University
of Glasgow

Briggs, A.H. and Goeree, R. and Blackhouse, G. and O'Brien, B.J. (2002) Probabilistic analysis of cost-effectiveness models: choosing between treatment strategies for gastroesophageal reflux disease. *Medical Decision Making* 22(4):pp. 290-308.

http://eprints.gla.ac.uk/4147/

Deposited on: 8 May 2008

Glasgow ePrints Service
http://eprints.gla.ac.uk

Probabilistic Analysis of Cost-Effectiveness Models: Choosing between Treatment Strategies for Gastroesophageal Reflux Disease

Andrew H. Briggs, DPhil, Ron Goeree, MA, Gord Blackhouse, MBA, Bernie J. O'Brien, PhD

When choosing between mutually exclusive treatment options, it is common to construct a cost-effectiveness frontier on the cost-effectiveness plane that represents efficient points from among the treatment choices. Treatment options internal to the frontier are considered inefficient and are excluded either by strict dominance or by appealing to the principle of extended dominance. However, when uncertainty is considered, options excluded under the baseline analysis may form part of the cost-effectiveness frontier. By adopting a Bayesian approach, where distributions for model parameters are specified, uncertainty in the decision concerning which treatment option should be implemented is addressed directly. The approach is illustrated using an example from a recently published cost-effectiveness analysis of different possible treatment strategies for gastroesophageal reflux disease. It is argued that probabilistic analyses should be encouraged because they have potential to quantify the strength of evidence in favor of particular treatment choices.

It is now increasingly common for economic evaluations to be conducted alongside clinical trials. Recent research attention has been focused on how to handle uncertainty in these so-called stochastic cost-effectiveness analyses where patient-level data are available on the costs and effects of treatment options.[1–3] However, the majority of economic evaluations still employ a decision analytic modeling framework to synthesize data from a number of sources.[4] Such cost-effectiveness models are often described as deterministic analyses. Although the limitations of simple univariate sensitivity analysis are well known, this remains the most popular technique for handling uncertainty in cost-effectiveness models.

Probabilistic sensitivity analysis is an alternative approach that involves specifying distributions for input parameters in the model and employing Monte Carlo simulation to sample from these distributions, allowing the joint effect of parameter uncertainty to be assessed.[5,6] A number of commentators have suggested that probabilistic sensitivity analysis methods be used to handle uncertainty in cost-effectiveness models,[1,7] including the US panel on cost-effectiveness analysis.[8] Despite these recommendations, few probabilistic analyses of cost-effectiveness models have been undertaken. The relative paucity of probabilistic analyses may be due to the increased complexity of the ap-

proach and a lack of clarity concerning which distributions for input parameters are appropriate.

The aim of this article is to demonstrate the use of probabilistic sensitivity analysis to handle uncertainty in a cost-effectiveness decision problem relating to alternative treatment options for gastroesophageal reflux disease (GERD). We argue that adopting a Bayesian approach to uncertainty offers both technical and conceptual advantages over traditional sensitivity analyses. In particular, a Bayesian approach allows a more intuitive interpretation of probability—we show how the study question of whether a treatment is cost-effective can be answered directly in the form of a probability that the intervention is cost-effective. Furthermore, we demonstrate this approach in the case of multiple treatment options for GERD, rather than the standard 2-treatment approach that is the norm in the majority of analyses.

The article is structured as follows. In the next section we give a brief introduction to the decision problem, the structure of the model, and the results of the previously published cost-effectiveness model—where uncertainty was handled through standard deterministic sensitivity analysis methods. The section that follows demonstrates how the model can be made probabilistic by specifying distributions for the input parameters following standard principles of Bayesian methods. The assumptions and calculations involved in specifying these distributions are discussed in detail. Results of the probabilistic analysis are then presented on the cost-effectiveness plane and summarized through the use of cost-effectiveness acceptability curves. These results, and the general probabilistic approach to cost-effectiveness modeling, are discussed in the final section of the article.

TREATMENT STRATEGIES FOR GERD

In this section, a model for assessing the cost-effectiveness of 6 management strategies for the treatment of GERD is briefly outlined. Full details of the model were presented in detail in a previous publication.[9] First, the structure and assumptions concerning the decision model are discussed. Second, the results of the deterministic analysis are presented. Finally, the limitations of the originally reported univariate sensitivity analysis are highlighted.

A Model for Assessing the Cost-Effectiveness of GERD Treatment

GERD is a common condition that results from regurgitation of acid from the stomach into the esophagus. The most frequent symptom of GERD is heartburn,

and the majority of patients with GERD require pharmacotherapy to reduce acid secretion. Currently, the choice of first-line antisecretory therapy is between the H_2-receptor antagonists (H_2RAs), such as ranitidine and cimetidine, and proton pump inhibitors (PPIs), such as omeprazole. Although they have higher acquisition costs, PPIs have been found to be more efficacious than H_2RAs in terms of both the rate and speed of healing.[10,11]

The objective of the original study was to compare, over a 1-year period, the expected costs and outcomes of alternative drug treatment strategies for the management of patients with erosive esophagitis confirmed by endoscopy, but without complications such as Barrett's esophagus or stricture. Outcomes are quantified in terms of GERD recurrence and weeks per year without GERD as indicated by data from clinical trials on healing and recurrence of esophagitis.

Treatment Strategies and Model Structure

Six strategies involving different combinations of first-line agents and change of therapy conditional on failure to heal or recurrence of GERD were modeled.

Strategy A: Intermittent PPI. Acute treatment with a PPI for 8 weeks and then no further treatment with prescription medication until recurrence.

Strategy B: Maintenance PPI. Acute treatment with a PPI for 8 weeks then continuous maintenance treatment with a PPI (same dose).

Strategy C: Maintenance H_2RA. Acute treatment with an H_2RA for 8 weeks and then continuous maintenance treatment with an H_2RA (same dose).

Strategy D: Step-down maintenance prokinetic agent (PA). Acute treatment with a PA for 12 weeks and then continuous maintenance treatment with a lower dose of PA.

Strategy E: Step-down maintenance H_2RA. Acute treatment with a PPI for 8 weeks and then continuous maintenance treatment with an H_2RA.

Strategy F: Step-down maintenance PPI. Acute treatment with a PPI for 8 weeks and then continuous maintenance treatment with a lower dose PPI.

Treatment options A to F represent clinical strategies rather than single-drug treatments for the management of erosive esophagitis where the physician is assumed to increase the dose of a drug or switch to another drug if the patient fails to respond to the first-line treatment. The logic of these assumptions with regard to stepping up dosage or switching can be found in Table 1. The structure of the decision tree that was developed is shown in Figure 1 and is based on the treatment strategies and step-up switching algorithms in Ta-

Table 1. Step-Up and Switching Algorithms Conditional on Healing Failure or Recurrence

Strategy A: Intermittent PPI			Strategy B: Maintenance PPI		
Healing	Maintenance	1ST Recurrence	Healing	Maintenance	1ST Recurrence
PPI →	No therapy →	PPI to heal	PPI →	PPI →	DD PPI to heal
↓ unhealed			↓ unhealed		
DD PPI →	PPI →	DD PPI to heal	DD PPI →	PPI →	DD PPI to heal

Strategy C: Maintenance H₂RA			Strategy D: Step-Down Maintenance PA		
Healing	Maintenance	1ST Recurrence	Healing	Maintenance	1ST Recurrence
H₂RA →	H₂RA →	DD H₂RA	PA →	LD PA →	PA to heal
↓ unhealed			↓ unhealed		
PPI →	H₂RA →	PPI to heal	PPI →	LD PA →	PPI to heal
↓ unhealed			↓ unhealed		
DD PPI →	PPI →	DD PPI to heal	DD PPI →	PPI →	DD PPI to heal

Strategy E: Step-Down Maintenance H₂RA			Strategy F: Step-Down Maintenance PPI		
Healing	Maintenance	1ST Recurrence	Healing	Maintenance	1ST Recurrence
PPI →	H₂RA →	PPI to heal	PPI →	LD PPI →	PPI to heal
↓ unhealed			↓ unhealed		
DD PPI →	PPI →	DD PPI to heal	DD PPI →	PPI →	DD PPI to heal

Note: PPI = proton pump inhibitor (e.g., omeprazole 20 mg OD).
DD PPI = double dose proton pump inhibitor (e.g., omeprazole 40 mg OD).
LD PPI = low dose proton pump inhibitor (e.g., omeprazole 10 mg OD).
H₂ RA = H₂ receptor antagonists (e.g., ranitidine 150 mg BID).
DD H₂ RA = double dose H₂ receptor antagonists (e.g., ranitidine 300 mg BID).
PA = prokinetic agent (e.g., cisapride 10 mg QID).
LD PA - low dose prokinetic agent (e.g., cisapride 10 mg BID).
Reprinted with permission from Goeree et al. (1999), Adis International.

ble 1. The model is recursive in two 6-month periods; hence, probabilities of recurrence in the period to 12 months are conditional upon recurrence or non-recurrence in the period from 0 to 6 months.

Treatment Outcomes

For GERD, the most commonly used formulation of outcome for economic evaluation has been either esophagitis-free or symptom-free time in a period of follow-up. The advantage of such a measure is that it combines 2 important aspects of efficacy: (1) the speed with which esophagitis is healed and (2) the likelihood of esophagitis recurring. In this analysis, the primary outcome measure is GERD-free time during the 12-month period of the model, defined as the time when the esophagitis is healed. A meta-analysis of healing and recurrence studies published to November 1997 was undertaken to estimate healing and recurrence probabilities together with associated GERD-free time. Full details of this analysis are given in the original study.[9]

Resource Use and Unit Costs

Generic prices were used for drugs when a generic equivalent was available, employing the "best available price" from the Ontario Drug Benefit (ODB) program[12] together with a 10% pharmacy markup charge. A dispensing fee of Can$4.11 was used (i.e., ODB program fee of Can$6.11 less a Can$2.00 patient copayment).

Cost estimates for physician fees were taken from the physician fee schedule for Ontario,[13] and procedure costs, such as endoscopy, were estimated from a hospital participating in the Ontario Case Costing Project in Southwestern Ontario.[14]

To estimate the costs associated with the management of patients with symptoms of GERD recurrence, information on clinical practice patterns and resource utilization was obtained by convening an expert physician panel and using a modified Delphi technique.[15] Estimated resource utilization was then combined with unit cost information to give the average cost asso-
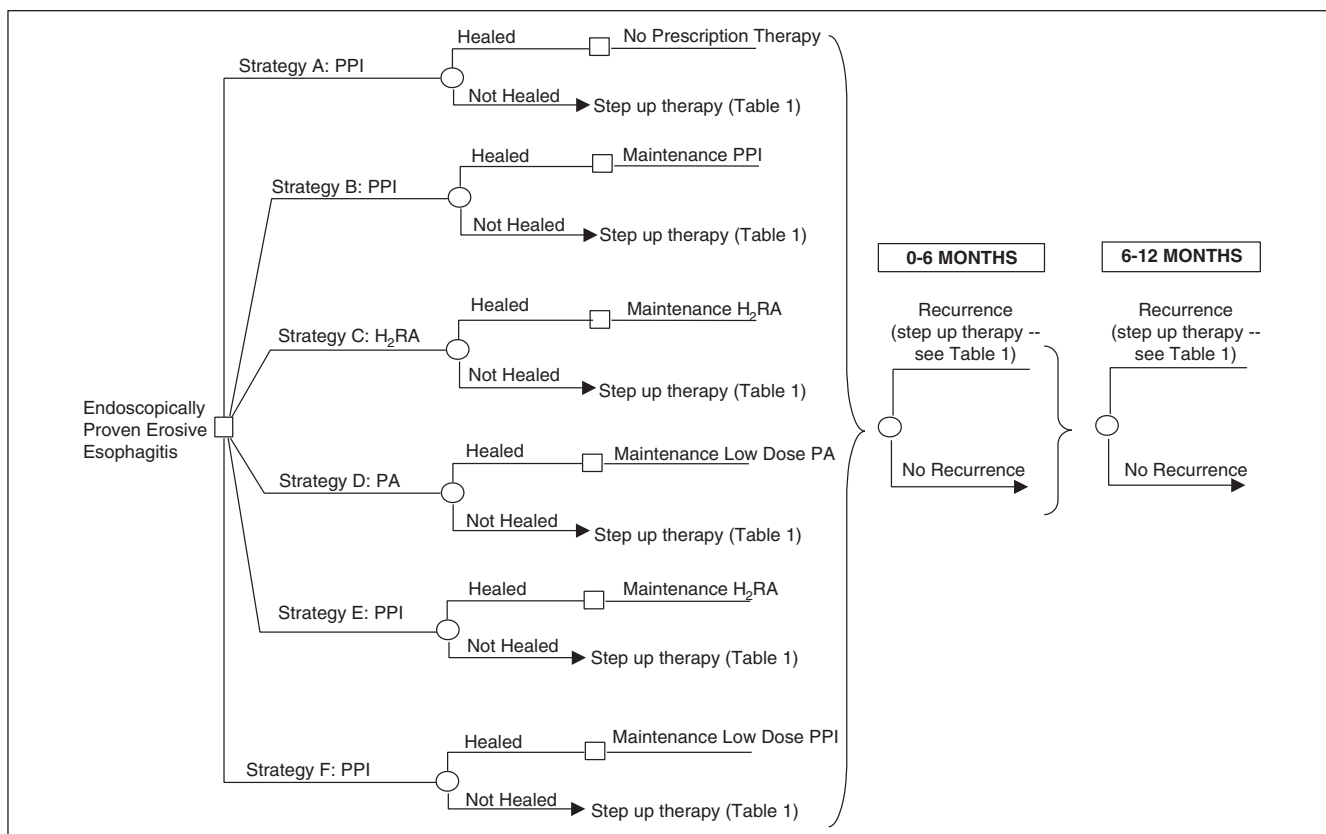
Figure 1. Decision tree for the management of erosive esophagitis. PPI = proton pump inhibitor; H_2RA = H_2-receptor antagonist; PA = prokinetic agent. Reprinted with permission from Goeree et al. (1999), Adis International.

ciated with each recurrence under each management strategy.

Results of the Deterministic Cost-Effectiveness Analysis

The decision tree model outlined above was evaluated to estimate the expected costs and the expected weeks without GERD in the 12-month period of the model. The analysts of the original study took the conventional approach to examining the cost-effectiveness of the alternative strategies.[16–18] First, it was determined whether any strategies were simply dominated by other strategies having both lower costs and greater therapeutic effects. Second, it was determined whether any strategies were dominated through the principles of extended dominance (i.e., whether linear combinations of other strategies can produce greater benefit at lower cost).[19] Finally, among nondominated treatment options, incremental cost-effectiveness ratios were calculated by comparing each option to the next more costly and more effective intervention. This process produces an "efficiency frontier" of increasingly more costly and more effective strategies. The results of the analysis are presented on the cost-effectiveness (CE) plane in Figure 2, which also shows the efficiency frontier.

The figure clearly shows that step-down maintenance PA (strategy D) is dominated by maintenance H_2RA (strategy C), intermittent PPI (strategy A), and step-down maintenance H_2RA (strategy E). The efficiency frontier is given by the lines joining strategies C (the origin), A, E, and B. Strategy F is internal to this frontier, indicating that it also can be ruled out through the principle of extended dominance (i.e., a linear combination of strategies E and B would strongly dominate F). The slope of the frontier at any point reflects incremental cost-effectiveness—the additional cost at which additional effects can be purchased.

Limitations of Conventional Sensitivity Analysis

Parameter estimates employed in the model are not known with certainty; therefore, it is important to ex-
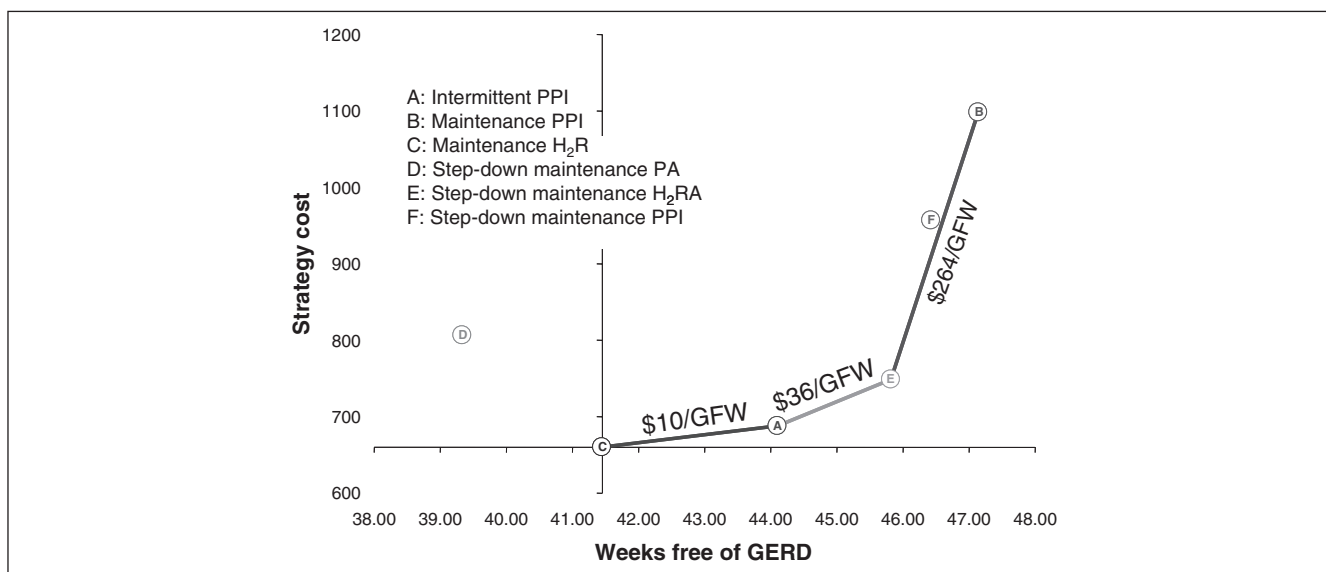
Figure 2.   Baseline cost-effectiveness results on the cost-effectiveness plane showing the "efficient frontier." PPI = proton pump inhibitor; H_2RA = H_2-receptor antagonist; PA = prokinetic agent; GERD = gastroesophageal reflux disease; GFW = GERD-free week.

plore the implications of parameter uncertainty for the results of the analysis. In particular, given the importance of excluding dominated interventions from the analysis before calculating the frontier, the extent to which particular strategies are part of (or can be excluded from) the frontier should be assessed.

In the original analysis, the authors examined the effect of a number of the parameters in their model using conventional sensitivity analysis techniques. They showed how the uncertainty in the parameter values chosen for the baseline analysis might affect the frontier. The sensitivity analysis that the authors presented was much less arbitrary than that of many reported cost-effectiveness analyses because the outcome ranges chosen were the 95% confidence intervals (CIs) from the reported meta-analyses of healing and recurrence rates. The authors reported that

> there were marked differences in expected costs, recurrences and weeks with GERD when using the lower and upper CIs for both healing and recurrence rates. However, there were no changes in the relative ranking of strategies for either costs or outcomes. The basic conclusions of the base-case analysis were not altered by using the lower or upper 95% CIs for healing or recurrence rates.[9(pp689,691)]

Despite this convincing argument, we might still be concerned that the full effects of uncertainty are more important than the authors suggest. It is well known[7,8] that conventional univariate sensitivity analysis, whereby individual parameters are varied while maintaining all remaining parameters at their baseline value, is likely to underestimate uncertainty because, in reality, parameters will not vary in isolation.

A BAYESIAN APPROACH TO PROBABILISTIC SENSITIVITY ANALYSIS

In this section, the general appeal of adopting a Bayesian approach to probabilistic analysis of cost-effectiveness models is presented. In the past, the use of probabilistic sensitivity analysis has been employed without reference to the statistical framework of the approach.[5,6] We argue that probabilistic sensitivity analysis (and indeed decision analysis itself) is naturally Bayesian and adopting this approach offers both technical and conceptual advantages. Particular emphasis is given to the choice of distributions for the different types of parameters commonly encountered in cost-effectiveness models, and it is argued that Bayesian methods provide the solution to which distributions should be used for parameters estimated from different types of data.

Probabilistic Analysis

Probabilistic sensitivity analysis involves specifying distributions for model parameters to represent uncertainty in their estimation and employing Monte Carlo simulation to select values at random from those distributions.[5,6] In this way, probabilistic models allow the effects of joint uncertainty across all the parameters

of the model to be considered. Note that for standard frequentist analyses (such as practiced in almost all clinical trials), parameters to be estimated from the data are considered to have true values that do not vary. Probabilities attached to CIs relate to the long-run coverage probabilities of the intervals if the same experiment were to be repeated many times.

By contrast, in probabilistic modeling, parameters are considered random variables, which can take a range of values described by the specified distribution. Although these distributions will represent "degrees of belief" in the parameters of interest, it does not necessarily follow that the analysis will become automatically "subjective" (the great fear of many of those who object to Bayesian methodology). When data are lacking and it becomes necessary to engage experts to provide information on prior distributions, then a number of experts should be consulted in order that the distributions reflect uncertainty between experts rather than representing the subjective beliefs of a single expert. Eddy et al.[20,21] outlined just such an approach to synthesizing data based on Bayesian methods that they termed the "confidence profile" technique.

Choosing Distributions for the Parameters

Parameters in decision models represent summary values related to the average experience across a population of (potential) patients. Therefore, the relevant uncertainty to capture in the formation of a distribution for the parameter is 2nd-order uncertainty related to the sampling distribution of the parameter, not the variability in the values observed in a particular population (1st-order uncertainty; see Stinnett and Paltiel[3] for further discussion). Although an assumption of normality for parameters is widely used in statistics, it is worth remembering that the assumption is based on asymptotics (the central limit theorem) and that the normal distribution has no bounds on values it can take. In practice, parameters of the model will have logical limitations on the values they can take. In this section, we discuss 4 different types of parameters commonly employed in cost-effectiveness models: probabilities, resource items, unit costs, and relative risks. For each, we discuss the nature of the data informing parameter estimates, the logical bounds on the parameter, and the way in which Bayesian methods can help to select distributions for parameters.

Probability Parameters

Probabilities for cost-effectiveness models are often based on the observed proportions of the event of interest (e.g., the number of successfully treated cases). At an individual level, a treated patient is classed as either a success or a failure; therefore, the data can be considered as independent Bernoulli trials leading to a binomial form of the data likelihood. With such data, it is natural to use the proportion of successful patients as the estimate of the corresponding probability in the model. However, in considering the distribution of that probability, note that the binomial distribution is a discrete distribution related to the sample size of the study generating the data, whereas it makes sense to model the distribution of probability in the model as continuous.

Standard frequentist methods for estimating a CI for a proportion involve calculating the binomial estimate of variance and assuming a normal sampling distribution in order to generate the interval

$$\left(p - 1.96 \times \sqrt{p(1-p)/n}, \quad p + 1.96 \times \sqrt{p(1-p)/n}\right)$$

(where p is the proportion and n is the sample size).[22] Although this method gives a good approximation to the true CI when p is not close to 0 or 1, the assumption of normality is not appropriate for probabilistic sensitivity analysis. This is because the probability is known to be bounded on the interval 0-1 whereas the normal distribution will (eventually) generate values outside this interval in a Monte Carlo simulation because it is unbounded.

One solution to this is to make the transformation to the logistic scale, which is unbounded, assume normality on this scale, then transform back to the original 0-1 probability scale, as described by Doubilet et al.[6] in their early contribution on the use of probabilistic methods for clinical decision making. Although this provides a solution to the problem of bounding, it turns out to be a reasonably sophisticated problem to choose the parameters for the normal distribution on the logit scale in order to match the desired mean and standard deviation on the probability scale (see the derivations in Appendix A of Doubilet et al.[6] for details).

Fortunately, Bayesian methods provide a straightforward method for moving from the discrete binomial likelihood to the continuous uncertainty concerning the probability parameter. The beta distribution is a continuous distribution on the interval 0-1 and is conjugate to the binomial distribution. This means that if it is possible to represent prior belief using a beta distribution, then the integration of that prior belief with the binomial data has a closed form, with the result that the posterior distribution of the probability will also follow a beta distribution. Fortunately, by varying the 2 parameters of the beta distribution, a wide variety of possible shapes to the distribution over the interval can be

obtained: skewed, symmetric, uniform, near normal, and even U-shaped. One parameterization of the beta distribution, beta(r,n), has a similar interpretation to r successes from n trials. A 2nd common parameterization is beta(α, β), where $\alpha = r$ and $\alpha + \beta = n$. The mean and variance of beta(r,n) distribution are given by

$$mean = \frac{r}{n} \qquad sd = \sqrt{\frac{r(n-r)}{n^2(n+1)}}.$$

Furthermore, if we can specify a prior distribution as beta(r', $n' - r'$), then following an observation of r successes and $n - r$ failures in n trials, the application of Bayes's theorem yields the result that the posterior distribution is beta($r' + r, n' - r' + n - r$). Where no prior information exists as to the probability, it is appropriate to use an "uninformative" or reference prior. Although beta(1,1), which yields a uniform distribution over the interval 0-1, seems an intuitively obvious choice of prior, in fact what constitutes uninformative in this context is not as straightforward as it appears,[23] and beta(0.5,0.5), which yields a U-shaped distribution, is also a popular choice for a reference prior. Because uninformative priors will be dominated by the data, the issue of which uninformative prior to employ is unlikely to be of practical importance when data are available to update that prior.

Resource Item Parameters

All economic analyses are concerned with the use of resources. The numbers of resource items that a patient uses can be considered a count variable. The Poisson distribution with parameter λ (which gives both the mean and variance of the distribution) is often used to model count data. If we are interested in the distribution of the mean resource use for a group of patients, we could use the Poisson estimate of variance to obtain a standard error for the mean resource use, relying on the central limit theorem to give a normal sampling distribution. However, this may be problematic for smaller samples due to a nonnegligible probability that the normal distribution could take a value less than 0 when it is clear that mean resource use cannot be negative.

Again, the Bayesian approach provides a solution. The gamma distribution is conjugate to the Poisson distribution, is constrained to be positive, and is fully continuous. Therefore, the gamma distribution for the mean resource use can be specified without fear of generating inconsistent values in a probabilistic analysis.

Unit Cost Parameters

Unit costs are applied to resource volumes in order to evaluate all resource use on a common (monetary) scale. Note that the unit of analysis for such costs is different from that for other parameters—unit costs are typically calculated across a broad group of patients. The unit cost of a surgical procedure or stay in a particular ward will typically be given at the level of the hospital (or similar provider unit). By contrast, the unit cost of a drug or device may be set provincially or nationally and may not vary at all within the context of a country-specific cost-effectiveness analysis. Furthermore, the unit cost of a resource item is strictly continuous, unlike the data on resource use considered above. Because unit costs are constrained to be positive, a gamma distribution could be used to represent uncertainty in these costs. However, it is less clear that unit costs will be highly variable than the resource items they are employed to value, which means that a normal distribution may be safely employed. It is perhaps telling that most economic analyses conducted alongside clinical trials treat unit costs as fixed rather than stochastic.

Relative Risk Parameters

It is very common for economic models to include relative risks as parameters. This mirrors the fact that relative risk is often the primary outcome in clinical trials. Methods for calculating CIs for relative risk estimated in such trials assume that the central limit theorem will lead to the natural logarithm of relative risk (which is additive) being normally distributed such that CIs can be determined in the standard way. A CI for a relative risk is then obtained by exponentiating the CIs on the log scale. This standard approach to CI estimation clearly suggests an equivalent approach to specifying a log-normally distributed parameter for relative risk to be used in a probabilistic sensitivity analysis. Furthermore, because the normal distribution is self-conjugate (a normal prior and a normal data likelihood generate a normal posterior distribution), the application of Bayes theorem on a normally distributed parameter is especially straightforward.

A PROBABILISTIC ANALYSIS OF TREATMENT STRATEGIES FOR GERD

In this section, we describe how a probabilistic sensitivity analysis of this decision problem was undertaken in order to more fully account for uncertainty in the choice of treatment strategy for GERD. Within the model, there are 3 main categories of parameters: model probabilities relating to the healing and recurrence rates of GERD symptoms, parameters relating to the level of resource consumption by patients with GERD symptoms, and unit costs of those resources.

Each of these parameter categories are discussed in detail below, and a full list of the parameters of the model is given in Table 2. Note that the outcome variables in the model—weeks free of GERD—are completely determined by the healing and recurrence rates and, therefore, are endogenous variables in the model.[9]

Parameter Distributions

Distributions for the Healing and Recurrence Probabilities

All patients begin the model with GERD. Following first-line therapy, there is a probability that their GERD will have healed. Once GERD has healed, there is then the probability that it will recur. The healing and recurrence probabilities were estimated from the literature. Consider that at an individual level, a patient with GERD has either been healed or has experienced a recurrence. Therefore, the clinical investigation of healing and recurrence can be considered as leading to a binomial form of the data likelihood as described above; hence, a beta distribution was chosen to represent uncertainty in the healing/recurrence parameters.

The original study went to some lengths to present a rigorous meta-analysis of healing and recurrence probabilities. This method resulted in estimates of constant hazards for healing probabilities and estimates of proportions of patients recurring in 2 periods—0 to 6 months following healing and 6 to 12 months following healing—together with associated estimates of standard error. Due to the random-effects assumption, this is more conservative than the Bayesian updating approach described above, which would be equivalent to simply pooling all the studies directly. Therefore, beta distributions were fitted by method of moments[23]: the mean and standard errors from the meta-analysis were equated to the estimates of mean and standard error of the beta distribution given, and these equations were then solved to give the appropriate beta distribution parameters (see the appendix for this derivation). Details of the distributions for the healing hazards and the recurrence probabilities fitted by this method are given in Table 3.

Distributions for Resource Use Assumptions

In contrast to the rigorous meta-analytic approach employed to summarize the wealth of information on the healing and recurrence rates associated with different drug interventions for GERD, the information on resource use, particularly the level of investigations received by patients following a recurrence, was extremely sparse. Although in the original study a Delphi panel of experts was convened in order to estimate the likely experience of patients, the purpose of the panel was to forge consensus, and no information on the variance of estimates that emerged prior to consensus of the experts remains. Therefore, the assumptions concerning the distributions of estimated resource use parameters are much more arbitrary.

For the estimated number of visits to general practitioners and for endoscopic investigation, a gamma distribution was assumed. This is because the number of visits is constrained to be positive and the gamma distribution is only defined for positive values. Again a method-of-moments approach to fitting was employed such that the mean of the gamma distribution was equal to the point estimates of the visits generated by the expert panel and assuming that the standard error was half that value (i.e., assuming the coefficient of variation was 0.5; see the appendix for the derivation).

For the proportions of patients receiving the various investigative procedures, it was assumed that the expert panel had related its estimates to a hypothetical cohort of 100 patients. Therefore, a beta distribution was again employed as if the event rates given by the expert panel were per 100. Given the considerable experience of the panel with GERD treatment, it is likely that this approach is conservative.

It is assumed that variation in medication use is negligible, such that all patients obtain their prescriptions and all prescriptions accord with the treatment strategies under evaluation. The chosen distributions for the resource use parameters are presented in Table 4.

Unit Costs of Resources

The 2nd component of uncertainty in cost is the potential uncertainty in the unit cost estimates employed to value resource items. We do not believe that it is appropriate to handle uncertainty in drug prices probabilistically, since, at the point of the evaluation, drug prices are determined by the manufacturers. Hence, although drug prices might be considered variable (because they are under the control of the manufacturer and may change over time), they are not uncertain. Of course, there may be some uncertainty concerning which drugs it is appropriate to prescribe, but that is part of the decision problem and is best handled outside of the probabilistic component of the analysis. Therefore, drug prices were not varied in this analysis.

A separate issue relates to the use of scheduled information of the cost of resource items in Ontario, Canada, where the original study was carried out. Although there is certainly an issue concerning whether scheduled reimbursement values for resources reflect the
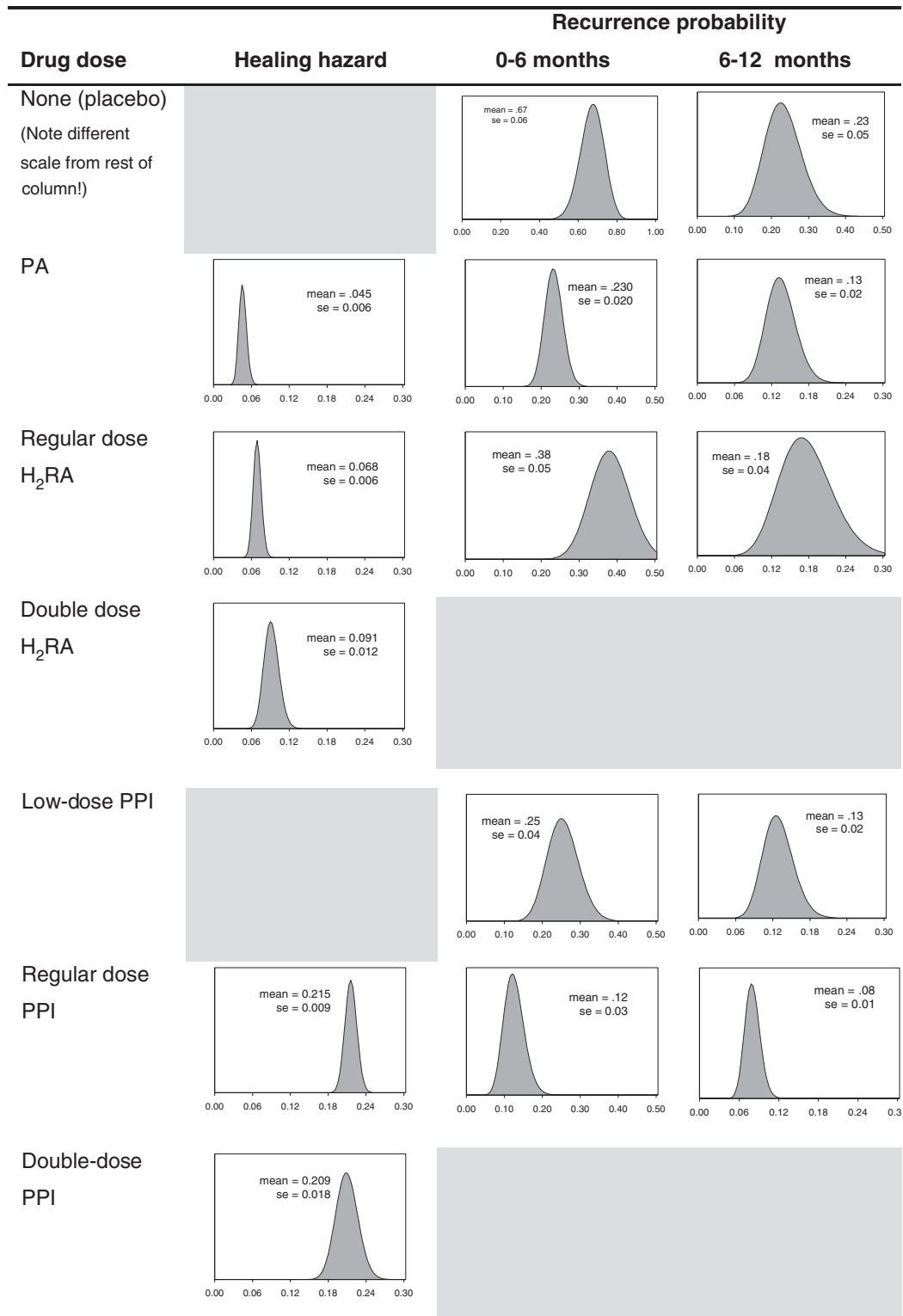
Table 2. Parameters in the Model

Name	Value	Description
Healing hazards		
hPA	0.045	Hazard for healing on PA
hH2	0.068	Hazard for healing on H_2RAs
hDDH2	0.091	Hazard for healing on double-dose H_2RAs
hPPI	0.215	Hazard for healing on PPIs
hDDPPI	0.209	Hazard for healing on double-dose PPIs
Corresponding transition probabilities		
pPA	0.42	Probability of healing on PA (at 12 weeks)
pH2	0.42	Probability of healing on H_2RAs (at 8 weeks)
pDDH2	0.52	Probability of healing on double-dose H_2RAs (at 8 weeks)
pPPI	0.82	Probability of healing on PPIs (at 8 weeks)
pDDPPI	0.81	Probability of healing on double-dose PPIs (at 8 weeks)
Recurrence variables		
p06PL	0.67	Probability of recurrence on placebo (0-6 months)
p06PA	0.23	Probability of recurrence on PA (0-6 months)
p06H2	0.38	Probability of recurrence on H_2RAs (0-6 months)
p06PPI	0.12	Probability of recurrence on PPIs (0-6 months)
p06LDPPI	0.25	Probability of recurrence on low-dose PPIs (0-6 months)
p06SU	0.12	Probability of recurrence after surgery (0-6 months)
p612PL	0.23	Probability of recurrence on placebo (0-6 months)
p612PA	0.13	Probability of recurrence on PA (6-12 months)
p612H2	0.18	Probability of recurrence on H_2RAs (6-12 months)
p612PPI	0.08	Probability of recurrence on PPIs (6-12 months)
p612LDPPI	0.13	Probability of recurrence on low-dose PPIs (6-12 months)
Symptom-week variables		
SWPA	2.280	Symptom weeks on PA
SWH2	1.521	Symptom weeks on H_2RAs
SWDDH2	1.816	Symptom weeks on double dose of H_2RAs
SWPPI	2.386	Symptom weeks on PPIs
SWDDPPI	2.383	Symptom weeks on double dose of PPIs
SWSU	0.539	Symptom weeks after surgery

Table 2 Continued

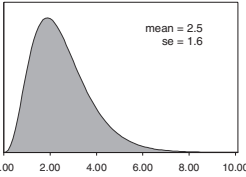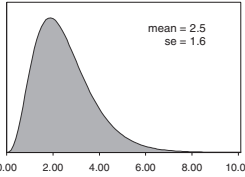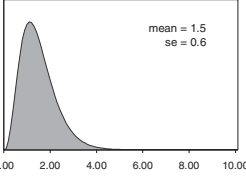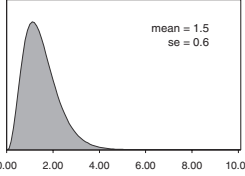| Name	Value	Description
Unit cost variables		
Fee	$4.11	Dispensing fee
H_2RA	0.44	150 mg H_2RA
CIS	0.61	10 mg prokinetic agent
PPI	2.42	20 mg PPI
LDPPI	1.93	10 mg PPI
GPGA	$48.20	General practitioner general assessment
GPRA	$28.10	General practitioner reassessment
GPMA	$16.25	General practitioner minor assessment
GERA	$38.65	Gastroenterologist reassessment
GEPA	$23.10	Gastroenterologist partial assessment
UGIE	$118.22	Upper gastrointestinal endoscopy
UGIS	$141.43	Upper gastrointestinal series
CST	$84.81	Cardiac stress test
ECG	$42.77	Electrocardiogram
BS	$135.10	Barium swallow
NF	$2462.60	Laparoscopic fundal plication
Resource use variables		
nGPR1	2.5	Visits to general practitioner (1st recurrence)
nGER1	1.5	Visits to gastroenterologist (1st recurrence)
nBSR1	0.1	Percentage getting a barium swallow (1st recurrence)
nCSTR1	0.0	Percentage getting cardiac stress test (1st recurrence)
nECGR1	0.0	Percentage getting electrocardiogram (1st recurrence)
nUGIER1	0.1	Percentage getting upper gastro-intestinal endoscopy (1st recurrence)
nUGISR1	0.1	Percentage getting upper gastro-intestinal series (1st recurrence)
nGPR2	2.5	Visits to general practitioner (2nd recurrence)
nGER2	1.5	Visits to gastroenterologist (2nd recurrence)
nBSR2		Percentage getting a barium swallow (2nd recurrence)
nCSTR2	0.0	Percentage getting cardiac stress test (2nd recurrence)
nECGR2	0.0	Percentage getting electrocardiogram (2nd recurrence)
nUGIER2	0.6	Percentage getting upper gastro-intestinal endoscopy (2nd recurrence)
nUGISR2	0.1	Percentage getting upper gastro-intestinal series (2nd recurrence)
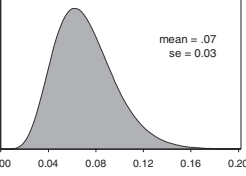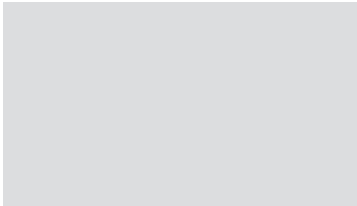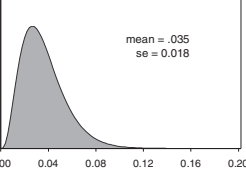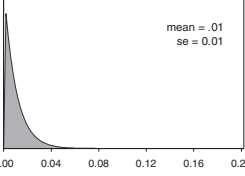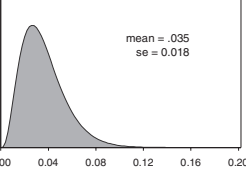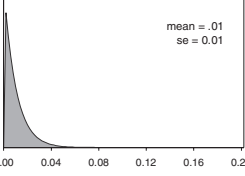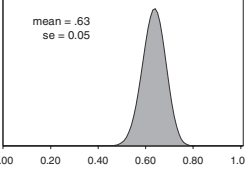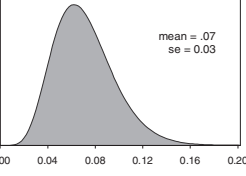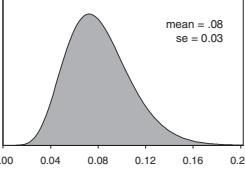
Note: PPI = proton pump inhibitor; H_2RA = H_2-receptor antagonist; PA = prokinetic agent.

Table 3. Distributions for Healing Hazards and Recurrence Probabilities

Drug dose	Healing hazard	Recurrence probability	
		0-6 months	6-12 months
None (placebo) (Note different scale from rest of column!)		mean = .67, se = 0.06	mean = .23, se = 0.05
PA	mean = .045, se = 0.006	mean = .230, se = 0.020	mean = .13, se = 0.02
Regular dose H$_2$RA	mean = 0.068, se = 0.006	mean = .38, se = 0.05	mean = .18, se = 0.04
Double dose H$_2$RA	mean = 0.091, se = 0.012		
Low-dose PPI		mean = .25, se = 0.04	mean = .13, se = 0.02
Regular dose PPI	mean = 0.215, se = 0.009	mean = .12, se = 0.03	mean = .08, se = 0.01
Double-dose PPI	mean = 0.209, se = 0.018		

Note: PA = prokinetic agent; H$_2$RA = H$_2$-receptor antagonist; PPI = proton pump inhibitor.

Table 4. Distributions for Resource Use Parameters

Resource item	First recurrence	Second recurrence
Visits to GP *(Gamma distribution)*	mean = 2.5, se = 1.6	mean = 2.5, se = 1.6
Visits to gastroenterologist *(Gamma distribution)*	mean = 1.5, se = 0.6	mean = 1.5, se = 0.6
Proportion getting a barium swallow	mean = .07, se = 0.03	
Proportion getting cardiac stress test	mean = .035, se = 0.018	mean = .01, se = 0.01
Proportion getting ECG	mean = .035, se = 0.018	mean = .01, se = 0.01
Proportion getting upper GI endoscopy	mean = .10, se = 0.03	mean = .63, se = 0.05
Proportion getting upper GI series	mean = .07, se = 0.03	mean = .08, se = 0.03

Note: GP = general practitioner; ECG = electrocardiogram; GI = gastrointestinal.

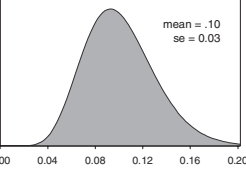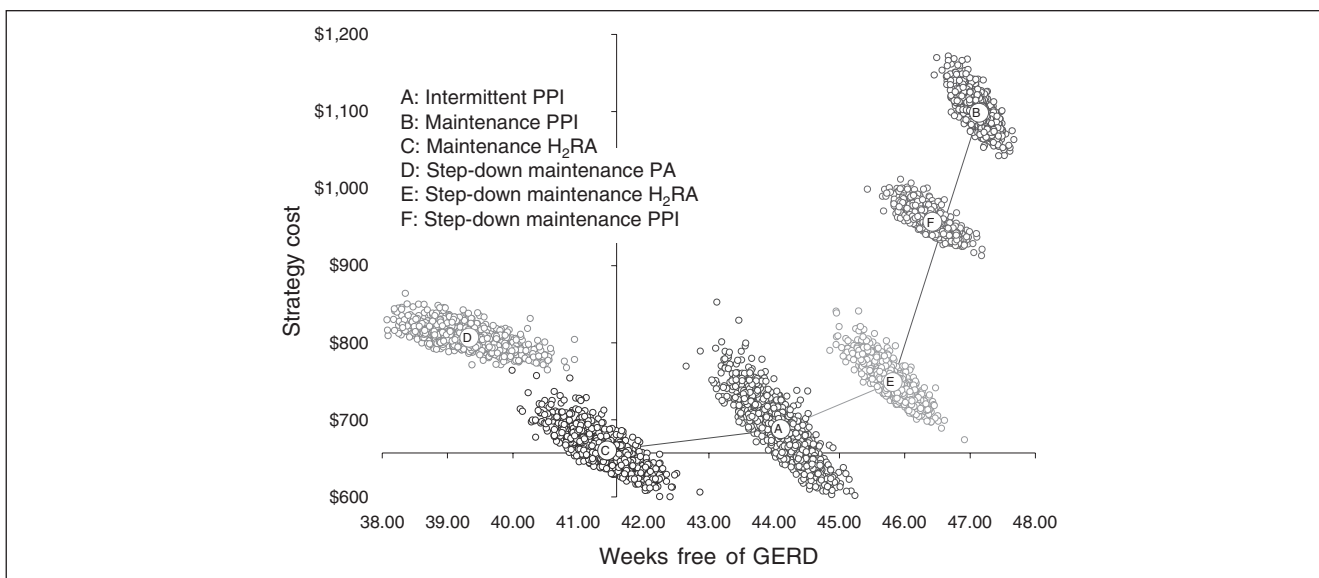*Figure 3. Results of 10,000 Monte Carlo simulation evaluations of the gastroesophageal reflux disease (GERD) model presented on the cost-effectiveness plane. PPI = proton pump inhibitor; H_2RA = H_2-receptor antagonist; PA = prokinetic agent.*

true opportunity cost of those resources, it is not clear how such uncertainty could be represented in this model. Therefore, all unit costs were taken as being deterministic and were not ascribed distributions in this analysis. Although there are some problems with this approach,[24] this is not conceptually different from the approach taken in stochastic cost-effectiveness analysis alongside clinical trials where it is typical for unit costs to be treated as fixed.

Results of the Probabilistic Analysis

Having specified distributions for all the relevant parameters of the model, the probabilistic analysis was undertaken by randomly sampling from each of the parameter distributions and calculating the expected costs and expected weeks free of GERD for that combination of parameter values. This process formed a single replication of the model results, and a total of 10,000 replications were performed in order to examine the distribution of the resulting cost and outcomes for each strategy. The results of these 10,000 replications from the model are presented on the cost-effectiveness plane in Figure 3 together with the baseline estimate of the efficient frontier.

It is clear that for each of the individual replications, an efficient frontier could be calculated together with the incremental cost-effectiveness ratios for treatments on the frontier. In particular, Figure 3 shows how it may not be possible to rule out strategy F, the strategy based

Table 5. Percentage of Times Each Strategy Formed Part of the Cost-Effectiveness Efficiency Frontier

Strategy	Percentage of Times on the Frontier
A	72
B	100
C	76
D	0
E	98
F	27

on step-down maintenance PPI, since it potentially forms part of the frontier in many replications. Note, however, that it is not possible to gain a clear view from Figure 3 as to how often strategy F forms part of the frontier: this is because there can be substantial covariance between the simulations plotted in the figure. Table 5 summarizes the proportion of the 10,000 simulations in which each strategy forms part of the cost-effectiveness frontier. It is immediately clear from this table (as it was from Figure 3) that strategy D can be ruled out. By contrast, it turns out that strategy F forms part of the frontier in 27% of simulations, and it is not clear how this result should be interpreted.

If the shadow price for a week free of GERD symptoms (the maximum willingness to pay or 'ceiling ratio') were known, it would be possible to choose between all of the treatment strategies, not just identify those that form the efficient frontier. Therefore, condi-
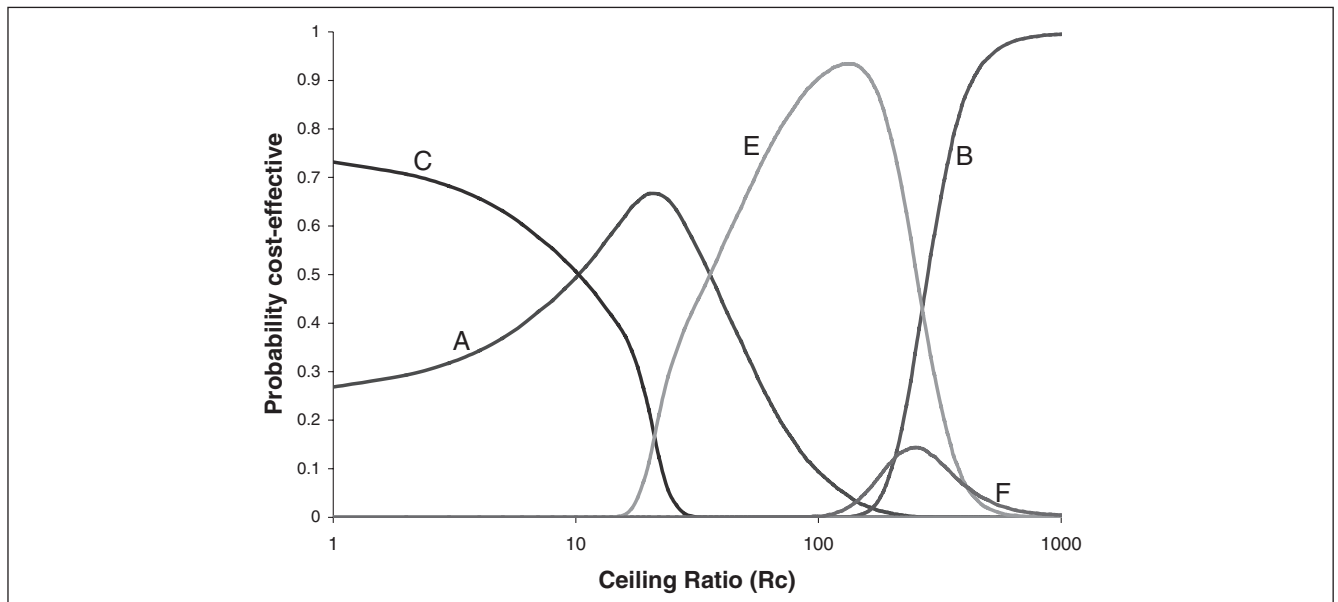
Figure 4.   Acceptability curves for the choice of treatment strategy (a log scale is employed to better illustrate the low values).

tional upon knowing the ceiling ratio, there is only 1 treatment of choice from the 6 strategies under evaluation, and the proportion of times that an intervention is the treatment of choice from the 10,000 replications of the model gives the strength of evidence in favor of that treatment. Although it is possible to identify the efficient frontier, calculate the incremental cost-effectiveness ratios, and choose 1 strategy from the 6 available for each of the 10,000 replications, a much more straightforward approach exists.

The net benefit framework has been argued to offer many advantages for handling uncertainty in cost-effectiveness analysis[25] and overcomes the particular problem associated with negative incremental cost-effectiveness ratios.[26] A further property is that whereas average cost-effectiveness ratios have no meaningful interpretation, average net benefits have the useful property that the incremental net benefit between any 2 treatments can be calculated from the difference between their individual average net benefits.[25] Therefore, the treatment of choice from the 6 strategies under evaluation will be the treatment with the greatest average net benefit. This must be the case, as only that treatment will have a positive incremental net benefit when compared to any other treatment alternative. The proportion of times a strategy has the highest net benefit among the 10,000 replications of the model gives the strength of evidence in favor of that strategy being cost-effective. This ability of the net benefit framework to

handle multiple mutually exclusive treatment options is a very strong advantage of the approach, as pointed out by the authors in their original article.[25]

Of course, in reality the shadow price of a week free of GERD symptoms is not known. However, by plotting out, for all possible values of the ceiling ratio (R_c), the proportion of times the intervention has the greatest net benefit, much can be learned concerning the implications of the estimated uncertainty for the treatment decision. Figure 4 shows the result of just such an exercise for the probabilistic evaluation of the GERD model presented in Figure 3. These curves are conceptually the same as the use of acceptability curves to summarize uncertainty on the cost-effectiveness plane in a 2-treatment decision problem.[2]

As expected, strategy D does not feature in Figure 4, indicating that it is never a contender for cost-effectiveness. Strategy F does feature, although it never achieves more than 13% of simulations, suggesting it is cost-effective, even at the most favorable ceiling ratio (about $260 per day free of GERD symptoms). In these situations, Stinnett and Mullahy[25] suggested that we look to the principle of stochastic dominance in order to rule out interventions. Because average net benefits are separable (in contrast to cost-effectiveness ratios), if it can be shown that the cumulative distribution of one treatment option is always greater than that of another option, the original treatment is said to dominate. This means that a utility-maximizing decision maker
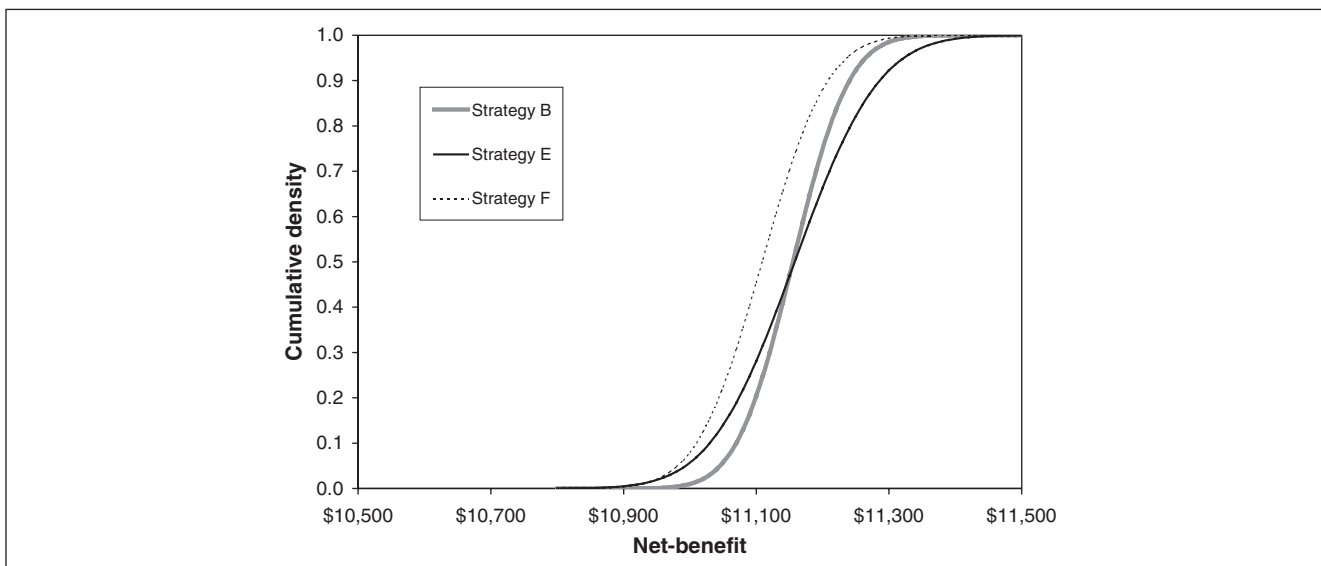
Figure 5.   Cumulative density functions over net benefit (R_c = $260 per day free of gastroesophageal reflux disease symptoms).
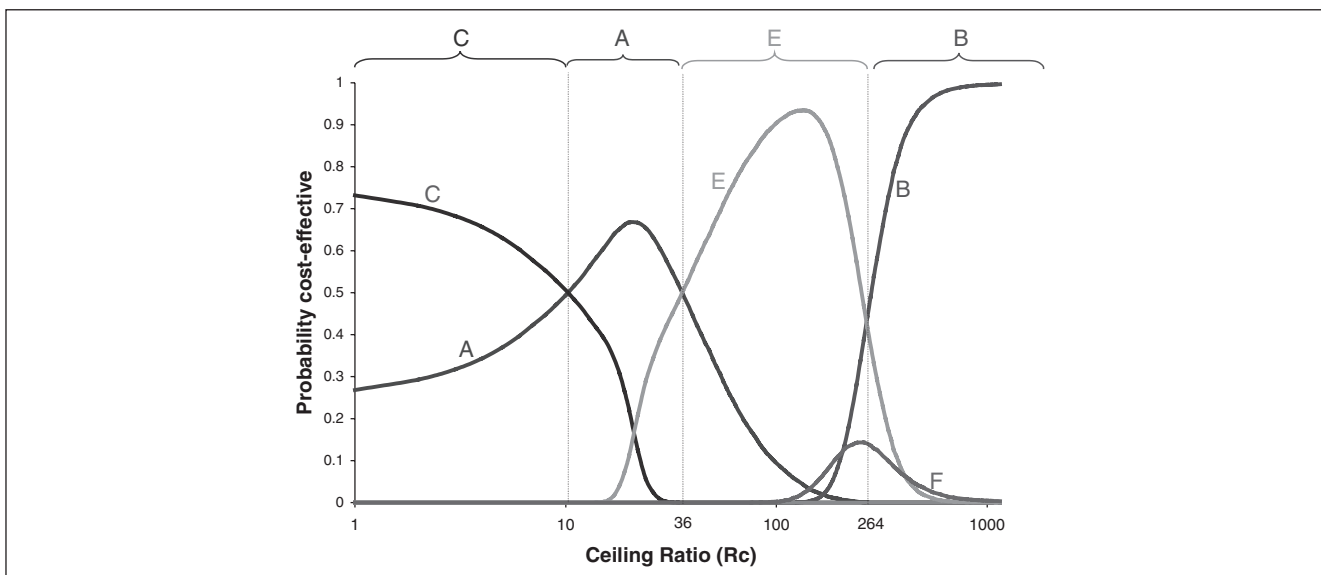

Figure 6.   Decision making using management strategy "most likely" to be cost-effective.

should always prefer the 1st strategy to the 2nd no matter what their risk preference. This condition is known as 1st-order stochastic dominance. A slightly weaker form is 2nd-order stochastic dominance, which allows the cumulative density functions over net benefit to cross, but holds that a treatment dominates another if the area under its cumulative density function is always greater than that under the alternative strategy. In Figures 5 and 6, we take the most favorable value of the ceiling ratio from the perspective of strategy F and plot the cumulative density functions of strategies B, E, and F. It turns out that we cannot quite rule out strategy F under 1st-order stochastic dominance, but strategy F is very clearly ruled out under the 2nd-order condition.

DISCUSSION

Probabilistic sensitivity analysis methods have been proposed for decision analysis and cost-effectiveness models for some time. Despite this, the vast majority of
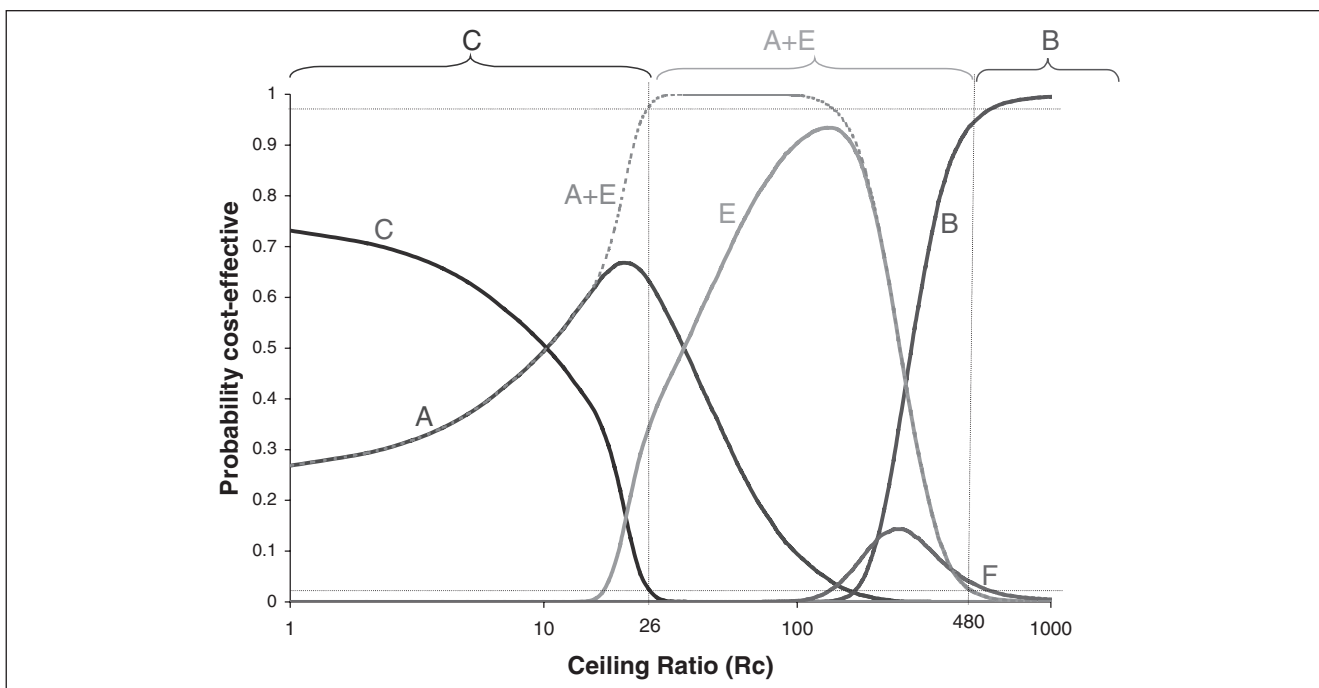
Figure 7.   Decision making assuming more expensive strategies have to be shown to be significantly more cost-effective than the current strategy (at conventional 5% level).

published economic evaluations do not make use of probabilistic methods. Although a number of authors have begun to emphasize the advantages of adopting a Bayesian perspective in probabilistic modeling,[27–29] many probabilistic analyses do not adopt a Bayesian approach, and for such studies the choice of distributions to represent uncertainty in parameters may appear rather ad hoc.[30–36] Furthermore, the Bayesian approach may afford a more intuitive interpretation of the probabilistic sensitivity results that has been highlighted as lacking in a review of the use of medical decision analysis models.[37]

Adopting a Bayesian approach to the probabilistic analysis of the model of GERD management allows the intuitive interpretation often afforded to acceptability curves as showing the probability that the intervention is cost-effective. Note that the curves all sum to 1 on the vertical axis (this clearly must be the case, since only 1 strategy is chosen for each value of the ceiling ratio and for each replication of the model). It is immediately apparent from Figures 3 and 4 that strategy D is always dominated. The acceptability curves in Figure 4 clearly show that initial concern that strategy F might form part of the frontier and might therefore be a cost-effectiveness choice in some situations was unwarranted. In fact, the conditions necessary for strategy F to

be considered the most cost-effective option rarely arose in the simulations. Appealing to the principle of stochastic dominance (2nd-order) shows strategy F to be clearly rejected even under the most favorable assumptions concerning the ceiling ratio (Figs. 5, 6). This illustrates how by employing probabilistic methods that directly quantify the strength of evidence in favor of alternative treatment options, it is possible to be confident about excluding particular strategies, in contrast to the traditional sensitivity analysis approaches.

Although this sort of presentation of the choice between mutually exclusive treatments in the face of many options is a natural extension of the use of cost-effectiveness acceptability curves in the 2-treatment case, the issue arises of how exactly decision makers should use this information to choose between the remaining strategies that form part of the frontier. One approach (as illustrated in Fig. 7) would be to say that, for any given value of the ceiling ratio, the optimal decision would be to choose the strategy that is most likely to be cost-effective. But, of course, this decision rule gives the exact same treatment recommendations as the baseline estimates in Figure 2, where uncertainty was not considered.

The conventional approach to statistical decision making is based on the adoption of a 5% type I error

rate. We might therefore adopt a decision rule that a more effective and more expensive treatment strategy should replace the currently provided treatment only if it can be shown to be significantly more cost-effective. This approach to decision making is illustrated in Figure 7 and gives markedly different cutoff points for decision maker's ceiling ratios for the different strategies to be considered cost-effective. For example, the most effective therapy, maintenance PPI, would only be considered the appropriate treatment option if decision makers had an underlying willingness to pay of $480 per week free of GERD symptoms, assuming that F is not part of the choice set.

Note that under a 5% significance decision rule, neither strategy A nor E would be considered significantly cost-effective to be a clear decision choice. However, it is clear that strategy C is not cost-effective (at the 5% error rate) for a shadow price greater than $26 per week free of GERD symptoms. Recall that strategy A involves healing with PPI without maintenance and strategy E involves healing with PPI and then maintenance using H_2RAs. The choice between these strategies is between no maintenance and maintenance with H_2RAs following healing. If the shadow price of a week free of GERD symptoms is between $26 and $480 per week free of GERD symptoms, then we know that strategies B, C, D, and F are not cost-effective. We cannot distinguish strategies A and E at conventional significance levels, so for purposes of decision making it should be clear that the strategy to heal with PPI should be adopted, but that it is unlikely to be important whether, subsequent to healing, no maintenance or maintenance with H_2RAs is undertaken. This is illustrated in Figure 7 by combining the 2 strategies in 1 acceptability curve.

It is important to recognize, however, the arbitrary nature of the conventional (and frequentist) decision rule. Consider whether instead of placing the "burden of proof" for cost-effectiveness on more expensive and more effective strategies, it is the cheaper but less effective strategies that would be used only if they were shown to be significantly cost-effective. Although not shown in Figure 7, it should be clear from the above exposition that this change in the burden of proof would result in a new set of threshold values such that strategy B would be the treatment of choice unless the shadow price were below $170 per week free of GERD symptoms; between $17.50 and $170 per week free of GERD symptoms, either strategy A or E would be considered cost-effective; and below a shadow price of $17.50, either A or C would be the strategy of choice.

Of course the arbitrary nature of such decision making under uncertainty emphasizes the inadequacies of such a simple decision rule—Claxton[38] argued that significance testing of this sort is irrelevant. Instead, he suggested that decision making be concerned fundamentally with expected values. That is not to say that the decisions should be made on the basis of the baseline point estimates as presented in Figure 2 without reference to uncertainty in obtaining those estimates. Rather, the expected returns to obtaining further information should be assessed in order to determine whether it is worth commissioning more research to obtain improved estimates of the decision parameters. Such an approach would require estimates of the loss function associated with incorrect decision making, the size of the population relevant to the decision, the lifetime of the technologies associated with each management strategy, and the returns to sampling. Each is itself subject to a great deal of uncertainty, and the methods for incorporating all this information into a single overall analysis are currently under development. We have not attempted such an analysis in this article, although we recognize that it is a possible extension to the work presented here and that such an analysis would be considered a fully Bayesian decision model.

In summary, probabilistic modeling of deterministic models is a practical solution to the problems of conventional sensitivity analysis. Adopting a Bayesian approach encourages analysts and users to think carefully about the state of evidence relating to the parameters of the model. Single-parameter specifications are straightforward to apply in a Bayesian framework and provide a simple way to update parameter distributions as new data become available. The use of acceptability curves to present information on the probability of multiple treatment options is a natural extension of the 2-alternative case usually presented in the literature. Much current research interest is focused on expected value of information methods. It is clear that such methods will have to be predicated on a well-specified probabilistic model.

APPENDIX

The model that formed the basis of this article is available for download from the World Wide Web at www.ihs.ox. ac.uk/herc/downloads.

Two formats are available: a spreadsheet model (Microsoft Excel 2000) and a TreeAge DATA™ model (version 4.0). Both models use the same parameters and naming conventions reported in this article and both are fully probabilistic. Full details of the modeling assumptions and parameters are given in the original (deterministic) article,[9] and readers are advised to familiarize themselves with this article to aid understanding of the model structure.

Note that the model implemented in TreeAge DATA™ is complete and can be run as either a deterministic model or a probabilistic model. An evaluation copy of DATA is available from www.treeage.com.

The spreadsheet model implemented in Excel has been stripped of the simulation results to reduce the file size and to facilitate speed of downloading. When opening the file, the simulation results can be generated by pressing the "Run Monte Carlo Simulation" button on the 1st page, which employs a visual basic macro to record simulation results. Note that no calculations are embodied in this macro—the macro simply implements a loop to repeatedly copy the results of calculations into the spreadsheet.

The purpose of this appendix is to describe in more detail the fitting of distributions in the probabilistic modeling exercise and the sampling from those distributions in relation to the 2 software platforms employed to illustrate the model.

Beta Distribution

As described in the article, obtaining parameters for a beta(r,n) distribution is straightforward when modeling a probability from binomial data because the parameters have an interpretation as r = events and n = total sample size. This is the default parameterization of the beta distribution in DATA; however, this only works with integer values of r and n. For the method-of-moments fitting described in this article, it is necessary to choose the "real numbered parameters" version of the beta distribution: beta(α, β), where $\alpha = r$ and $\alpha + \beta = n$.

Method-of-moments fitting involves equating the means and variances observed in the data to the expressions for the mean and variance of the distribution and solving for the parameters of that distribution. Therefore, for observed mean, $\bar{\mu}$, and standard error, s, we formulate the expressions

$$\bar{\mu} = \frac{r}{n} \qquad s^2 = \frac{r(n-r)}{n^2(n+1)}$$

and rearrange to give

$$n = \frac{\bar{\mu}(1-\bar{\mu})}{s^2} - 1$$
$$r = \bar{\mu}n$$

So, for example, the instantaneous hazard for the prokinetic agent in the model was estimated to be 0.045 with a standard error of 0.006. Using the above expressions generates parameters for the beta distribution of approximately $n = 1194$ and $r = 54$, or $\alpha = 54$ and $\beta = 1140$ (these are approximate due to rounding—the models use full accuracy).

Gamma Distribution

Parameters involving numbers of visits to health care professionals (primary care physicians and gastroenterologists) were modeled using the gamma distribution, which has 2 parameters, α and β. The same method-of-moments approach was used to obtain these parameters. The mean and variance of the gamma(α, β) distribution are given by

$$mean = \frac{\alpha}{\beta}, \qquad var = \frac{\alpha}{\beta^2}.$$

Setting these equal to the corresponding estimates, and then rearranging, gives expressions for the parameters

$$\alpha = \frac{\bar{\mu}^2}{s^2}, \qquad \beta = \frac{\bar{\mu}}{s^2}.$$

For example, visits to the general practitioner after a recurrence of GERD are estimated to occur with a mean 2.5 visits and a standard error of 1.6, which is modeled in the DATA model as gamma(4,1.6).

Note, however, that the parameterization of the gamma distribution in Excel is slightly different from that commonly found in other packages. Excel uses exactly the same distribution but has parameterized β as $1/\beta$, which means that in Excel the appropriate version of the expression for the parameters is

$$\alpha = \frac{\bar{\mu}^2}{s^2}, \qquad \beta = \frac{s^2}{\bar{\mu}}.$$

For the same example of general practitioner visits after recurrence of GERD, the same distribution modeled in Excel is gamma(4,0.625).

Generating Random Draws from the Specified Distribution

In DATA, there is no need to be concerned about drawing random values from the specified distribution because, having specified the parameters, the software takes care of simulation. In Excel (and many other software packages), there is no direct simulation function. However, it is straightforward to use the distribution functions provided together with the random number generator function to obtain random draws. The upper panel of Figure A1 shows a cumulative distribution function (cdf) for a (standard normal) distribution—the
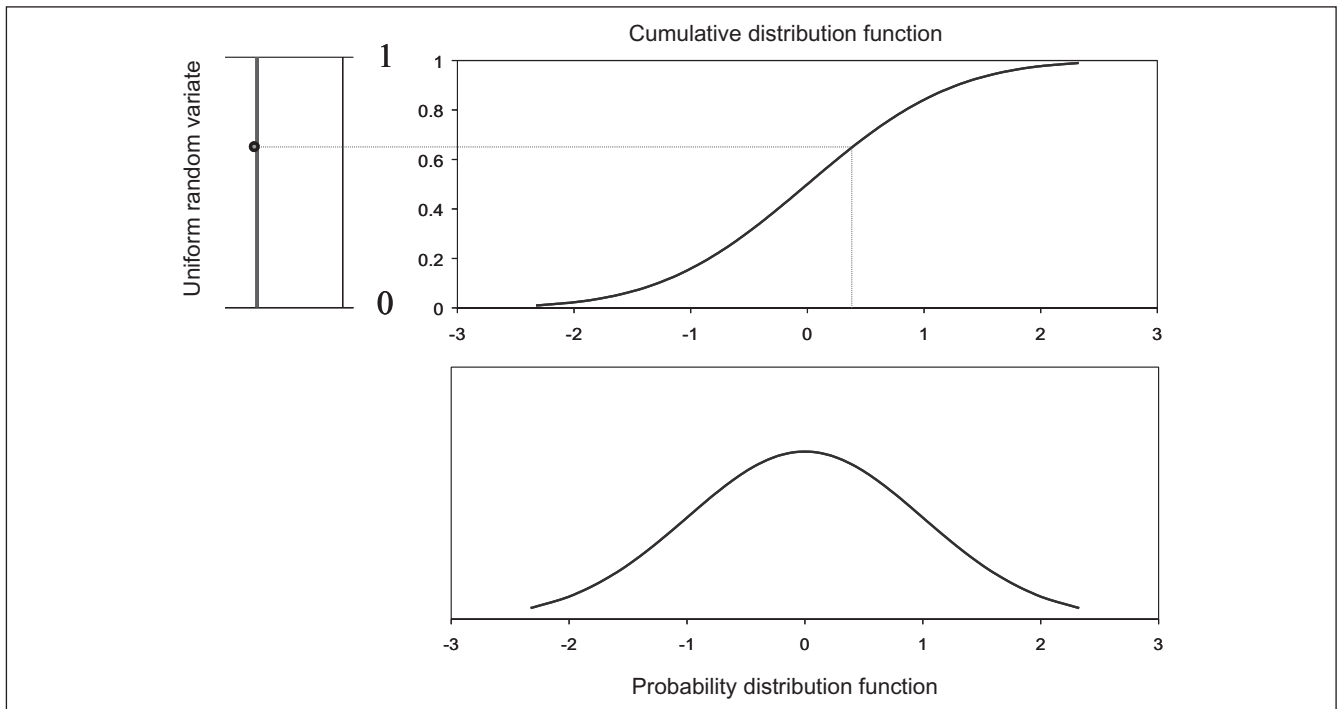
Figure A1. Random draws from a parametric distribution.

values of the distribution are shown on the horizontal axis, and the vertical axis shows the proportion of the probability density function (pdf) falling below that value. The vertical access therefore has a scale of 0 to 1, and the cdf can be thought of as a function that maps the 0-to-1 scale onto the scale of the pdf. We can use the cdf together with a uniform random variate to obtain a random draw from the pdf simply by generating a random value between 0 and 1 and reading across from the vertical access to the curve to obtain a value from the pdf (see Fig. A1). Repeating this process a large number of times generates random values from the pdf shown at the bottom of Figure A1.

In Excel, the required functions for the cdfs of distributions have an INV suffix (because we are using the inverse of the cdf) and each function has 3 arguments—for example,

BETAINV(p, α, β) and GAMMAINV(p, α, β)–and the function returns a value, q. The 1st argument represents the proportion of the pdf up to the value q, and the 2 remaining arguments are the parameters of the distribution function. The function RAND() in Excel generates a uniform random variate on the interval 0 to 1. Hence, to obtain a random draw from the beta and gamma distributions specified in the preceding sections, we use the BETAINV(RAND(),54,1140) and GAMMAINV (RAND(),4,0.625) in Excel.

Note that a 0-1 switch is used in the Excel model to allow the user to choose between a probabilistic and a deterministic version of the model. This switch is located in cell B3 of the <Parameters> worksheet. Setting this cell to 1 allows the user to observe simulated values in real time—pressing the <F9> key draws another set of random values for the parameters.

REFERENCES

1. O'Brien BJ, Drummond MF, Labelle RJ, Willan A. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. Med Care. 1994;32:150–63.

2. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ. 1994;3:309–19.

3. Stinnett AA, Paltiel AD. Estimating CE ratios under second-order uncertainty: the mean ratio versus the ratio of means. Med Decis Making. 1997;17:483–9.

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

5. Critchfield GC, Willard KE, Connelly DP. Probabilistic sensitivity analysis methods for general decision models. Computers Biomed Res. 1986;19:254–65.

6. Doubilet P, Begg CB, Weinstein MC, et al. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. Med Decis Making. 1985;5:157–77.

7. Briggs AH. Handling uncertainty in cost-effectiveness models. PharmacoEcon. 2000;17:479–500.

8. Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost-effectiveness analysis. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press, 1996. p. 247–75.

9. Goeree R, O'Brien B, Hunt R, Blackhouse G, Willan A, Watson J. Economic evaluation of long term management strategies for erosive oesophagitis. PharmacoEcon. 1999;16:679–97.

10. Chiba N, de Gara CJ, Berget D. Rapidity of healing in GERD [abstract]. Gastroenterology. 1993;104:A53.

11. Chiba N, Wilkinson J, Hunt RH. Symptom relief in erosive GERD [abstract]. Am J Gastroenterol. 1993;88:A1486.

12. Ministry of Health. Drug Benefit Formulary: Comparative Drug Index no. 35. Toronto: Publications Ontario, 1996.

13. Ontario Ministry of Health. Schedule of Benefits: Physician Services under the Health Insurance Act. Toronto: Publications Ontario, 1992.

14. Ontario Hospital Association. Ontario Case Costing Project: Ontario Guide to Case Costing. Toronto: Publications Ontario, 1995.

15. Park RE, Fink A, Brook RH. Physician ratings of appropriate indications for six medical and surgical procedures. Am J Public Health. 1986;76:766–72.

16. Weinstein MC. From cost-effectiveness ratios to resource allocation: where to draw the line? In: Sloan FA, ed. Valuing Health Care. Cambridge, UK: Cambridge University Press, 1995. p. 77–98.

17. Gold MR, Siegel JE, Russell LB, et al. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press, 1996.

18. Drummond MF, O'Brien B, Stoddart GL, et al. Methods for the Economic Evaluation of Health Care Programmes. 2nd ed. Oxford, UK: Oxford University Press, 1997.

19. Cantor SB. Cost-effectiveness analysis, extended dominance, and ethics: a quantitative assessment. Med Decis Making. 1994;14:259–65.

20. Eddy DM, Hasselblad V, Shachter R. A Bayesian method for synthesizing evidence: the confidence profile method. Int J Technol Assess Health Care. 1990;6(1):31–55.

21. Eddy DM, Hasselblad V, Shachter R. An introduction to a Bayesian method for meta-analysis: the confidence profile method. Med Decis Making. 1990;10(1):15–23.

22. Altman DG. Practical Statistics for Medical Research. London: Chapman & Hall, 1991.

23. Pratt JW, Raiffa H, Schlaifer R. Introduction to Statistical Decision Theory. Cambridge, MA: MIT Press, 1995.

24. Rittenhouse BE, Dulisse B, Stinnett AA. At what price significance? The effect of price estimates on statistical inference in economic evaluation. Health Econ. 1999;8:213–9.

25. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making. 1998;18(2 suppl):S65–S80.

26. Stinnett AA, Mullahy J. The negative side of cost-effectiveness analysis [letter; comment]. JAMA. 1997;277:1931–2.

27. Parmigiani G, Samsa GP, Ancukiewicz M, Lipscomb J, Hasselblad V, Matchar DB. Assessing uncertainty in cost-effectiveness analyses: application to a complex decision model. Med Decis Making. 1997;17:390–401.

28. Sendi PP, Craig BA, Meier G, Pfluger D, Gafni A, Opravil M, Battegay M, Bucher HC. Cost-effectiveness of azithromycin for preventing Mycobacterium avium complex infection in HIV-positive patients in the era of highly active antiretroviral therapy: the Swiss HIV Cohort Study. J Antimicrob Chemother. 1999;44:811–7.

29. Fryback DG, Chinnis JO, Ulvila JW. Bayesian cost-effectiveness analysis: an example using the GUSTO trial. Int J Technol Assess Health Care. 2001;17(1):83–97.

30. Berwick DM, Cretin S, Keeler E. Cholesterol, children, and heart disease: an analysis of alternatives. Pediatrics. 1981;68:721–30.

31. Hornberger JC. The hemodialysis prescription and cost effectiveness. J Am Soc Nephrol. 1993;4:1021–7.

32. Gabriel SE, Campion ME, O'Fallon WM. A cost-utility analysis of misoprostol prophylaxis for rheumatoid arthritis patients receiving nonsteroidal antiinflammatory drugs. Arthritis Rheum. 1994;37:333–41.

33. Pharoah PD, Hollingworth W. Cost effectiveness of lowering cholesterol concentration with statins in patients with and without pre-existing coronary heart disease: life table method applied to health authority population. Br Med J. 1996;312:1443–8.

34. Fiscella K, Franks P. Cost-effectiveness of the transdermal nicotine patch as an adjunct to physicians' smoking cessation counseling [see comments]. JAMA. 1996;275:1247–51.

35. Oh PI, Maerov P, Pritchard D, Knowles SR, Einarson TR, Shear NH. A cost-utility analysis of second-line antibiotics in the treatment of acute otitis media in children. Clin Ther. 1996;18:160–82.

36. Pasta DJ, Taylor JL, Henning JM. Probabilistic sensitivity analysis incorporating the bootstrap: an example comparing treatments for the eradication of Helicobacter pylori. Med Decis Making. 1999;19:353–63.

37. Sonnenberg FA, Roberts MS, Tsevat J, Wong JB, Barry M, Kent DL. Toward a peer review process for medical decision analysis models. Med Care. 1994;32(7 suppl):JS52–JS64.

38. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999;18:341–64.