) and then applying fixed unit costs. All the costs in the current model are treated as unit costs.

This leaves the question of how to handle "other causes" death. Strictly speaking, this should be treated as endogenous to the model, in that the more money spent on the intervention modelled, the less there is to be spent on other treatments, and hence the higher the mortality from other causes. However, in practice, this effect is likely to be negligible and it makes sense to model "other causes" death as exogenous. It is then reasonable to note the uncertainty in the parameter(s) used to represent "other causes" death. It is, however, not obvious whether to treat these in just the same way as any other parameters, as has been done in this paper.

Conclusions

The Cost-Effectiveness Acceptability Frontier is of limited use when the options modelled are an essentially arbitrary selection from a very large set of possible options. In contrast, the Expected Value of Perfect Information can be reasonably approximated by a model with a limited choice set. It should, however, be noted that a model with a very restricted choice set may overestimate the total uncertainty measured through the EVPI. This can happen when the limited model does not include an option close to the true optimal choice under current uncertainty that would be found from the "perfect" model including all choices.

Appendix 1

This appendix shows simple examples of various phenomena to do with the cost-effectiveness acceptability curve. They are not claimed as original and the author would welcome pointers to citations, especially any citeable examples of "real life" models showing these effects.

1. Option X dominates option Y in the mean results, but the probability shown for option X at any threshold ICER is arbitrarily low

With arbitrarily low probability p, option X costs $2C$ less than option Y and produces $2Q$ extra QALYs. With probability $1-p$, option X costs pC more than option Y and produces pQ QALYs less than Y. In the mean, option X costs $p(1+p)C$ less than option Y and produces $p(1+p)Q$ extra QALYs.

2. For each pair of options, the one with a higher probability of cost-effectiveness in a bilateral comparison has a lower probability of cost-effectiveness in the three-way comparison

Suppose for example that a model compares three options X, Y, and Z, where X and Y are quite similar and Z is somewhat different from the other two. Suppose that the order of preference in these options at some threshold ICER under different replications of the model is as follows:

30% X Y Z; 28% Y X Z; 23% Z Y X; 17% Z X Y; 1% X Z Y; 1% Y Z X

(These are entirely plausible results given the assumed similarity between X and Y.)

Then the CEAC at that threshold will show probabilities as follows: X 31%, Y 29%, Z 40%, showing (correctly) that Z is the most likely of the three to be the preferred option. However, if we compare the options two at a time, we have the following:

X v Y: X 49%, Y 51%; X v Z: X 59%, Z 41%; Y v Z: Y 59%, Z 41%

In all three cases, the order between options in the pairwise comparison is the reverse of the order in the three-way comparison (albeit marginally in the case of X and Y).

3. Option Z is preferred to options X and Y in the mean results, but has zero probability of being cost-effective with arbitrarily small (but nonzero) skewness in the output distributions

Figure 16. Net benefit for three options

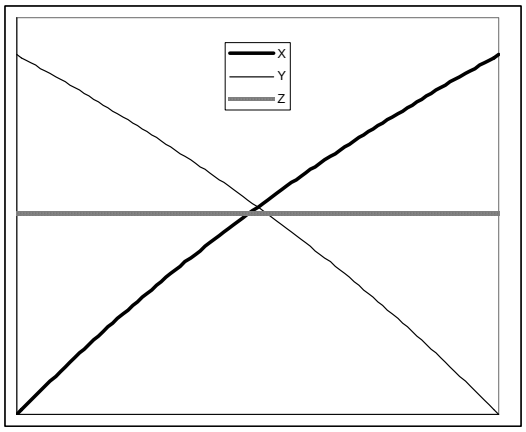


Figure 16 shows the net benefit for the three options at some threshold ICER as a function of a single uniformly distributed uncertain parameter. There is just enough skewness in the distributions of net benefits for X and Y that the mean values of the net benefit are below the constant net benefit of Z. Clearly, the skewness can be made arbitrarily small while preserving the overall pattern.

4. Option Z is preferred to options X and Y in the mean results, but has arbitrarily low probability of being cost-effective, with symmetrical output distributions

Consider a figure similar to Figure 16 but with straight line graphs for X and Y. The graph for Z must now be above the crossing point of the graphs for X and Y, but can be as close as desired.

Each of the last three was described in terms of preferences at a single threshold ICER. Sufficient negative correlation between costs and outcomes would allow the effect to be maintained across the whole range of thresholds.

Appendix 2

Suppose someone incurs an annual cost or accumulates QALYs at a rate *r* per year for *s* years. Apply a continuous discounting function so that costs or QALYs generated at time *t* are multiplied by $\exp(-\lambda t)$. Then the total discounted cost or QALYs are given by

$$\int_0^s r\exp(-\lambda t)dt = -\frac{r}{\lambda}\left[\exp(-\lambda t)\right]_0^s = -\frac{r}{\lambda}\left(\exp(-\lambda s)-1\right) = r\frac{1-\exp(-\lambda s)}{\lambda}.$$

Now consider the individual illustrated in Figure 1 on page 5. If he is screened at age 80 and tests positive, he will incur the costs of the confirmatory test, surgery, and maintenance therapy for just over 31 years. The cost of this will be $1000 + 50000 + 500 \times \frac{1-\exp(-31.08\lambda)}{\lambda} = 60540,$ where $\lambda = \ln(1.035)$ representing the discount rate. These costs are only incurred in the case of a positive result, so the expected cost of screening at that age is $50 + 0.9784 \times 60540 = 59290.$

References

Barton P, Jobanputra P, Wilson J, Bryan S, Burls A (2004) The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. *Health Technology Assessment* 8(11).

Claxton K (1999) The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics* 18: 341-364.

Fenwick E, Claxton K, Sculpher M (2001) Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics* 100: 779-787.

Van Hout BA, Al MJ, Gordon GS, Rutten FFH (1994) Costs, effects and C/E-ratios alongside a clinical trial. *Health Economics* 3: 309-319.

The author would welcome suggestions of appropriate references to illustrate the various points made in this paper.