THE UNIVERSITY of York

CENTRE FOR HEALTH ECONOMICS

Selecting Treatments: A Decision Theoretic Approach

Karl Claxton
Larry F. Lacey
Stephen G. Walker

CHE Technical Paper Series 15

Selecting Treatments: A Decision Theoretic Approach

Karl Claxton[1] Larry F. Lacey[2]
Stephen G. Walker[3]

Abstract

This paper looks at the problem of comparing two treatments, for a particular patient population, where one is the current standard treatment and the other a possible alternative under investigation. With limited (finite) financial resources the decision whether to replace one by the other is not going to be based on health benefits alone. This motivates an economic evaluation of the two competing treatments where the cost of any gain in health benefit is scrutinised; it is whether this cost is acceptable to the relevant authorities which decides if the new treatment can become the standard. We adopt a Bayesian decision theoretic framework in which a utility function is introduced describing the consequences of making a particular decision when the true state of nature is expressed via an unknown parameter θ (this parameter denotes cost, effectiveness, etc.) The treatment providing the maximium posterior expected utility summarises the decision rule, expectations taken over the posterior distribution of the parameter θ.

1. Introduction. A common decision problem, which arises in a medical context, consists of comparing two treatments T_0 and T_1; T_0 is the current treatment and T_1 a new treatment under investigation as a possible alternative to T_0. The question — whether to replace T_0 by T_1 or stay with T_0. This is a decision problem and as much as possible we should try and stay within a decision theoretic framework. We confine our attention to the Bayesian framework and an excellent exposition describing our position is to be found in the paper by Lindley (1961) and also in the book by De Groot (1970) among others.

There are a number of possible perspectives which can be taken when considering this decision problem. These include the individual patient, the individual clinician, the service provider (an NHS Trust and its formulary committee), the purchasing authority (Health authority or GP fundholder), the health care sector (NHS), a societal decision maker considering the impact of the decision on society as a whole as well as the perspective of the pharmaceutical industry. Here we focus

on two broad perspectives: that of a health care sector decision maker (either in purchasing or providing care) and that of society as a whole (see Gold et al., 1996b)

The idea is straightforward: we have two treatments and we are uncertain about performance (equivalently, the consequences of using a treatment). We let the parameter θ_j denote the performance or consequence of using T_j. The consequences of importance are discussed in Section 3. Information is collected about θ_j, via experimentation on a sample from the population, and the posterior distribution for θ_j constructed. If θ_j is known then we assume the consequences of the decision are fully known and can be valued via a utility function $U(j, \theta_j)$. The decision rule is to select the treatment which has the *maximum* utility. Instead, if we have a posterior distribution for θ_j to represent beliefs after seeing the data (instead of knowing its true value) then the decision rule is to select the treatment which has the maximum *expected* utility; expectation taken over the posterior distribution. The following quote is from Lindley (1961):

> The best decision to take is that which maximises the expected utility, the expectation being taken over θ_j.

It is becoming increasingly recognised that an economic evaluation alongside a clinical trial, assessing health benefits and costs simultaneously, is both feasible and necessary (Drummond and Stoddart, 1984; Drummond and Davies, 1991; O'Brien et al., 1994; Van Hoult, 1994; Drummond, 1995). It is not difficult to motivate such an economic evaluation of health care. A new treatment will have a 'cost' associated with it. The new treatment would obviously be chosen if it demonstrated improved benefits at reduced costs and would obviously not be chosen if costs increased but benefits decreased. These are two extreme cases pointing to yes/no regions in the cost/benefit space. The more difficult decisions arise when the new treatment has improved benefit at increased cost or reduced benefit at reduced cost — does the increase in cost justify the benefit? This should not depend on the chronological order in which the treatments are investigated. Would the current treatment have replaced the new treatment if the latter had been discovered first? The historical accident dictating which of the alternatives is regarded as current practice should be irrelevant.

Without knowlegde of the costs associated with treatments it is not possible to discuss these fundamental questions which arise naturally in any decision making process. Spiegelhalter and Smith (1981) discuss a utility based approach to decision making in clinical decisions but confine their attention to quantifying utilities for health benefit alone. They state "In general, of course, the problem of trading off benefits versus costs of health states is a fundamental one and cannot be ignored". However, their approach is based on relative utilities and so it is not obvious how costs could be incorporated, using their ideas.

There is now a substantial amount of literature on the evaluation of the cost/benefit of treatments with a view to selecting the optimal treatment: essentially, the optimal allocation of limited resources to maximising a particular output (health benefit). This is mostly, if not entirely, to be found in the *health economic* literature.

On the other hand, the problem can genuinely be regarded as a statistical decision problem and the theory is well documented in the statistics literature. Spiegelhalter et al. (1994) acknowledge that a utility/decision theoretic based approach is "correct" but think it unfeasible. In a discussion to the paper of Spiegelhalter et al., D. V. Lindley hopes it will not take 25 years for a *statistician* to tackle (and solve) this problem, apparently unaware that health economists have been working on it for many years.

Let us therefore accept that an economic evaluation alongside health benefits is useful (if not necessary) in the decision making process. The question that arises is how to use the information to decide on which treatment is 'better'. The most common economic evaluation of health care has been a *cost-effectiveness* (CE) analysis which measures health benefit in either natural physical units or health related quality of life (see, for example, Drummond et al., 1987; Weinstein, 1995; Gold et al., 1996b). Cost-effectiveness analysis is concerned with the effectiveness of treatment in routine clinical practice (a pragmatic attitude) rather than efficacy in a controlled and atypical setting (an explanatory attitude) (see Schwartz and Lellouch, 1967). This has apparently been more popular to the medical community than the *cost-benefit* (CB) analysis which explicitly requires the conversion of health outcome into monetary units. We discuss in Section 2.

2. Economic evaluation of health care. Here we briefly discuss the two types of economic evaluations described in the literature (see Weinstein and Stason, 1977; Warner and Luce 1982; Drummond et al., 1987; Gold et al., 1996b).

Cost effectiveness. The CE approach is currently the most popular method for comparing the usefulness of two competing treatments for limited resources. A CE analysis compares incremental cost and effectiveness of a new treatment with the current treatment. Specifically, the ratio of interest is given by

$$R = \frac{C_1 - C_0}{H_1 - H_0},$$

where H_j is the health benefit of the jth treatment, typically measured in physical units such as life years (LY) or measures of health related quality of life (HRQL) such as quality adjusted life years (QALY) or heathy years equivalent (HYE), among others. C_j is the associated cost of treatment j. So R quantifies the cost of a gain in one unit of health benefit. If both $\Delta C = C_1 - C_0$ and $\Delta H = H_1 - H_0$ are positive then the decision rule is to select the new treatment T_1 to replace T_0 if $R < \lambda$, where the threshold value λ is the maximum amount the decision maker is willing to pay for 1 unit of effectiveness. On the other hand, if both are negative, the decision rule is to select the new treatment T_1 to replace T_0 if $R > \lambda$. If one is negative and the other positive the decision rule is obvious.

Cost benefit. In a CB analysis, all quantities of interest are expressed in the same, usually monetary, units. Since health benefit is the only quantity not already in these units, this type of analysis depends on these benefits being converted into

money by eliciting individual marginal willingness to pay, typically using contingent valuation methods. (see Johansson, 1995 and references there in). Once this has been done it is possible to obtain utilities $\{U_j\}$ by adding and subtracting, and we are in familiar territory. The treatment with the largest utility is selected.

There has been a lot of discussion on the merits and drawbacks to the use of these two approaches (Donaldson, 1990; Johannesson and Jonsson, 1991; Pauly, 1995; Culyer and Evans, 1996). Our discussion is focused on the goal of the analysis: to make a decision. In a CE analysis we only care how R compares with the choice of λ and in a CB analyis if $U_1 > U_0$. However, if ΔC and ΔH are both positive and $R < \lambda$ then

$$U_1 = \lambda H_1 - C_1 > \lambda H_0 - C_0 = U_0$$

which is the CB decision rule when $\lambda (= 1/g)$ is equal to the individual marginal willingness to pay per effectiveness unit. In fact, Phelps and Mushlin (1991) argue that in these circumstances as far as decision making is concerned a CB and CE analysis are mathematically equivalent. Both require a price per effectiveness unit to be determined. Phelps and Mushlin (1991) say:

> Each CE ratio implies a dollar value per QALY and therefore provides the necessary information to conduct a straightforward CB analysis. Resource allocation using CE methods requires or produces a value of g that would in turn allow direct use of CB methods. Thus CE and CB methods are indistinguishable when used for selecting among competing treatments in a resource-scarce environment.

This reformulation of cost-effectiveness analysis has been called net health benefits analysis (NHB) (see Claxton and Posnett 1996; Stinnett and Mullahy, 1998; and Claxton, 1999). It differs from the traditional cost benefit formulation (where benefits are valued according to individual willingness to pay) because the same monetary valuation is applied to health benefits across all individuals. For further discussion of this distinction, the circumstances when CE/NHB will lead to the same decisions as CB and some of the equity implications of each approach, see Garber et al. (1996); Weinstein and Manning (1997); and Garber and Phelps (1997).

Based on the material in this section, the appropriate utility function to consider is

$$U_j = g^{-1} H_j - C_j.$$

Since this utility function depends on cost and effectiveness we obviously take $\theta_j = (H_j, C_j)$ and construct the posterior distribution for θ_j. We collect information about (H_j, C_j), derive the posterior distribution for $[H_j, C_j|\text{data}]$, from which we obtain

$$\bar{U}_j = g^{-1} \bar{H}_j - \bar{C}_j,$$

where $\bar{H}_j, \bar{C}_j$ are the posterior means. We choose the appropriate treatment depending on the sign of $\Delta = \bar{U}_1 - \bar{U}_0$. Information about H and C is gathered via

the economic evaluation conducted alongside the clinical trial. Assigning a value to g is more problematic. We discuss in Section 3.

Before proceeding to Section 3, we note that we can use this decision theoretic approach to determine a starting dose. Essentially we replace the j indexing competing treatments with a j indexing doses for a particular drug. The theory is the same.

3. Quantifying the unknowns. In this section we consider the components of C and H and the problem of detemining the value of g. There is substantial literature quantifying health benefits, particularly in the medical and medical statistics literature, and so we do not discuss it here. In the context of this discussion the quantification of health benefits can be done using some single dimensional physical measures of effectiveness such as life years or some generic measure of health related quality life (see Torrance 1986; The EuroQol Group 1990; Mehrez and Gafni, 1993; Kaplan, 1995; and Gold et al., 1996a)

For a recent review of the use of QALYs and more generally Quality-of-Life (QOL) measures, see Cox et al. (1992) and references cited in this paper. Trevor Sheldon, in the discussion of the paper by Cox et al., says:

> For individuals, collapsing QOL dimensions into a single score is reasonable in that they can be involved in a practical decision analytic exercise. The weighting will reflect individual preference sets in the context of real treatment opinions. The results are not generalised to others. However, economists, via the QALY, are developing a standardised, non-disease-specific instrument — a generic single-score outcome measure which, when combined with cost data, can be used to rank treatments for rationing.

First, we discuss which costs should be included in the utility function and second, we discuss how to quantify g.

Quantifying costs. Quantifying costs associated with health care is not so well documented and, not suprisingly, is concentrated in the health economic literature. The fundamental question is exactly which costs should be included: a recent survey and recommendations were given by the US Panel on cost-effectiveness analysis (see Luce et al., 1996; Gold et al., 1996b). Aside from the direct cost of providing the treatment; that is, the cost of the medication, administration, nursing etc., what else should be included?

Which cost should be included depends on the perspective of the evaluation. There are a number of possible perspectives which can be taken but here we focus on two broad perspectives: that of a health care sector decision maker (either in purchasing or providing care) and that of a society as a whole. The key distinction is that in the former the budget for health care can be taken to be exogenously determined where as in the later it is endogenous. We consider both because the perspective of health care sector decision maker is most common in economic evaluations of health care technologies and the societal perspective (although less com-

mon) is the most appropriate when considering public policy decisions (see Gold et al., 1996b) such as setting priorities in research and development and the efficient regulation of new pharmaceuticals.

From the perspective of a health care sector decision maker the budget for service provision is exogenous, so the costs should only include those which draw on the specific budget at the time the decision is made to proceed with the treatment. As far as the health care sector decision maker is concerned it is only the cost (from the decision makers budget) of implementing a treatment which will influence the decision. A discussion of methods for estimating these direct costs can be found in Drummond (1987); Dranove (1995); and Luce et al. (1996).

Some authors argue that future related medical costs, arising as a consequence of extending life, should also be included (see Weinstein and Manning, 1997; and Garber and Phelps, 1997). Weinstein and Stason (1977) and Willems et al. (1980) both discuss the inclusion of costs that patients would expect to incur as a result of their lives being extended. As Weinstein (1995) points out "the implicit budget in such analyses is the total expenditure on health care." Russell (1986) argues that:

> if the purpose of the analysis is ... to determine whether the program is a good investment, only the costs of the preventive program should be counted. Added years of life involve added expenditures for food, clothes, and housing as well as medical care. None ... is relevant to deciding whether the program is a good investment ...

This suggests that costs arising after the completion of an epsiode of treatment must be dealt with separately since the decision to treat or not to treat in the future is still open. Costs should only incorporate those relevant to the present decision, and not future costs for which no decision has yet been made to incur.

However if we take the perspective of society as a whole then not only do future related medical costs become relevant to a societal decision maker but also all future unrelated medical and non medical costs (for the current debate see Meltzer, 1997; Garber and Phelps 1997; and Weinstien and Manning, 1997). Also, indirect costs and benefits such as productivity losses and gains in the labour market, which are irrelevant to the health care sector budget, are relevant to a societal decision maker. Weinstein (1995) states:

> from a societal perspective, indirect costs and benefits, that is, the value of time spent or saved, can represent real resources forgone or recovered. These costs and benefits are not relevant if the constrained resource is health care cost and are therefore excluded if the health sector ... perspective is used.

Clearly the range of costs that are considered relevant is determined by the perspective adopted and whether the budget for health care is regarded as exogenous or endogenous. Some of these issues, particularly including future costs, remain controversial but will be resolved as the methodological and applied literature develops.

Quantifying g. There are a number of approaches to establishing a value of g that have been considered. We will concentrate on two ideas which we believe to be consistent with the perspective of a health care sector decision maker facing an exogenous budget constraint and a societal decision maker where the budget for health care is endogenous.

An exogenous budget constraint. The health care sector will be supporting a number of programs and investing in a treatment for each program. If a rational approach has been taken then the treatments and programs will maximise health benefits within the budget constraint. Of all these treatments there will be the *marginal* program, the program having the least cost-effective treatment currently supported. If $\Delta\bar{H}, \Delta\bar{C} > 0$, and the new treatment is going to be financed, then it is from the marginal program that the necessary resources will be taken (assuming a fixed budget). These resources will be found by selecting a less expensive alternative treatment within the marginal program. If we denote the incremental cost-effective ratio for these two treatments in the marginal program by λ_M then it is obvious that we would only proceed with the new treatment if $\Delta\bar{C}/\Delta\bar{H} < \lambda_M$. Although $\lambda \geq \lambda_M$ we only need to know λ_M to know whether we choose T_0 or T_1. Even if we knew λ and $\Delta\bar{C}/\Delta\bar{H} < \lambda$ but $\Delta\bar{C}/\Delta\bar{H} > \lambda_M$ then the new treatment would not be adopted because there would not be the motivation to finance the extra cost, as no resources would be avaliable, unless the decision maker behaved irrationally. Essentially, λ_M plays the role of λ. This is the argument of Phelps and Mushlin (1991), where $1/\lambda_M$ can be regarded as the shadow price of the budget constraint (see Stinnett and Paltiel, 1996).

According to Weinstein (1995), from the perspective of a health care sector decision maker, this is "the most direct and theortically correct criterion for judging the acceptability of a $\Delta C/\Delta H$ ratio ... the $\Delta C/\Delta H$ ratio of the least cost-effective funded program provides the standard against which competing uses of resources must be measured." Note, that as the budget is increased, the incremental CE ratio of the marginal project will increase as more effective but more costly programmes are adopted.

Basically, the idea is simple: is the incremental cost-effectiveness ratio of the new treatment more acceptable than that of the marginal program? This provides the threshold value of interest to the decision maker. Clearly, without full and perfect information about existing programs the value of λ will be uncertain and will change over time (the budget and the performance of funded programs may change). However what is required for decision making is the expected value of λ at the time the decision must be made. In this sense it seems reasonable to regard it as a constant at the point that the decision must be made.

An endogenous budget constraint. The above ideas assume the new treatment is going to be funded from the decision makers exisiting constrained budget. It may be that this is not the case and funds will be available from other sources, such as other public services. From a societal perspective, there is no reason to regard

existing budgets for health care as fixed. The question now is how much is society willing to give up for an additional year of life or the societal willingness to pay? In these circumstances the value of g is the normative choice of a social decision maker rather then a positive empirical question. The budget is now endogenous because all programs which have positive NHB will be funded and the budget will be increased until the NHB of the marginal project is equal to zero. Although an analyst may not know with certainty which value of g will be chosen it will not be uncertain to the societal decision maker. We suggest that analysis is conducted conditional on a range of values of g. It is then the task of societal decision makers to make a normative choice of which value of g is acceptable. In practice generally agreed fixed prices are often used ($50,000 per life year in the US).

Whatever view is taken concerning the appropriate characterisation of g, it is clear that a price per effectiveness unit must be assigned implicitly or explicitly (Weinstein and Zeckhauser, 1972; Phelps and Mushlin, 1988; Stinnett and Paltiel, 1996). That is, even if no g is designated and a decision maker thinks they have avoided the problem of assigning such a value, one must be able to investigate the decision and imply a value of g. Surely, it is more desirable to acknowledge the existence of g and tackle the problem 'head-on' rather than try and side-step the issue.

4. Relation to previous work. The influential paper of Spiegelhalter et al. (1994), from now on SFP, is our main source of reference and starting point for this section. The main aim of the paper by SFP was to promote the Bayesian approach to randomised clinical trials, as an alternative to the classical significance test of the null hypothesis and p-values.

Let H_0 be a parameter associated with treatment T_0 and H_1 a parameter associated with treatment T_1, where $H_1 - H_0$ measures a difference of interest between the two treatments. This difference will most likely be quantified in terms of health benefit. The classical null hypothesis is $H_0 : \delta = 0$, where $\delta = H_1 - H_0 = \Delta H$. The approach of SFP is to construct a posterior distribution for δ, and to use this to investigate the usefulness of the treatments, by comparing the posterior distribution with a *range of equivalence* (Freedman, 1984). This range of equivalence consists of a specified range (δ_I, δ_S). If $\delta < \delta_I$ then T_1 is regarded as inferior to T_0, whereas, if $\delta > \delta_S$, then T_1 is regarded as superior to T_0. If $\delta_I < \delta < \delta_S$ then the treatments are regarded as equivalent, and according to SFP "we would be unable to make a definitive choice of treatment", presumably remaining with T_0.

Some guidelines for the construction of this range are provided by Freedman et al. (1984), Armitage (1989) and Fleming and Watelet (1989). It is clear, as SFP point out

> Treatments may be so unequal in their costs, in toxicity, inconvenience or monetary costs, that it is commonly accepted that the more costly treatment will be required to achieve at least a certain margin of benefit δ_I, before it can even be considered.

Obviously, since δ is unknown, the posterior for δ is used to ascertain the relevant probabilities of δ being in a particular region relative to the range of equivalence. According to SFP, the monitoring of the trial involves making decisions based on the posterior probabilities $P(\delta < \delta_I)$, $P(\delta_I < \delta < \delta_S)$ and $P(\delta > \delta_S)$. If $P(\delta > \delta_S|\text{data}) > 1 - \epsilon$ then the decision is made to go with the new treatment T_1 and if $P(\delta < \delta_I|\text{data}) > 1 - \epsilon$ then the decision is made to stay with the old treatment T_0. Otherwise, the decision is to continue with randomisation. But what should ϵ be? SFP provide the following quote:

> We are reluctant to express a firm opinion about what this critical size ϵ should be, since the choice should, in principle anyway, be made from decision theoretic considerations of expected utility. But we have already said that this is unrealisitic, and so we are left with the conventional bench-marks such as 2.5% and 5%.

This does seem somewhat unsatisfactory.

SFP think the 'correct' procedure of using decision theoretic considerations of maximising expected utility as unrealistic:

> Although there have been many attempts to place clinical trials within such a decision theoretic framework, in our formulation we specifically do not include utility assessments. Our reason is that when a decision is whether or not to discontinue the trial, coupled with whether or not to recommend one treatment in preference to the other, the consequences of any particular course of action are so uncertain that they make the meaningful specification of utilities rather speculative.

Berry (1994) disagrees:

> I disagree that a decision theoretic approach using utilities is unrealistic in clinical trials ...speculation and assessing uncertainty are the stuff of the Bayesian approach. In deciding whether to stop a trial the authors spurn utilities and are left with conventional benchmarks such as 2.5% and 5%. In my view, deciding when to stop a trial requires considering why we are running it in the first place, and this means assessing utilities.

Lindley (1994) adds:

> It must be recognised that clinical trials are not there for inference but to reach a decision ... expected utility is realistic and, indeed, necessary. It is only by using expected utility that we can be sure that our actions will fit together sensibly.

In fact, the two decisions referred to by SFP can be dealt with separately. If we agree to continue with randomisation then we do not need to suggest a treatment preference — we continue with randomisation. It is only when we decide not to

continue with the trial that we need to indicate treatment preference, which we have already discussed.

The problem of establishing a range of equivalence can be compared to the problem of finding g. Incidentally, Freedman (1984) uses a procedure of 'iterative questioning' of clinicians to establish the range of equivalence. However, in spite of a range of equivalence, the SFP decision rule is to choose T_1 if $\Delta H > \delta_S$. Our decision rule is to choose T_1 if $\Delta H > g\Delta C$. Regardless of how we implement these decision rules in the light of uncertainty, the connection between δ_S and g is apparent:

$$g = \delta_S / \Delta C.$$

If SFP provide δ_S then evaluation of ΔC provides a value for g.

Our evaluation of health benefit will allow comparison across diseases. The problem of using a measure of benefit which is not comparable across diseases is that even if the new treatment has improved benefit, but is more expensive, where are the necessary finances going to come from to fund the extra cost? If the budget is fixed, money must come from an alternative program and necessarily one must then make comparisons across diseases; and have a measure of benefit which allows this.

5. Continuing with the trial. Knowing when to stop a trial is crucial. According to our approach, once the decision has been taken to stop the trial, then the choice of preferred treatment is the one with maximum expected utility at this point of termination. Before looking at the problem, we consider why the trial is being conducted in the first place.

A clinical trial, combined with an economic evaluation of health care, is carried out because it is thought that the costs of the trial will not exceed the benefits (in monetary terms) of discovering the new treatment is more cost effective than the old. That is, the benefits gained by changing are greater than the loss incurred by running the trial. Let us formalise this with the use of utilities.

The benefit in not conducting a trial and staying with the old treatment is simply NU_0, where N is the population size, that can benefit from information generated by a trial. N can be calculated by considering the incident rate of the disease, the effective lifetime of the treatment(s) and the discount rate applied to health benefits. The benefit of running a trial on a sample of size n is given, initially, by $n(U_0+U_1)/2$, assuming 50% randomisation (for a discussion of optimal allocation in a fixed sample design see Claxton, 1999a). The cost of the trial is denoted by C_n. After the trial is completed, the remaining $N-n$ members of the population will provide benefit U_0 each, if $\Delta(n) < 0$, and provide U_1 each, if $\Delta(n) > 0$. Here $\Delta(n) = n^{-1}(\Delta_1 + \cdots + \Delta_n)$, where $\Delta_1, \ldots, \Delta_n$ are iid copies of $\Delta = U_1 - U_0$. Therefore, the gain in conducting a trial of size n is given by

$$\mathcal{U}_n = n(U_0 + U_1)/2 - C_n + (N-n)\left[U_0 I(\Delta(n) < 0) + U_1 I(\Delta(n) > 0)\right] - NU_0.$$

If we take $\Delta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ a priori, where $\mathcal{N}$ denotes a Normal density, then, noting

that $U_0 + \Delta = U_1$, we have

$$E(\mathcal{U}_n) = n\mu_0/2 - C_n + (N-n)E\left[\Delta I(\Delta(n) > 0)\right].$$

If we further assume

$$\Delta_i \sim \mathcal{N}(\Delta, \phi^2),$$

then

$$\Delta|\Delta(n) \sim \mathcal{N}(\mu_n, \sigma_n^2),$$

where

$$\mu_n = \frac{\mu_0\phi^2 + n\Delta(n)\sigma_0^2}{\phi^2 + n\sigma_0^2}$$

and

$$\sigma_n^2 = \frac{\phi^2\sigma_0^2}{\phi^2 + n\sigma_0^2}.$$

Then $E\left[\Delta I(\Delta(n) > 0)\right]$ is given by

$$\frac{\mu_0\phi^2 + n\sigma_0^2 L(\mu_0, \sigma_0)}{\phi^2 + n\sigma_0^2},$$

where

$$L(\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x>0} x \exp\left\{-0.5(x-\mu)^2/\sigma^2\right\}\, dx.$$

According to prior beliefs, it is coherent to conduct a trial of size n if $E(\mathcal{U}_n) > 0$. Obviously, as information is collected, beliefs will change and so it is necessary to monitor results, checking whether the updated beliefs are compatible with continuing the trial.

We can work this out in the most general terms. Suppose that m observations have been made; that is, we have seen $\Delta_1, \ldots, \Delta_m$. The expected gain in continuing the trial for a further n observations is given by

$$E(\mathcal{U}_{nm}) = n\mu_m/2 - C_n + (N-m-n)\left\{\frac{\mu_m\phi^2 + n\sigma_m^2 L(\mu_m, \sigma_m)}{\phi^2 + n\sigma_m^2}\right\}.$$

This will provide the decision whether to continue the trial or not. See also Claxton and Posnett (1996) and Claxton (1999a).

Clearly different perspectives embody different objectives and constraints (utility functions) which in practice will lead to different decision about adoption of a technology and conducting a trial. From a societal perspective this represents a failure of markets, incentives and regulation to achieve socially desirable outcome. One reason for this apparent failure is that research in general and the information generated by a clinical trial in particular is non-rival (once it is produced it can be used by all at no additional cost) and has public good characteristics (see Culyer, 1999). This means that lower level decision makers will underestimate the value of

research (conducting and continuing a trial) to society as a whole and the amount of research conducted will be less than socially optimal. In these circumstances it is appropriate to take a societal perspective when considering public policy issues such as the regulation of new pharmaceuticals and setting priorities in publicly funded research.

For example, a pharmaceutical company funds a trial to establish a new treatment on the market and benefits from the sale of the treamtent to the health care sector. The funding of a trial is a 'gamble' taken by the pharmaceutical company. It pays out the cost of the trial with the 'chance' of reaping the sales if the treatment gets on the market. This depends on the criterion by which a new treatment can get on the market. This is the 'chance' referred to: the result is unkown prior to the decision to conduct the trial. If *equivalence* with the current treatment is all that is required, the risk of failure might be small and the gamble seen as a good one.

However, once on the market, the new treatment still needs to be purchased in sufficient quantities. For this, the new treatment might have to be shown to be 'better' than the current treatment (see Backhouse, 1998). 'Better' could mean more cost-effective, according to the utility function of those purchasing the new drug. The pharmaceutical company might have been happy about demonstrating equivalence (and paying for this), but not for demonstrating improved cost effectiveness, which might be seen as too risky a gamble to make. This is where the health care sector or societal decision maker can intervene in two ways: firstly, they could make a coherent decision to fund a trial, after a pharmaceutical company has got the new treatment on the market. This is precisely the role of some aspects of the NHS Health Technology Assessment Program and some MRC funded trials. Therefore, the pharmaceutical company funds an equivalence trial, with a view to getting the treatment accepted, and, on the success of this, the health care sector funds a cost-effectiveness study. If not, they simply remain with the current treatment.

Alternatively, a societal decision maker could consider adopting regulations which would require evidence of cost-effectiveness, rather then simply equivalence, before the approval of a new drug (the Australian guidelines are an example of such an approach). However all current regulatory regimes can be described as arbitrary for three reasons. Firstly the standard of demonstrating efficacy or even cost-effectiveness at arbitrarily selected levels of significance and power relies on a traditional power calculation which excludes the marginal cost of sampling and implicitly places an infinite value on the benefits of sample information (see Claxton and Posnett, 1996). This leads to either unbounded or arbitrary sample sizes, consequently regulation based on this calculation demands either infinite or arbitrary amounts of information (see Claxton, 1999b). Secondly, the type of information demanded by regulatory regimens is not directly relevant to the decisions of "a formulary committee or similar entity ... [or in] the selection of drugs for managed care or other similar organisation" (Section 114, Food and Drug Administration Modernization and Accountability Act of 1997). Useful information for decision

making should take the form pr[hypothesis|data] but the information provided by a traditional frequentist clinical trial takes the form pr[data|hypothesis] which by itself provides little useful information to decision makers. Finally any regulatory regimen that demands the same standards of evidence in all circumstance and across all technologies irrespective of any evidence already available, the size of the patient population that could benefit from the new technology, and the costs of gathering more information simply cannot be efficient. These issues seem to be recognised in the more recent US Food and Drug Administration legislation (see FDA Modernization and Accountability Act 1997; and FDA Prescription Drug User Fee Act 1997), and a definition of competent and reliable evidence which references the Federal Trade Commissions standards (1984):

> ... a reasonable basis [for a claim of cost-effectiveness] depends ... on a number of factors relevant to the benefits and costs of substantiating a particular claim. These factors include: the type of product, the consequences of a false claim, the benefits of a truthful claim, the costs of developing substantiation for the claim ...

This standard of evidence requires explicit consideration of the marginal benefits and costs of acquiring additional information but no method for estimating these costs and benefits has been suggested. The approach outlined in this paper does provide such a method and a framework which can define a claim as "substantiated" and evidence as "competent and reliable" when it is not efficient to gather any more information. Efficient regulation would demand more information for some new technologies as compared to others and require different amounts of information for the same technology in different circumstances (for some stylised examples see Claxton, 1999b). The appropriate role of regulatory authorities should be to police the prior information which is explicitly used in this type analysis and implicity used in the classical approach (see Berger and Berry, 1988). The consequences of the existing arbitrary regulation will be distorted research and development priorities, inappropriate approval decisions and less than optimal flow of phamacoeconomic information. All of which will have tangible costs in terms of health gains forgone. These issues seem to be recognised by regulatory agencies and Bayesian decision theory provides the practical tools to implement a more rational approach.

The approach outlined in this paper can provide an estimate of the expected net benefits of proposed research before sample information is collected. This can be used to prioritise proposed research (which have a high expected net benefit) within a clinical area and allocate research and development resources between broad areas of clinical research, as well as considering the allocation of resources between health care provision and research and development (see Claxton, 1996). Currently complete applications of this approach are limited but it has been applied to a policy model of Alzheimers disease (Claxton et al., 1998). The purpose of this analysis was to inform the decision as to whether an additional clinical trial would be worth while, if so whether a longer follow-up (the existing trial was 24 weeks)

would be required and also which endpoints (costs and HRQL) should be included in the design. It was found the value of perfect information (the maximum value of a trial) for the decision to adopt this new technology (Donepezil) was in the region of $160 million for the US population and that information about efficacy duration (a longer follow-up) was valued at $130 million. The value of additional information about indirect and direct costs, and health related quality of life measures was substantially lower (less than $10 million). This prior analysis does start to inform important policy issues and subsequent work will estimate the value of sample information using different trial designs. Fenwick et al. (1999) also presented an analysis of diagnostic and treatment strategies for urinary tract infection. She considered the value of additional information associated with each parameter in a decision analytic model and considered the value of information surrounding each of the many possible strategies of diagnosis and treatment. This work continues and it is hoped that it will inform (at least in retrospect) the decision of the NHS Health Technology Assessment Program to call for a clinical trial in this area in 1998. The practical issues which arise when implementing this approach are not insurmountable and if in the process of application the analysis does become "rather speculative" it will simply reflect the fact that (currently) there may be limited information about important parameters that cannot be ignored when decisions must be made. The solution to this problem is to acquire more information about these parameters (where that is efficient) rather than adopt methods which exclude these important uncertainties.

6. Discussion. In this paper we have argued that explicit monetary valuation of health care is an essential and unavoidable procedure in the cost-effective comparison of a new treatment with current practice. The shadow price of the budget constraint determines which of the alternatives will be regarded as cost-effective and establishes the key parameter g. The value of this crucial parameter should not be abdicated to the vagaries of uninformed, ad hoc or implicit judgements. We have suggested other approaches make such an evaluation implicitly. Reasons why decision theoretic approaches have been rejected in the past have been primarily due to the problem of the explicit valuation of g. However, if the budget for service prevision is fixed then the shadow price of the budget constraint is also fixed allowing health outcome to be rescaled into monetary units. If a societal perspective is taken then the budget for health care is endogenous and λ represents societies marginal willingness to pay for improvements in health outcome. Which ever view is taken it is clear that the value of λ can and should be made explicit.

The measure of NHB presented are based on a particular objective or social welfare function which may be judged inappropriate and consequently the measures of NHB may be regarded as incomplete. However, if there are equity issues which need to be incorporated they can be made explicit with appropriate adjustments to the measure of outcome (Deber and Goel, 1990). Similarly, on the issue of the safety of a new technology, if there is particular concern for rare but catastrophic events, then these undesirable outcomes should be given appropriate weight in

the calculation of expected net benefits. If decision makers wish to adopt some type of voting rule then the decision may focus on the median rather than the mean net benefits and the treatment which benefits the greatest number could be adopted Attitudes to risk can be incorporated in the measures of outcome but if we wish to incorporate the fact that some individuals preferences violate the axioms of expected utility theory then prospect theory or some notion of regret (Loomes and Sugden, 1982) can be used in the measure of health benefits. It is not necessary to take a position on the appropriate social welfare function; different definitions of need; the normative content of expected utility theory; or the best way to incorporate important and legitimate equity concerns (although it is worth noting that they will all have implications for the measures of outcome and the design of studies not just their interpretation) only to note that these issues do not change the fundamental point that inference is irrelevant and decisions need to be made.

The framework for decision making outlined in this paper can distinguish the simultaneous but conceptually separate steps of deciding which alternatives should be chosen, given existing information, from the question of whether more information should be acquired. It mirrors the sequential nature of decision making: making an initial decision; deciding to gather evidence; revising decisions in the light of this new information; and again considering whether more information is required. It also ensures that the type of information acquired (research) is driven by the objectives of the health care system. This approach to the value of information means that clinical research can be designed and research priorities can be set in a way which is consistent with the objectives of health service provision and the resource constraint faced by clinicians and society as a whole.

Acknowledgements. The authors would like to thank two referees for comments and questions relating to a previous draft of the paper.

References

Backhouse M.E. (1998). An investment appraisal approach to clinical trial design. *Health Economics* **7**, 605-619

Berger, J.O. and Berry, D.A. (1998). Statistical analysis and the illusion of objectivity. *American Scientist* **76**, 159-165.

Berry, D.A. (1994). Discussion of the paper by Spiegelhalter et al. (1994). *Journal of the Royal Statistical Society, Series A* **157**, 357-416.

Boyle, M.H., Torrance, G.W., Sinclair, J.C. and Horwood, S.P. (1983). Economic evaluation of neonatal intensive care of very low birth weight infants. *New England Journal of Medicine* **308**, 1330-1337.

Chancellor, J.V., Hill, A.M., Sabin, C.A., Simpson, K.N. and Youle, M. (1997). Modelling the cost effectiveness of Lamivudine/Zidovudine combination therapy in HIV infection. *PharmacoEconomics* **12**, 1-13.

Claxton, K and Posnett, J. (1996). An economic approach to clinical trial design and research priority-setting. *Health Economics* **5**, 513-524.

Claxton, K., Neuman, P.J., Araki, S.S. and Weinstein, M.C. (1998). The value of information: application to a policy model of Alzheimers disease. Paper presented to the Society for Medical Decision Making, Boston (MA), October, 1998. Abstract published in Medical Decision Making, 1998 (18) 4.

Claxton, K. (1999a). The irrelevance of inference: a decision making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics* (in press).

Claxton K. (1999b). Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. *Health Economics* (in press).

Cox, D.R., Fitzpatrick, R., Fletcher, A.E., Gore, S.M., Spiegelhalter, D.J. and Jones, D.R. (1992). Quality-of-life assessment: can we keep it simple? *Journal of the Royal Statistical Society, Series A* **155**, 353-393.

Coyle, D. (1996). Statistical Analysis in Pharmacoeconomic Studies. *PharmacoEconomics* **9**, 506-516.

Culyer, A.J. (1999). Economics and public policy: NHS research and development as a public good. Centre for Health Economics Discussion Paper No. 163.

Culyer, A.J. and Evans, R.G. (1996). Mark Pauly on welfare economics: Normative rabbits from positive hats. *Journal of Health Economics* **15**, 243-251.

Deber, R. and Goel, V. (1990). Using explicit decision rules to manage issues of justice, risk and ethics in decision analysis: when is it not rational to maximise expected utility? *Medical Decision Making* **10**, 181-194.

De Groot, M. (1970). *Optimal Statistical Decisions.* Addison-Wesley, New York.

Donaldson, C. (1990). Willingness to pay for publicly provided goods: a possible measure of benefit? *Journal of Health Economics* **9**, 103-118.

Dranove, D. (1995). Measuring costs. In *Valuing Health Care.* F.A. Sloan (Ed). Chapter 4.

Drummond, M.F., Stoddart, G.L. and Torrance, G.W. (1987). *Methods for the economic evaluation of health care programmes.* Oxford University Press.

Drummond M. (1995). Economic analysis alongside clinical trials: problems and potential. *Journal of Rheumatology* **22**, 1403-1407.

Drummond, M.F. and Davies, L. (1991). Economic analysis alongside clinical trials: revisiting the methodological issues. *International Journal of Technology Assessment in Health Care* **7**, 561-573.

Drummond, M.F. and Stoddart, G.L. (1984). Economic analysis and clinical trials. *Controlled Clinical Trials* **5**, 115-128.

Fenwick, E., Sculpher, M., Claxton, K. and Briggs, A. (1999). Improving the efficiency and relevance of health technology assessment: the role of decision analytic modelling. Health Economics Study Group, Birmingham, January, 1999.

Fleming, T.R. and Watelet, L.F. (1989). Approaches to monitoring clinical trials. *J. Natn. Cancer Inst.* **81**, 188-193.

Freedman, L.S., Lowe, D. and Macaskill, P. (1984). Stopping rules for clinical trials incorporating clinical opinion. *Biometrics* **40**, 575-586.

Garber A.M., Weinstein, M.C., Torrance, G.W. and Kamlet M.S. (1996). Theoretical foundations of cost-effectiveness analysis. In *Cost-effectiv-eness in health and medicine.* Oxford University Press.

Garber, A.M. and Phelps, C.E. (1997). Economic foundations of cost effectiveness analysis. *Journal of Health Economics* **16**, 1-31.

Gold, M.R., Siegel, J.E., Russell, L.B. and Weinstein, M.C. (1996a). Identifying and valuing outcomes. In *Cost-effectiveness in Health and Medicine.* Oxford University Press.

Gold, M.R., Siegel, J.E., Russell, L.B. and Weinstein, M.C. (1996b). *Cost-effectiveness in Health and Medicine.* Oxford University Press.

Johansson, P.O. (1995). *Evaluating Health Risks.* Cambridge University Press.

Johannesson, M. and Jonsson, B. (1991). Economic evaluation in health: is there a role for cost-benefit analysis? *Health Policy*, 1-19.

Kaplen, R.M. (1995). Utility assessment for estimating quality-adjusted life years. In *Valuing Health Care.* F.A. Sloan (Ed). Chapter 3.

Lindley, D.V. (1961). Dynamic programming and decision theory. *Applied Statistics* **10**, 39-51.

Lindley, D.V. (1994). Discussion of the paper by Spiegelhalter et al. (1994). *Journal of the Royal Statistical Society, Series A* **157**, 357-416.

Loomes, G. and Sugden, R. (1982). Regret theory; an alternative theory of rational choice under uncertainty. *Economics Journal* **92**, 805-824.

Luce, B.R., Manning, J.E. Siegel J.E. and Lipscomb, J. (1996). Estimating costs in cost-effectiveness analysis. In *Cost-effectiveness in Health and Medicine.* Gold et al. (Eds). Oxford University Press.

Mehrez, A. and Gafni, A. (1993) Healthy years equivalents versus quality adjusted life years: in pursuit of progress. *Medical Decision Making* **13**, 287-292.

Meltzer, D. (1997). Accounting for future costs in medical cost-effectiveness analysis. *Journal of Health Economics* **16**, 33-64.

Mullahy, J. (1997). Which cost-effectiveness ratio? Evaluation of health policies and medical technologies in a stochastic context. Preprint.

O'Brien, B.J., Drummond, M.F. and Labelle, R.J. (1994). In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. *Medical Care* **32**, 150-63.

Olsen, J. (1993). On what basis should health be discounted? *Journal of Health Economics* **12**, 39-54.

Pauly, M.V. (1995). Valuing health care benefits in money terms. In *Valuing Health Care.* F.A. Sloan (Ed). Chapter 6.

Phelps, C.E. and Mushlin, A.I. (1988). Focusing technology assessment using medical decision theory. *Medical Decision Making* **8**, 279-289.

Phelps, C.E. and Mushlin, A.I. (1991). On the (near) equivalence of cost-effectiveness and cost-benefit analyses. *International Journal of Technology Assessment in Health Care* **7**, 12-21.

Russell, L.B. (1986). Is prevention better than cure? Brookings Institute, Washington, DC.

Schwartz, D. and Lellouch, J. (1967). Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Chronic Disease*, 637-648.

F.A. Sloan. (1995). *Valuing Health Care.* Cambridge University Press.

Spiegelhalter, D.J. and Smith, A.F.M. (1981). Decision analysis and clinical decisions. In *Perspectives in Medical Statistics* (J.F.Bithell and R.Coppi, eds.) 103-131. Academic Press, London.

Spiegelhalter, D.J., Freedman, L.S. and Parmar, M.K.B. (1994). Bayesian approaches to randomized trials. *Journal of the Royal Statistical Society, Series A* **157**, 357-416.

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.

Stinnett, 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.

Torrance, G.W. (1986). Measurement of health state utilities for economic appraisal. *J. Health Economics* **5**, 1-30.

Torrance, G.W. (1987). Utility approach to measuring health-related quality of life. *J. Chronic Diseases* **40**, 593-600.

Van Hoult, B.A. (1994). Costs, effects and CE ratios alongside a clinical trial. *Health Economics* **3**, 309-19.

Viscusi, W.K. (1995). Discounting the health effects for medical decisions. In *Valuing Health Care.* F.A. Sloan (Ed). Chapter 7.

Warner, K. and Luce, B. (1982). *Cost benefit and cost effectiveness analysis in health care: Principles, practice and potential.* Ann Arbor, MI: Health Administration Press.

Weinstein, M.C. and Zeckauser, R.J. (1972). Critical ratios and efficient allocation. *Journal of Public Economics* **2**, 716-721.

Weinstein, M.C. and Stason, W.B. (1977). Foundations of cos-effectiveness analysis for health and medical practices. *New England Journal of Medicine* **296**, 716-721.

Weinstein, M.C. (1995) From cost-effectiveness ratios to resource allocation: where to draw the line? In *Valuing health care: cost, benefits and effectiveness of pharmaceuticals and other medical technologies.* Cambridge University Press.

Weinstein, M.C. and Manning, W.G. (1997). Theoretical issues in cost-effectiveness analysis. *Journal of Health Economics* **16**, 121-128.