*Work in progress*

**Cost-Effectiveness Analyses: A Stochastic Programming Approach for Optimal Decision-Making**

**Zaid Chalabi*  and Jack Dowie**

**Department of Public Health and Policy**

**London School of Hygiene and Tropical Medicine**

**Keppel Street, London WC1E 7HT, United Kingdom**

## Abstract

There has been an increasing interest recently in stochastic cost-effectiveness analyses (CEA) of healthcare interventions. The objective of CEA is to determine the *optimal* intervention (or subset of interventions) – from a specified set of interventions – where *optimality* is defined in some (often *weak*) sense. CEA involve two components: models to establish the costs and effects of the interventions concerned and decision rules to determine the optimal intervention given the costs and effects. It has been shown that Bayesian hierarchical models are best for modelling under (inevitable) uncertainty and that Bayesian decision techniques are most appropriate for delivering decision-rules that yield the optimal intervention. It is argued in this paper that debates on what constitutes the optimal intervention in the presence of uncertainty can be resolved most transparently if the decision problem is cast in a mathematical framework where optimality is defined in a *strong* sense. A new Bayesian stochastic framework is therefore proposed for CEA which combines the hierarchical modelling approach with the mathematical programming approach for optimal decision-making. It provides a logical and a transparent platform for representing, propagating and handling uncertainty for optimal decision-making. Some of the main practical implications are explored.

* Author for correspondence
Tel: +44 20 7927 2453; fax: +44 20 7580 4524
Email address: zaid.chalabi@lshtm.ac.uk

## 1. Introduction

Several methods have been proposed for handling uncertainty in Cost-Effectiveness Analyses (CEA) of healthcare interventions (Stinnett and Mullahy, 1998; Briggs and Gray, 1999; Briggs, 2000; Spiegelhalter *et al.*, 2000; Briggs *et al.*, 2002; Sendi *et al.*, 2002). Most of these methods consist of mathematical models to evaluate the costs and effects of interventions and mathematical decision rules to determine the *optimal* intervention from a defined set of interventions. *Optimality* however is often defined in what, mathematically speaking, is a non-rigorous or *weak* sense.

It is now acknowledged that it is necessary to represent and propagate the uncertainty in CE models in a manner which exposes the effect of the uncertainty on the decision rules. One of the many reasons in the low uptake of CEA results by decision-makers may be the difficulty in communicating and interpreting results which are heavily embedded in model uncertainty (Hutubessy *et al.*, 2001).

The Bayesian approach has become the accepted approach to CEA under uncertainty (Stinnett and Mullahy, 1998; Briggs, 1999; Claxton *et al.*, 2001; O'Hagan *et al.*, 2001; Briggs *et al.*, 2002; Spiegelhalter and Best, 2002). From a mathematical perspective, this approach is justified because of the nature of the uncertainty in CEA models. The models are best constructed in a hierarchical manner where both model parameters and model variables are regarded equally as statistical variables whose probability distributions are to be modelled.

Although there have been much interest in the use of Bayesian decision theoretic approaches in CEA (Claxton *et al.*, 2000), relatively little attention has been paid to the use of formal optimisation methods in the decision technology. Furthermore, the focus in CEA has in general been more on *optimal modelling* than *optimal decision-making* per se. This imbalance between modelling and optimisation, in relation to handling uncertainty, is evident in the use of Cost-Effectiveness Acceptability Curves (CEACs) to represent uncertainty in economic evaluations of healthcare interventions (Van Hout *et al.*, 1994; Lothgren and Zethraeus, 2000; O'Hagan *et al.*, 2000; Fenwick *et al.*, 2001; Spiegelhalter and Best, 2002). It will be argued below that the mathematical formulation of CEAC – particularly in the case of the multiple-interventions decision problem – obscures some of the key issues concerning optimality under uncertainty.

Although there have been some attempts recently to formulate CEA problems as *strong* optimisation problems using mathematical programming techniques (Stinnett and Paltiel, 1996), the resulting mathematical formulations suffer from two main weaknesses:

(i) They do not take into account the uncertainty in the costs and the effects of the interventions.

(ii) They assume that the total budget available to implement the interventions is known to the analyst. Mathematically, this means that the decision problem contains an added constraint which sets an upper bound to the total costs of the interventions.

The above weaknesses are disadvantageous in practice for two reasons. Firstly, there is undoubtedly uncertainty in the costs and effects of the interventions and, secondly, the total budget that is available to implement the interventions is not known. For example, the terms of reference of a national public health body such as NICE (National Institute for Clinical Excellence) – required to determine the cost-effectiveness of interventions – preclude it from assuming a budgetary constraint explicitly in its deliberations. But if the budgetary constraint is not known, it is impossible to formulate the decision problem as a deterministic mathematical programming problem (Birch and Gafni, 2002).

The main objective of this paper is to propose a stochastic mathematical programming framework for CEA which takes into account both the uncertainties in the costs and the effects of interventions and the uncertainty in the budgetary constraint. The framework combines the Bayesian modelling hierarchical approach to simulate costs and effects of interventions and stochastic mathematical programming methods to determine the optimal intervention.

The paper is divided into nine sections in addition to the Introduction and an appendix. The second section highlights general issues concerning uncertainty in

CEA. The third section reviews the CEAC approach and highlights the weakness of its optimality criterion. The fourth section reviews the deterministic mathematical programming approach for CEA and outlines its disadvantages. The fifth section introduces the stochastic mathematical programming framework. The sixth section gives a motivating example to explain the basis of the stochastic framework. The seventh section outlines the overall stochastic mathematical framework. The eighth section discusses the practical implications of the new framework and the final section concludes. The Appendix summarizes the mathematical notation used in this paper.

## 2. Dealing with uncertainty in CEA

CEA are based on models that map interventions to costs and effects and on decision rules that determine the *optimal* intervention given this mapping. When dealing with uncertainty in CEA, two issues need to be addressed:

- How to represent uncertainty in the models
- How to determine the optimal intervention under model uncertainty

Consider first the representation of uncertainty in CEA models. It seems useful to differentiate between two types of model uncertainty: extrinsic and intrinsic. We define intrinsic uncertainty as the uncertainty associated with elements of a CEA model that are known to the analyst at least to some degree. Typical examples are the costs and effects of interventions. We define extrinsic uncertainty on the other hand as the uncertainty associated with elements of a model that in principle are not known to the analyst. A typical example is the actual budgetary constraint. In the case of intrinsic uncertainty, additional information can be sought to reduce the uncertainty whereas in the case of extrinsic uncertainty the uncertainty cannot be reduced or eliminated. Although this distinction between intrinsic and extrinsic uncertainty in a CEA model may seem artificial and unnerving – as we believe that the Bayesian hierarchical modelling approach is best in capturing both types – it has implications for the formulation of the decision problem, as will be shown later.

The case for adopting a Bayesian decision theoretic approach in CEA has been well made in health economics (Claxton *et al.*, 2000; Claxton *et al.*, 2001). The decision rule advocated by this approach is to select the intervention which maximises the *expected* net health benefit metric, where expectation is carried over the posterior distribution. It has been shown by these authors however that the intervention that maximises the expected net benefit is not necessarily the one that has the highest probability of maximising net benefit. It will be argued below that the tradeoffs between such different concepts of *optimality* can be made more explicit if the decision problem is cast in a strict mathematical programming framework.

## 3. Cost-effectiveness acceptability curves and frontiers

CEAC curves or frontiers are being increasingly adopted by health economists as one of the leading methods for CEA. Their attraction lies in the way they bring model uncertainty to the fore. However the decision rules in the CEAC approach at present are problematic, particularly in the case of the multiple-interventions decision problem.

Consider separately the two-interventions decision problem and the multiple-interventions decision problem.

In the first, the comparison is between an old (or a current) intervention and a new intervention. For this problem, the CEAC curve is defined as (Van Hout *et al.*, 1994; Lothgren and Zethraeus, 2000; O'Hagan *et al.*, 2000; Briggs *et al.*, 2002)

$$z(l) = P\left(l\,\Delta_e - \Delta_c \geq 0\right) \tag{1}$$

where $\Delta_e$ and $\Delta_c$ are respectively the incremental effect and incremental cost of the new intervention (*new*) relative to the old intervention (*old*)

$$\Delta_e = e_{new} - e_{old} \tag{2}$$

$$\Delta_c = c_{new} - c_{old} \tag{3}$$

$l$ is the willingness to pay for an additional unit of effect and $P$ is the probability measure.

In the two-interventions decision problem, the normal decision rule is to specify $l$ and then accept that *new* is cost-effective at a certainty level $z(l)$, alternatively (Briggs *et al.*, 2002) specify the level of certainty $z$ at which *new* is cost-effective and accept the implied level of $l$, where $l$ is read-off the CEAC curve. Note that the intervention selected by the above decision rule is not necessarily the same as that which maximises expected net benefit unless the probability distribution of incremental net benefit is normal (Fenwick *et al.*, 2001).

In the multiple-interventions decision problem, the above definition of a CEAC curve is extended to describe a CEAC frontier (Fenwick *et al.*, 2001; Briggs *et al.*, 2002). Consider in this case $n$ interventions and denote the set of interventions by $I = \{1 \ldots n\}$ where $i \in I$ is the $i^{\text{th}}$ intervention. Let $e_i$ and $c_i$ denote respectively the effect and cost of intervention $i$ and define the net benefit of intervention $i$ by

$$b_i = l\, e_i - c_i \tag{4}$$

The set of CEAC curves for the multiple-interventions decision problem is

$$z_i(l) = P\left(b_i > b_j \mid \forall j \neq i, j \in I, l\right) \tag{5}$$

where $z_i(l)$ is the probability that the net benefit of intervention $i$ is higher than all other interventions. The CEAC frontier ( $f$ ) is defined as

$$f(l) = \max_{i \in I}\left(z_i(l)\right) \tag{6}$$

In other words, the CEAC frontier is the upper bound of the set of CEAC curves.

The normal CEAC decision rule for the multiple-decisions problem is to select the intervention which has the highest probability of yielding the maximum net benefit, i.e. the intervention which solves problem (6). The above rule selects for each $l$ an optimal intervention at a given (not necessarily the same) degree of certainty.

The above rules of CEAC are narrow in scope. The CEAC approach uses one measure to characterise 'maximum net benefit': the highest probability of its occurrence. Note that 'maximum net benefit' is a random variable. Other statistical measures could be used to quantify its uncertainty and high order statistical measures of maximum net benefit could be relevant for equity considerations. Its probability distribution is induced by the probability distributions of the costs and effects of the interventions..

The relationship between the uncertainty of the CE model and the uncertainty of the maximum net benefit is not made explicit in the CEAC approach. This should be made more explicit to increase confidence in the decision technology. It is shown below that using a stochastic mathematical programming framework provides a coherent and transparent approach in modelling and propagating the model uncertainty in decision-making.

## 4. Formulation of the multiple-interventions decision problem as a deterministic mathematical programming problem

Let us now formulate the multiple-interventions decision problem as a deterministic mathematical programming problem, initially ignoring uncertainty.

The reformulated multiple-interventions decision problem takes the form (Stinnett and Paltiel, 1996)

$$\max_{\{x_i, i\in I\}} \left( \left( \sum_{i=1}^{n} x_i e_i \right) l - \sum_{i=1}^{n} x_i c_i \right) \qquad (7)$$

The maximisation is carried over the decision variables $\{x_i, i \in I\}$. There are three cases to consider (Stinnett and Paltiel, 1996):

- $x_i \in [0,1], \forall i \in I$. Here $x_i$ is a Boolean variable (0 or 1) which means that an intervention is either implemented fully or is not implemented at all (indivisible).
- $0 \le x_i \le 1, \forall i \in I$. This means that all interventions are implemented partially (divisible).
- $0 \le x_i \le 1, \forall i \in I_1;\ x_i \in [0,1], \forall i \in I_1; I_1 \cap I_2 = f; I_1 \cup I_2 = I$. $I_1$ and $I_2$ represent respectively the sets of divisible ($I_1$) and indivisible ($I_2$) interventions. $I_1 \cup I_2 = I$ means that an intervention is either divisible or indivisible and $I_1 \cap I_2 = f$ means that an intervention cannot be both divisible and indivisible. This case is a mixture of the first two cases where some interventions are divisible and others are indivisible.

The mathematical programming problem corresponding to the first case is known as an integer programming problem. It is formulated as

$$\max_{\{x_i, i \in I\}} \left( \left( \sum_{i=1}^{n} x_i e_i \right) l - \sum_{i=1}^{n} x_i c_i \right)$$
$$\textit{such that} \qquad (8)$$
$$x_i \in [0,1] \qquad \forall i \in I$$

The mathematical programming problem corresponding to the second case is

$$\max_{\{x_i, i \in I\}} \left( \left( \sum_{i=1}^{n} x_i e_i \right) l - \sum_{i=1}^{n} x_i c_i \right)$$
$$\textit{such that} \qquad (9)$$
$$0 \le x_i \le 1 \quad \forall i \in I$$

Finally, the mathematical programming problem corresponding to the third case is known as a mixed integer programming problem and is described by

$$\max_{\{x_i, i \in I\}} \left( \left( \sum_{i=1}^{n} x_i e_i \right) l - \sum_{i=1}^{n} x_i c_i \right)$$

$$\begin{aligned} &\textit{such that} \\ &0 \le x_i \le 1 \quad \forall i \in I_1 \\ &x_i \in [0,1] \quad \forall i \in I_2 \\ &I_1 \cup I_2 = I \\ &I_1 \cap I_2 = f \end{aligned} \qquad (10)$$

The above mathematical formulations ignore the fact that it is possible that the solution of problem (9) – which is defined for divisible interventions – may yield an $x_i = 0$ or an $x_i = 1$ for some intervention $i$ making this intervention divisible. If this is to be avoided, than a strict inequality ( $0 < x_i < 1$ ) should be used as a constraint. This issue is of technical nature and has relevance only in implementation.

Consider now, without loss of generality, the case of divisible interventions. If the budgetary constraint $k$ is known, then the solution should satisfy the following constraint:

$$\sum_{i=1}^{n} x_i c_i \le k \qquad (11)$$

The complete mathematical programming problem in this case becomes:

$$\max_{\{x_i, i \in I\}} \left( \left( \sum_{i=1}^{n} x_i e_i \right) l - \sum_{i=1}^{n} x_i c_i \right)$$

$$\begin{aligned} &\textit{such that} \\ &0 \le x_i \le 1 \quad \forall i \in I \\ &\sum_{i=1}^{n} c_i x_i \le k \end{aligned} \qquad (12)$$

Contrary to previous assertions (Stinnett and Paltiel, 1996), the solutions of any of the above mathematical programming problems (8, 9, 10 or 12) may include an intervention whose cost-effectiveness ratio is greater than $l$ . Indeed if it is required that all selected interventions have cost-effectiveness ratio less than $l$ , the following set of $n$ inequalities must be included in the mathematical programming problem as constraints:

$$x_i(l\ e_i - c_i) \geq 0 \quad \forall i \in I \tag{13}$$

Because $x_i \geq 0$, inequality (13) is satisfied if $x_i > 0$ when $(l\ e_i - c_i) > 0$ and if $x_i = 0$ when $(l\ e_i - c_i) < 0$. In other words by satisfying this constraint, any selected intervention will have its cost-effectiveness ratio less than $l$ .

There is a vast literature on methods to solve deterministic mathematical programming problems of the types described above. Unfortunately, as has been discussed previously, deterministic formulations have several drawbacks for CEA. They do not take model uncertainty into account, and furthermore, assume that the budgetary constraint is known. The problems should be formulated stochastically to address the uncertainty in the models and in the constraints. The literature on stochastic mathematical programming is not as vast however as the literature on deterministic mathematical programming.

## 5. Formulation of the multiple-interventions decision problem as a stochastic mathematical programming problem

In order to simplify the algebra, we introduce the following notation.

Denote by $\hat{x}$ the intervention vector whose components are $x_1, x_2 ...., x_n$ i.e.

$$\hat{x} = \begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix} \tag{14}$$

Throughout this paper we will use the 'hat' symbol on a variable name to indicate that the variable is a vector. The transpose of vector $\hat{x}$ is denoted by $\hat{x}' = (x_1 \quad \ldots \quad x_n)$. Similarly, introduce the vectors $\hat{e}$ and $\hat{c}$ whose elements are the effects and costs of the interventions respectively.

Consider, again without loss of generality, the case where all interventions are divisible. (Extension to the other two cases is straightforward). The mathematical programming problem which includes a budgetary constraint becomes

$$
\begin{aligned}
&\max_{\hat{x}} \left( l\hat{x}'\hat{e} - \hat{x}'\hat{c} \right) \\
&\textit{such that} \\
&0 \le x_i \le 1 \quad \forall i \in I \\
&x_i (l\, e_i - c_i) \ge 0 \quad \forall i \in I \\
&\hat{c}'\hat{x} \le k
\end{aligned}
\tag{15}
$$

where

$$
\begin{aligned}
\hat{x}'\hat{e} &= \sum_{i=1}^{n} x_i e_i \\
\hat{x}'\hat{c} &= \sum_{i=1}^{n} x_i c_i \\
\hat{c}'\hat{x} &= \sum_{i=1}^{n} c_i x_i
\end{aligned}
\tag{16}
$$

To represent model uncertainty, a stochastic counterpart to the deterministic mathematical programming problem (15) should be derived. In the reformulated problem, $\hat{e}$ and $\hat{c}$ become stochastic vectors and $k$ becomes a stochastic variable.

A stochastic mathematical programming problem requires careful interpretation. (A good introduction to this subject is given by (Ermoliev and Wets, 1988) ). To differentiate between deterministic vectors (or variables) and their stochastic counterparts, the functional dependence of $\hat{e}$, $\hat{c}$ and $k$ on chance is represented by $\hat{e}(w)$, $\hat{c}(w)$ and $k(w)$ where $w \in \Omega$ is the 'elementary chance event' and $\Omega$ is the set of all such events (i.e. $P(\Omega)=1$). In other words, each $w \in \Omega$ induces one realization of $\hat{e}$, $\hat{c}$ and $k$ .

The stochastic counterpart to problem (15) is

$$\begin{aligned}
&\max_{\hat{x}} \left( l\ \hat{x}'\hat{e}(w) - \hat{x}'\hat{c}(w) \right) \\
&\textit{such that} \\
&\hat{c}'(w)\,\hat{x} - k(w) \le 0 \\
&0 \le x_i \le 1 \qquad \forall i \in I \\
&x_i\,( l\ e_i(w) - c_i(w)) \ge 0 \quad \forall i \in I
\end{aligned} \tag{17}$$

The next section introduces an example to motivate the stochastic formulation.

## 6. Motivating example to illustrate the issues of the stochastic mathematical programming framework

This is adopted from a numerical example used to determine the distribution of the optimum of a stochastic linear programming problem (Stancu-Minasian, 1984). The example is modified here to formulate a simple multiple-interventions decision problem to illustrate the principles of stochastic optimization within the context of CEA.

The multiple-interventions decision problem concerns three divisible interventions ($i=1,2,3$). Denote the costs and effects of the interventions by $(c_i, i=1..3)$ and $(e_i, i=1..3)$ respectively. Assume for simplicity that the probability distributions of $c_i$ and $e_i$ are discrete and are described respectively by the matrices:

$$C_i = \begin{pmatrix} a_{i,1} & a_{i,2} & a_{i,3} \\ p_{i,1} & p_{i,2} & p_{i,3} \end{pmatrix} \tag{18}$$

and

$$E_i = \begin{pmatrix} b_{i,1} & b_{i,2} & b_{i,3} \\ q_{i,1} & q_{i,2} & q_{i,3} \end{pmatrix} \tag{19}$$

for $i = 1,2,3$. The top and bottom row of matrix $C_i$ give respectively the possible costs of intervention $i$ and their corresponding probabilities. The same definition applies to $E_i$. In other words,

$$\begin{aligned} P(c_i = a_{i,1}) &= p_{i,1} \\ P(c_i = a_{i,2}) &= p_{i,2} \\ P(c_i = a_{i,3}) &= p_{i,3} \\ P(e_i = b_{i,1}) &= q_{i,1} \\ P(e_i = b_{i,2}) &= q_{i,2} \\ P(e_i = b_{i,3}) &= q_{i,3} \end{aligned} \tag{20}$$

for $i = 1$, 2 and 3 where $p_{i,1} + p_{i,2} + p_{i,3} = 1$ and $q_{i,1} + q_{i,2} + q_{i,3} = 1$. Assume furthermore – also for simplicity reasons – that the total budget is also described by the discrete probability distribution

$$\mathrm{K} = \begin{pmatrix} g_1 & g_2 & g_3 \\ r_1 & r_2 & r_3 \end{pmatrix} \tag{21}$$

where

$$\begin{aligned} P(k = g_1) &= r_1 \\ P(k = g_2) &= r_2 \\ P(k = g_3) &= r_3 \end{aligned} \tag{22}$$

and $r_1 + r_2 + r_3 = 1$.

The stochastic mathematical programming problem is then given by

$$
\max_{x_1,x_2,x_3}\left(x_1\left(l\,e_1(w)-c_1(w)\right)+x_2\left(l\,e_2(w)-c_2(w)\right)+x_3\left(l\,e_3(w)-c_3(w)\right)\right)
$$
*such that*
$$
\begin{aligned}
&0\le x_1\le 1\\
&0\le x_2\le 1\\
&0\le x_3\le 1\\
&x_1c_1(w)+x_2c_2(w)+x_3c_3(w)-k(w)\le 0\\
&x_1\left(l\,e_1(w)-c_1(w)\right)\ge 0\\
&x_2\left(l\,e_2(w)-c_2(w)\right)\ge 0\\
&x_3\left(l\,e_3(w)-c_3(w)\right)\ge 0
\end{aligned}
\tag{23}
$$

In problem (23), $w$ denotes the chance event. The total number of chance events is $7^3$ ( = 343) corresponding to the number of permutations of the possible values of $e_1, e_2, e_3, c_1, c_2, c_3$ and $k$. Each chance event occurs with a certain probability which can be calculated exactly. For example, the probability of occurrence of the chance event $w$ which corresponds to the permutation

$$
w\equiv\left\{c_1=a_{1,1}\quad c_2=a_{2,3}\quad c_3=a_{3,2}\quad e_1=b_{1,3}\quad e_2=b_{2,2}\quad e_3=b_{3,1}\quad k=g_2\right\} \tag{24}
$$

is given by this product

$$
P(w)=p_{1,1}\,p_{2,3}\,p_{3,2}\,q_{1,3}\,q_{2,2}\,q_{3,1}\,r_2 \tag{25}
$$

The mathematical programming problem corresponding to each chance event is a deterministic problem. Denote by $\left(x_1^*(w)\ x_2^*(w)\ x_3^*(w)\right)'$ the optimal intervention vector corresponding to chance event $w$, $z^*(w)$ the corresponding maximum net benefit and $\Delta^*(w)$ its probability of occurrence. We can work out the probability of occurrence of all chance events, the corresponding optimal intervention vectors and the maximum net benefits as shown in the following table:

| *Chance event* | *Costs, effects and budget* | *Optimal intervention vector* | *Maximum net benefit* | *Probability* |
|---|---|---|---|---|
| $w_1$ | $e_1(w_1), e_2(w_1), e_3(w_1),$ $c_1(w_1), c_2(w_1), c_3(w_1),$ $k(w_1)$ | $x_1^*(w_1), x_2^*(w_1), x_3^*(w_1)$ | $z^*(w_1)$ | $\Delta^*(w_1)$ |
| ….. | … | …. | …. | …. |
| …. | …. | $x_1^{**}, x_2^{**}, x_3^{**}$ | $z^{**}$ | $\Delta^{**}$ |
| …. | … | … | … | … |
| $w_{343}$ | $e_1(w_{343}), e_2(w_{343}), e_3(w_{343}),$ $c_1(w_{343}), c_2(w_{343}), c_3(w_{343}),$ $k(w_{343})$ | $x_1^*(w_{343}), x_2^*(w_{343}), x_3^*(w_{343})$ | $z^*(w_{343})$ | $\Delta^*(w_{343})$ |

The CEAC decision rules select the *optimal* intervention vector which gives the highest probability of achieving the maximum net benefit. In the above table, this optimal intervention vector is denoted by $\left(x_1^{**}\ x_2^{**}\ x_3^{**}\right)'$, the corresponding maximum net benefit $z^{**}$ and its probability of occurrence as $\Delta^{**}$. The probability distribution function of the maximum net benefit is approximated from the last two columns in the above table.

Note there are many ways to proceed with problem (23). The easiest to conceptualise is that which involves taking the expectation of the maximand and the constraints. If we denote by $\langle\ \rangle$ the expectation operator, problem (23) is reduced to

$$\max_{x_1,x_2,x_3} \left\langle x_1\left(l\,e_1(w)-c_1(w)\right)+x_2\left(l\,e_2(w)-c_2(w)\right)+x_3\left(l\,e_3(w)-c_3(w)\right)\right\rangle$$
$$\begin{aligned}
&\textit{such that}\\
&0 \le x_1 \le 1\\
&0 \le x_2 \le 1\\
&0 \le x_3 \le 1\\
&\left\langle x_1 c_1(w)+x_2 c_2(w)+x_3 c_3(w)-k(w)\right\rangle \le 0\\
&\left\langle x_1\left(l\,e_1(w)-c_1(w)\right)\right\rangle \ge 0\\
&\left\langle x_2\left(l\,e_2(w)-c_2(w)\right)\right\rangle \ge 0\\
&\left\langle x_3\left(l\,e_3(w)-c_3(w)\right)\right\rangle \ge 0
\end{aligned} \tag{26}$$

which in turn is reduced, by propagating the expectation operator through each term, to:

$$\max_{x_1,x_2,x_3} \left(x_1\left(l\langle e_1\rangle-\langle c_1\rangle\right)+x_2\left(l\langle e_2\rangle-\langle c_2\rangle\right)+x_3\left(l\langle e_3\rangle-\langle c_3\rangle\right)\right)$$
$$\begin{aligned}
&\textit{such that}\\
&0 \le x_1 \le 1\\
&0 \le x_2 \le 1\\
&0 \le x_3 \le 1\\
&x_1\langle c_1\rangle+x_2\langle c_2\rangle+x_3\langle c_3\rangle-\langle k\rangle \le 0\\
&x_1\left(l\langle e_1\rangle-\langle c_1\rangle\right)\ge 0\\
&x_2\left(l\langle e_2\rangle-\langle c_2\rangle\right)\ge 0\\
&x_3\left(l\langle e_3\rangle-\langle c_3\rangle\right)\ge 0
\end{aligned} \tag{27}$$

Equation (27) is a deterministic mathematical programming. The expected values of $e_1,e_2,e_3,c_1,c_2,c_3,k$ are substituted in Eqn. (27) using the relationships:

$$\begin{aligned}
\langle c_i\rangle &= a_{i,1}\,p_{i,1}+a_{i,2}\,p_{i,2}+a_{i,3}\,p_{i,3}\\
\langle e_i\rangle &= b_{i,1}\,q_{i,1}+b_{i,2}\,q_{i,2}+b_{i,3}\,r_{i,3}\\
\langle k\rangle &= g_{i,1}\,r_{i,1}+g_{i,2}\,r_{i,2}+g_{i,3}\,r_{i,3}
\end{aligned} \tag{28}$$

for $i$=1, 2 and 3. Equation (28) is obtained from Eqns. (20) and (22).

If we donate the optimal intervention vector for problem (27) by $(\bar{x}_1\ \bar{x}_2\ \bar{x}_3)'$ and the corresponding maximum net benefit by $\bar{z}$, it is obvious that in general

- $(\bar{x}_1\ \bar{x}_2\ \bar{x}_3)' \neq (x_1^{**}\ x_2^{**}\ x_3^{**})'$ and $\hat{z} \neq z^{**}$
- $(\bar{x}_1\ \bar{x}_2\ \bar{x}_3)' \neq (\langle x_1^{*}\rangle \langle x_2^{*}\rangle \langle x_3^{*}\rangle)'$ and $\bar{z} \neq \langle z^{*}\rangle$

The first set of inequalities means that the solution of problem (27) is not the same as that obtained using CEAC decision rules. The second set of inequalities on the other hand means that the solution of problem (27) is not the same as the expectation of the solutions of problem (17) . One can think of many other alternatives to problem (23) Note therefore that – for transparency considerations – it is essential that the optimisation problem is formulated explicitly in order to appreciate the 'terms of reference' of the solution. The next section proposes a stochastic mathematical programming framework for CEA which we believe holds promise in terms of transparency and propagating and handling uncertainty.

## 7. The stochastic mathematical programming framework

For ease of representing the new framework, reformulate problem (17) as

$$\begin{aligned}
&\max_{\hat{x}} \left(\hat{h}'(w)\hat{x}\right) \\
&\textit{such that} \\
&\hat{g}'(w)\hat{x} - b(w) \leq 0 \\
&\hat{l}_b \leq \hat{x} \leq \hat{u}_b \\
&L(\hat{x})\hat{h}(w) \geq \hat{l}_b
\end{aligned} \qquad (29)$$

where

$$\hat{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \hat{l}_b = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \hat{u}_b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \hat{h}(w) = \begin{pmatrix} l\, e_1(w) - c_1(w) \\ l\, e_2(w) - c_2(w) \\ l\, e_3(w) - c_3(w) \end{pmatrix},$$

$$\hat{g}(w) = \begin{pmatrix} c_1(w) \\ c_2(w) \\ c_3(w) \end{pmatrix} \quad , \quad L(\hat{x}) = \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{pmatrix}$$

The last constraint in problem (29) ensures that only interventions with cost-effectiveness ratios less than $l$ are selected

$$L(\hat{x})\hat{h} = \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{pmatrix} \begin{pmatrix} l\, e_1 - c_1 \\ l\, e_2 - c_2 \\ l\, e_3 - c_3 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

For each chance event $w$, there exists an optimal set of interventions which solves problem (29). Denote $\hat{x}^*(w)$ the vector of optimal interventions and by $z^*(w)$ the corresponding maximum net benefit. $z^*(w)$ is a random variable which depends on $\hat{e}(w)$, $\hat{c}(w)$ and $k(w)$. We can estimate the probability density function of $z^*(w)$, $\Phi(z^*)$, from knowledge of the probability density functions of $\hat{e}(w)$, $\hat{c}(w)$, $k(w)$ as was demonstrated in the motivating example above. Of course, in the general case, approximations methods – different from the simple one used in the motivating example – are needed to approximate $\Phi(z^*)$. These include the discretization method and the incomplete description method (Stancu-Minasian, 1984). Knowledge of $\Phi(z^*)$ enables us to compute the high order statistics of the maximum net health benefit that could have relevance where equity is a consideration.

Now that we have established the association between model uncertainty and the uncertainty in the optimal net benefit, we can formulate the decision problem from a Bayesian perspective. Bayesian models are often used to describe $\hat{e}(w)$, $\hat{c}(w)$, $k(w)$. These models are represented hierarchically, with the prior values of $\hat{e}(w)$, $\hat{c}(w)$ at the lowest end of the hierarchy.

Our aim is to maximise the posterior expectation of total net benefit subject to constraints. Recall first that uncertainty in $\hat{e}(w)$ and $\hat{c}(w)$ (intrinsic uncertainty) is conceptually different from that of $k(w)$ (extrinsic uncertainty). Because the budgetary constraint is in principle unknown, the shape of the probability distribution of $k(w)$ can be modelled arbitrary but is assumed to have finite support

$$P(k_{\min} \leq k(w) \leq k_{\max}) = 1 \tag{30}$$

where $k_{\min}$ and $k_{\max}$ are constants suitably defined.

Because of the nature of the uncertainty in $k$, it is proposed to replace the constraint

$$\hat{c}'(w)\hat{x} - k(w) \leq 0 \tag{31}$$

with the relaxed and less restrictive constraint

$$P\left(\hat{c}'(w)\hat{x} - k(w) \leq 0\right) \geq d \tag{32}$$

where $0 \leq d \leq 1$. In other words, the constraint is to be satisfied in probability only.

Adopting a Bayesian decision theoretic perspective leads to solving the following stochastic programming problem

$$\begin{aligned}
&\max_{\hat{x} \in \hat{X}(w)} \left\langle l\ \hat{x}'\hat{e}(w) - \hat{x}'\hat{c}(w) \right\rangle \\
&\textit{such that} \\
&P\left(\hat{c}'(w)\,\hat{x} - k(w) \leq 0\right) \geq d \\
&\hat{l}_b \leq \hat{x} \leq \hat{u}_b \\
&L(\hat{x})\left\langle \hat{h} \right\rangle \geq \hat{l}_b
\end{aligned} \tag{33}$$

Equation (33) is known mathematically as a ‘chance constraint’ problem. A chance constraint problem can be turned into a deterministic mathematical programming problem by transforming the stochastic constraint into an equivalent deterministic constraint (Vajda, 1972; Kibzun and Kan, 1996). For simplicity of illustration, we shall ignore the uncertainty in $\hat{c}'$ and just concern ourselves with the uncertainty in $k$, where we assume that the probability distribution function of $k$ is given by $\Psi$.

Note that $\Psi(.)$ is a monotonically increasing function between 0 and 1 and therefore admits a unique inverse $\tilde{\Psi}(.)$ such that $\tilde{\Psi}(\Psi(.)) = 1$. It can be shown (Vajda, 1972) that the probabilistic constraint

$$P(\hat{c}'\hat{x} - k(w) \leq 0) \geq d \tag{34}$$

is equivalent to the deterministic constraint

$$\hat{c}'\hat{x} \leq \tilde{\Psi}(1-d) \tag{35}$$

It is straightforward to extend the above arguments to determine the equivalent deterministic constraint to the probabilistic constraint

$$P(\hat{c}'(w)\hat{x} - k(w) \leq 0) \geq d \tag{36}$$

using quantile rules (Vajda, 1972; Kibzun and Kan, 1996). Inequality (36) takes into account the uncertainties in the costs and effects of the interventions.

It is beyond the scope of this paper to discuss the numerical methods available to solve stochastic optimisation problems of the type described in problem (33) (Ermoliev and Wets, 1988; Mayer, 1998). These methods will need to be integrated with Markov Chain Monte Carlo methods which are used for stochastic simulation (Gilks *et al.*, 1996).

**8. Practical implications**

Although the above stochastic framework may initially seem to be too theoretical to be of any practical use, this is not the case. Indeed it could be argued that the complexity of the framework is due to our requirement that the decision-making process is made transparent. By formulating the decision problem explicitly as a mathematical programming problem, the 'terms of reference' of the solution are made very clear. The practical implications of the above framework for CEA are challenging because the analyst is forced to formulate the intervention-decision problems rigorously and explicitly

In practice, there are several steps towards formulating a decision problem in the above framework.

- The first step is to define the maximand. The maximand is the total net benefit which includes cost and effect terms associated with each intervention. The decision rules of this framework maximise the expected value of the maximand whilst satisfying some constraints.

- The second step is to define the set of constraints. The constraints confine the 'search region' where the solution is to be found. In addition to the roles described above, the constraints can include inequalities associated with equity. The set of constraints should not be contradictory and ideally should define a wide feasible region. There is a balance to achieve between constraint satisfaction and optimisation. The decision rules ensure that constraints are satisfied either 'on average' or with given probabilities.

- The third step models the uncertainties in the cost and effects of interventions and in the total budget. The costs and effects of the interventions and the budget are modelled appropriately by Bayesian hierarchical models. The decision rules compute the posterior expectation of the vectors and variables.

- The fourth step approximates the probability distribution of the optimal maximand. This information is used to expose the effect of the model uncertainties on the distribution of the optimal maximand. Although in theory this information does not influence the solution, it can be used a priori - if required - to define appropriate risk terms for inclusion in the maximand.

If we were to follow the above steps, then we can be confident that the solution that we obtain satisfies strict mathematical criteria in terms of optimality and feasibility. Of course, this is achieved at the expense of further complexity.

## 9. Conclusion

Much work has been done recently on developing stochastic CEA methods for the evaluation of healthcare interventions. The main aim of most CEA methods is to provide a decision tool to aid the decision-making process. CEA methods are based on models that simulate the costs and effects of interventions and on decision rules to determine the optimal intervention or set of interventions. However, one of the difficulties encountered by decision-analysts is the interpretation of CEA results when these are embedded in uncertainty. It is argued in this paper that by presenting the decision problem as a stochastic mathematical programming problem, a more coherent and transparent decision technology is formulated.

## Appendix: Mathematical notation

Scalars are denoted by small case letters ( $x$ ) and vectors by a 'hat' symbol at the top of the variable name ( $\hat{x}$ ). The transpose of a vector $\hat{x}$ is denoted by $\hat{x}'$, i.e. if $\hat{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, then $\hat{x}' = (x_1 \quad x_2 \quad x_3)$. The dot product of two vectors $\hat{x}$ and $\hat{y}$ of dimension $n$ is $\hat{x}'\hat{y} = \sum_{i=1}^{n} x_i y_i$ . Matrices are denoted by upper case letter, e.g. $X$ , except for the variable name $P$ (see below).

The expected value of $x$ (or vector $\hat{x}$ ) is denoted by $\langle x \rangle$ (or vector $\langle \hat{x} \rangle$ ). If $x$ denotes either a decision variable or the maximand in a mathematical programming problem, then $x^*$ denotes the optimal value of $x$ . $P(.)$ denotes the probability measure, $w$ is the elementary chance event and $\Omega$ is the space of all chance events. The symbol $\in$ means 'an element of', e.g. $w \in \Omega$. Stochastic variables or stochastic vectors are differentiated from their deterministic counterparts by giving their functional dependence on $w$, e.g. $x(w)$ or $\hat{x}(w)$.

The standard mathematical programming problem is represented by

$$\max_{\hat{x}}(h)$$
*such that*
$$\hat{u}(\hat{x}) \leq 0$$

where $\max_{\hat{x}}(.)$ is the maximisation operator, $h$ is the maximand and $\hat{u}$ is the vector of equalities or inequalities. The above problem is interpreted as follows: Determine the optimal vector $\hat{x}$ which maximises $h$ whilst satisfying $\hat{u}(\hat{x}) \leq 0$.

If $f(x)$ is a function of $x$, $\tilde{f}(x)$ denotes the inverse function of $f(x)$, i.e. $\tilde{f}(f(x)) = f\left(\tilde{f}(x)\right) = 1$.

**References**


Birch S and Gafni A (2002) On being NICE in the UK: guidelines for technology appraisal for the NHS in England and Wales. *Health Economics* 11: 185-191.

Briggs A H (1999) A Bayesian approach to stochastic cost-effectiveness analysis. *Health Economics* 8: 257-261.

Briggs A H (2000) Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 17: 479-500.

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

Briggs A H and Gray A M (1999) Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technology Assessment* 3(2).

Briggs A H, O'Brien B J and Blackhouse G (2002) Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. *Annuals Reviews in Public Health* 23: 377-401.

Claxton K, Lacey L F and Walker S G (2000) Selecting treatments: a decision theoretic approach. *Journal of the Royal Statistical Society A* 163: 211-225.

Claxton K, Neumann P J, Araki S and Weinstein M C (2001) Bayesian value-of-information analysis. An application to a policy model of Alzheimer's disease. *International Journal of Technology Assessment in Health Care* 17: 38-55.

Ermoliev Y and Wets R, J-B (1988). Numerical techniques for stochastic optimization. Berlin, Springer-Verlag.

Ermoliev Y and Wets R, J-B (1988). Stochastic programming, an introduction. Numerical techniques for stochastic optimization. Y. Ermoliev and R. Wets, J-B. Berlin, Springer-Verlag: 1-32.

Fenwick E, Claxton K and Sculpher M (2001) Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics Letters* 10: 779-787.

Gilks W R, Richardson S and Spiegelhalter D J (1996). Markov chain Monte Carlo in practice. London, Chapman & Hall.

Hutubessy R C W, Baltussen R M P M, Evans D B, Barendregt J J and Murray C J L (2001) Stochastic league tables: communicating cost-effectiveness results to decision-makers. *Health Economics Letters* 10: 473-477.

Kibzun A I and Kan Y S (1996). Stochastic prgramming problems with probability and quantile functions. Chichester, John Wiley & Sons.

Lothgren M and Zethraeus N (2000) Definition, interpretation and calculation of cost-effectivness acceptability curves. *Health Economics* 9: 623-630.

Mayer J (1998). Stochastic linear programming: a comparison based on a model management system. Amsterdam, Gordon and Breach Science Publishers.

O'Hagan A, Stevens J W and Montmartin J (2000) Inference for the cost-effectiveness acceptability curve and cost-effectiveness ratio. *Pharmacoeconomics* 17: 339-349.

O'Hagan A, Stevens J W and Montmartin J (2001) Bayesian cost-effectiveness analysis from clinical trial data. *Statistics in Medicine* 20: 733-753.

Sendi P, Gafni A and Birch S (2002) Opportunity costs and uncertainty in the economic evaluation of health care interventions. *Health Economics* 11: 23-31.

Spiegelhalter D J and Best N G (2002) Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statistics in Medicine (in press)*.

Spiegelhalter D J, Myles J P, Jones D R and Abrams K R (2000) Bayesian methods in health technology assessment: a review. *Health Technology Assessment* 4(38).

Stancu-Minasian I M (1984). Stochastic programming with multiple objective functions. Dordrecht, D.Reidel Publishing Company.

Stinnett A A and Mullahy J (1998) Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical Decision Making* 18: S68-S80.

Stinnett A A and Paltiel A D (1996) Mathematical programming for the efficient allocation of health care resources. *Journal of Health Economics* 15: 641-653.

Vajda S (1972). Probabilistic programming. New York, Academic Press.

Van Hout B A, AL M J, Gordon G S and Rutten F F H (1994) Costs, effects and C/R-ratios alongside a clinical trial. *Health Economics* 3: 309-319.