Chapter 2
Bayesian Decision Theory and the Design and Analysis of Randomized Clinical Trials

Andrew R. Willan

Abstract Traditional approaches to the analyses and sample size determinations for randomized clinical trials are based on tests of hypotheses and rely on arbitrarily set error probabilities and smallest clinically important differences. Recently Bayesian methods have been proposed as an alternative. In particular, many authors have argued that Bayesian decision theory and associated value of information methods can be used to the determine if current evidence in support of a new health care intervention is sufficient for adoption and, if not, the optimal sample size for a future trial. Value of information methods incorporate current knowledge, the value of health outcome, the incidence and accrual rates, time horizon and trial costs, while maximizing the expected net benefit of future patients and providing an operational definition of equipoise. In this chapter value of information methods are developed in detail and illustrated using a recent example from the literature.

2.1 Introduction

The standard approach to the analysis and sample size determination for a randomized clinical trial (RCT) is based on the use of tests of hypotheses and the frequentists definition of probability. Consider a randomized clinical trial in which a new health care intervention, referred to as *Treatment* and labeled T, is compared to an existing intervention, referred to as *Standard* and labeled as S. The trial is conducted for the purpose of considering the adoption of *Treatment* if it is superior to *Standard*. This type of trial is often referred to as a superiority trial. Let Y be the random variable representing the primary outcome where larger values of Y are preferred, such as survival (where $Y = 1$ if the patient survives, 0 otherwise), survival time, quality-adjusted survival time or net benefit. Let $E(Y|i), i = T, S$ be the expected value of the outcome for a patient randomized to i, and let $\theta = E(Y|T) - E(Y|S)$. Thus, larger values of θ favour *Treatment*. Typically, in a superiority trial the data is used to test the null hypothesis $H : \theta \leq 0$

A.R. Willan (✉)
SickKids Research Institute, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
e-mail: andy@andywillan.com

K. van Montfort et al. (eds.), *Developments in Statistical Evaluation of Clinical Trials*, DOI 10.1007/978-3-642-55345-5_2

versus the alternative hypothesis $A : \theta > 0$. *Treatment* is considered for adoption if, and only if, H is rejected in favour of A. The probability of falsely rejecting H, referred to as the Type I error probability, is set to some relatively small value. Sample size is determined by specifying the smallest clinically important (positive) difference for θ, labelled as θ_{SCID}, and requiring that the probability of failing to reject H when $\theta \geq \theta_{\text{SCID}}$ is less that some relatively small value, referred to as the Type II error probability.

There are many problems with this approach. Firstly, the value selected for the Type I error probability is somewhat arbitrary and is almost always set to 0.05. Using the same value for the probability of a Type I error for every trial ignores the seriousness of the error, which clearly varies from trial to trial. Thus, a trial that randomizes patients with age-related macular degeneration between two different wavelengths of laser coagulation [42] uses the same probability of falsely declaring *Treatment* superior, as does a trial of Caesarean section versus vaginal delivery for women presenting in the breech position [23]. Declaring one wavelength superior to another when they are the same is not a serious error since selecting the wavelength is a matter of simply dialing the appropriate frequency and the only difference to patients is the colour of the light observed during the procedure. However, in the latter, declaring Caesarean section superior when it is the same as vaginal delivery is a serious error. Assigning the same probability to the two errors makes no sense, quite apart from the fact that the value of 0.05 is somewhat arbitrary in the first place. Also somewhat arbitrary is the typical choice of 0.2 for the probability of a Type II error. It means that there is a 20 % chance that the effort and money invested in the trial will be wasted, even if a clinically important difference between the treatments exists. Again, it fails to reflect the seriousness of making the error. The choice of θ_{SCID} can be less arbitrary and can be estimated by polling clinicians and decision makers. However, in practice it is often back-solved from the sample size equation after substituting in a sample size that reflects constraints relating to patient recruitment and budget. Even if θ_{SCID} is a reasonable, clinically determined estimate of the smallest clinically important difference, there is a range of values for the true treatment difference that is less than the smallest clinically important difference, for which the probability of rejecting the null hypothesis and adopting *Treatment* is greater than 50 %. This sometimes referred to as a Type III error.

In response to these problems, many authors have proposed alternative methods [1, 3, 9, 11–16, 18–22, 25, 26, 28, 33–35, 43–49, 52]. In particular many authors have proposed the application of decision theory and associated expected value of information methods for assessing the evidence from RCTs and for determining optimal sample size for future trials. The application of decision theory to the design and sample size determination is the subject of the remainder of this chapter. In Sect. 2.2 an introduction to the cost-effectiveness analysis of RCTs is given, complete with an illustrative example. The use of decision theory in the design and analysis of RCTs is given in Sect. 2.3 and illustrated with the same example in Sect. 2.4. A summary and discussion are given in Sect. 2.5.

2.2 Cost-Effectiveness Analysis of Randomized Clinical Trials

Consider the cost-effectiveness comparison of a new health care intervention referred to as *Treatment* and labeled T, with an existing health care intervention referred to as *Standard* and labeled S. The health care interventions could be therapeutic, preventive or diagnostic. Let e_j and c_j be the respective mean measure of effectiveness and cost for patients receiving intervention j, where $j = T, S$. The measure of effectiveness is framed in the positive, such as surviving the duration of interest, survival time or quality-adjusted survival time. Cost includes not just cost of the interventions, but all down-stream health care cost over the duration of interest and might, depending on the perspective taken, include non-health care cost, such as time lost from work, etc. Let $\Delta_e = e_T - e_S$ and $\Delta_c = c_T - c_S$.

Initially, cost-effectiveness inference was centred on the parameter $R \equiv \Delta_c/\Delta_e$, which is referred to as the incremental cost-effectiveness ratio (ICER) and is the cost of achieving each additional unit of effectiveness from using *Treatment* rather than *Standard*. For example, suppose the probability of surviving the duration of interest was 0.6 for a patient receiving *Standard* and 0.7 for a patient receiving *Treatment* and the respective mean costs for *Standard* and *Treatment* over the duration of interest were \$14,000 and \$15,000 respectively. The ICER $= (15,000 - 14,000)/(0.7 - 0.6) = \$10,000$ per life saved or death averted. Many authors have discussed inference on the cost-effectiveness ratio [4–6, 8, 27, 30–32, 37, 40, 41, 50].

Due to the concerns regarding ratio statistics, focus has shifted from the incremental cost-effectiveness ratio to the incremental net benefit (INB). Let the net benefit (NB) of intervention j be defined as $\text{NB}_j \equiv e_j\lambda - c_j$ where λ is the threshold value for a unit of effectiveness (e.g., the value of saving a life or the value of a year of life gained). The INB is defined as $b \equiv \text{NB}_T - \text{NB}_S = e_T\lambda - c_T - (e_S\lambda - c_S) = \Delta_e\lambda - \Delta_c$. The term $\Delta_e\lambda$ is the incremental effectiveness (benefits) expressed in monetary terms and the term $-\Delta_c$ subtracts the incremental costs, leaving the incremental *net* benefit. When INB is positive, *Treatment* is considered value-for-money and should be considered for adoption, subject to budgetary constraints and the level of uncertainty. In the simple example above $b \equiv 0.1\lambda - 1{,}000$ and is positive for values of the threshold greater than \$10,000 (i.e., the ICER). Many authors have discussed inference on the incremental net benefit [2, 7, 24, 29, 36, 38, 39, 53–56].

Suppose $\hat{\Delta}_e$ and $\hat{\Delta}_c$ are the respective estimates of Δ_e and Δ_c from a study, such as a clinical trial or an observational study, where individual patient measures of effectiveness and cost have been recorded. Let $V(\hat{\Delta}_e)$, $V(\hat{\Delta}_c)$ and $C(\hat{\Delta}_e, \hat{\Delta}_c)$ be the relevant variances and covariance. For more on parameter estimation the reader is referred to Willan and Briggs [57]. Assuming no prior information, and invoking the central limit theorem, the posterior *pdf* for the incremental net benefit can be given by $N(b_0, v_0)$, where $b_0 = \hat{\Delta}_e\lambda - \hat{\Delta}_c$ and $v_0 = V(\hat{\Delta}_e)\lambda^2 + V(\hat{\Delta}_c) - 2C(\hat{\Delta}_e, \hat{\Delta}_c)\lambda$. Inference regarding INB, which is an attempt to characterize the cost-effectiveness of *Treatment* compared to *Standard* and the corresponding uncertainty, can best

be presented by a plot the cost-effective acceptability curve (CEAC), which is the probability that the INB is positive (i.e., that *Treatment* is cost-effective) as a function of the threshold value for a unit of effectiveness and, invoking the central limit theorem, can be calculated as $\Phi\left(b_0/\sqrt{v_0}\right)$, where $\Phi(\cdot)$ is the *cdf* for the standard normal random variable. The CEAC passes through 0.5 at $\lambda = \text{ICER}$, crosses the vertical axis at the probability that *Treatment* is cost saving (i.e., $\hat{\Delta}_c < 0$), and is asymptotic to the right to the probability that *Treatment* is more effective (i.e., $\hat{\Delta}_e > 0$). For more on the CEAC the reader is referred to Fenwick et al. [17].

2.2.1 The CADET-Hp Trial

The CADET-Hp Trial is a double-blind, placebo-controlled, parallel-group, multicentre, randomized controlled trial performed in 36 family practitioner centres across Canada. The results are published in Chiba et al. [10] and Willan [51]. Patients 18 years and over with uninvestigated dyspepsia of at least moderate severity presenting to their family physicians were eligible for randomization, provided they did not have any alarm symptoms and were eligible for empiric drug therapy. Patients were randomized between T: Omeprazole 20 mg, metronidazole 500 mg and clarithromycin 250 mg, and S: Omeprazole 20 mg, placebo metronidazole and placebo clarithromycin.

A total of 288 patients were randomized, 142 ($= n_T$) to *Treatment* and 146 ($= n_S$) to *Standard*. Both regimens were given twice daily for 7 days. The binary measure of effectiveness was treatment success and defined as the presence of no or minimal dyspepsia symptoms at 1 year. Total cost was determined from the societal perspective and are given in Canadian dollars. A summary of the parameter estimates is given in Table 2.1. The details regarding parameter estimation are given in the Appendix. Assuming no prior information and invoking the central limit theorem, the posterior *pdf* for the INB is normal with mean $0.1371\lambda + 53.01$

Table 2.1 Parameter estimates for the CADET-Hp trial

Treatment	Standard		
($n_T = 142$)	($n_S = 146$)		
Proportion of successes	0.5070	0.3699	Difference $= \hat{\Delta}_e = 0.1371$
Average cost	476.97	529.98	Difference $= \hat{\Delta}_c = -53.01$
V(Proportion of successes)	0.00176	0.001596	Sum $= \hat{V}\left(\hat{\Delta}_e\right) = 0.003356$
V(Average cost)	2,167	2,625	Sum $= \hat{V}\left(\hat{\Delta}_c\right) = 4,792$
C(Proportion of successes, mean cost)	−0.2963	−0.4166	sum $= \hat{C}\left(\hat{\Delta}_e, \hat{\Delta}_c\right) = -0.7129$

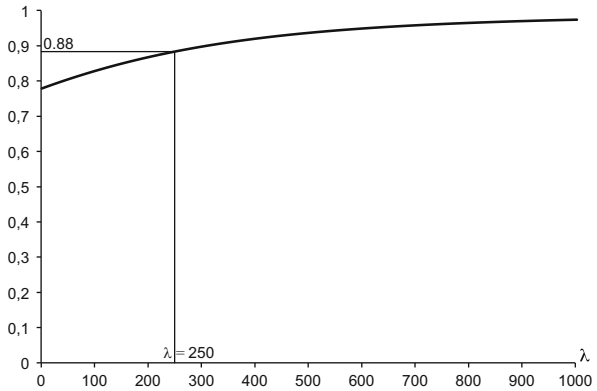

Fig. 2.1 The cost-effectiveness acceptability curve for the CADET-Hp trial

and variance $0.003356\lambda^2 + 4782 + 1.426\lambda$. The cost-effectiveness acceptability curve, given by $\Phi\left(0.1371\lambda + 53.01/\sqrt{0.003356\lambda^2 + 4782 + 1.426\lambda}\right)$, is shown in Fig. 2.1. Because *Treatment* is observed to increase effectiveness (i.e., $\hat{\Delta}_e > 0$) and decrease cost (i.e., $\hat{\Delta}_c < 0$), the INB will be positive and the CEAC will be greater than 0.5 for all positive values of the threshold value (λ).

Because the mean INB is positive regardless of the threshold value, and because *Treatment* is observed to reduce cost and therefore budget constraints may not be an issue, it may seem obvious that *Treatment* should be adopted. But this would ignore the uncertainty regarding the INB (i.e., $v_0 > 0$). Because of this uncertainty, there is a positive probability that the INB is negative. (For $\lambda = 250$, the probability that INB is negative, i.e., $1 - \text{CEAC}$ for $\lambda = 250$, is 0.12.) Therefore, there is a positive expected opportunity loss associated with the net benefit maximizing decision (action) to adopt *Treatment* and the optimal action might be to obtain more information (e.g., another trial) to reduce the uncertainty and decrease the expected opportunity loss. Whether or not another trial is optimal, and the optimal size of the trial if it is, will depend on the trade-offs between the additional cost and the reduction in expected opportunity loss. This is covered in the next section.

2.3 Decision Theory and Value of Information in RCT Research

2.3.1 Introduction

In response to the many problems associated with sample size determinations based on tests of hypotheses and power arguments, many authors have proposed alternatives [1, 3, 9, 11–16, 18–22, 25, 26, 28, 33–35, 43–49, 52]. In particular, among others, Willan and Pinto [43], Eckermann and Willan [14–16], Willan [44, 45], Willan and Kowgier [46], and Willan and Eckermann [47–49] propose methods

based on decision theory and the expected value of information that determines the sample size for maximizing the difference between the expected cost of the trial and the expected value of the information provided by the results. Fixed, variable and opportunity trial costs are considered. In addition to providing optimal sample sizes, these methods can identify circumstances when the current information is sufficient for decision making, see Willan [44]. Details of the approach are given below.

2.3.2 Opportunity Loss in Decision Making

To recognize the role that decision theory can play in the analysis and design of RCTs one must understand the definition of opportunity loss and how to determine its expected value. To that aim we use an example based on a simple bet on the toss of a (not necessarily fair) coin. The decision to accept the bet on the coin toss has an associated opportunity loss, and one can determine its expected value based on the current information regarding the outcome of a toss of the coin. The more information one has regarding the toss of the coin, the less is the expected opportunity loss. The chance to gather additional information should be accepted only if the cost of doing so is less than the reduction in the expected opportunity cost provided by the additional information. The reduction in the expected opportunity loss provided by additional information is referred to as the expected value of information (EVSI).

Suppose Karl has tossed a particular coin on 12 occasions and noted that it came up heads on 9 of them. He must now decide whether or not to accept the following bet: On a new toss of the coin, if it comes up heads he wins \$1,000 and if it comes up tails he loses \$1,000. Let the random variable $X = 1$ if the next toss of the coin is a head, and 0 otherwise. A reasonable *pdf* for X to reflect the uncertainty regarding the next toss (i.e., the value of X) is Bernulli(θ), given by $\Pr(X = x) = \theta^x(1-\theta)^{1-x}$, where θ is the probability that the next toss of the coin is a head. The utility of accepting the bet is \$1,000 if the toss is a head and −\$1,000 if it is tail, and as a function of X, equals $1{,}000X - 1{,}000(1 - X) = 1{,}000(2X - 1)$, with expectation $1{,}000(2\theta - 1)$. The utility of refusing the bet is zero, since nothing is gained or lost. Karl's previous experience with the coin has provided him with some knowledge regarding θ. In general, if Karl had observed r heads in n tosses, and assuming he had no other prior knowledge or opinions regarding θ, the posterior distribution for θ is Beta(a_0, b_0), where $a_0 = r + 1$ and $b_0 = n - r + 1$, with mean $a_0/(a_0+b_0)$ and variance $a_0b_0/\left\{(a_0 + b_0)^2(a_0 + b_0 + 1)\right\}$. The *pdf* and *cdf*, denoted by $f_B(\theta; a_0, b_0)$ and $F_B(\theta; a_0, b_0)$ respectively, are given by

$$f_B(\theta; a_0, b_0) = [(a_0 + b_0 - 1)!/\{(a_0 - 1)!(b_0 - 1)!\}]\,\theta^{a_0-1}(1-\theta)^{b_0-1} \quad \text{and}$$

$$F_B(\theta; a_0, b_0) = \sum_{j=a_0}^{a_0+b_0-1} \frac{(a_0 + b_0 - 1)!}{j!(a_0 + b_0 - 1 - j)!}\theta^j(1-\theta)^{a_0+b_0-1-j}$$

Therefore, Karl's current knowledge regarding θ after observing 9 heads in 12 tosses is characterized by a beta distribution with mean $10/14 = 0.7143$ and variance $10 \times 4/\left\{(10+4)^2(10+4+1)\right\} = 0.01361$, and his expected utility for the decision to accept the bet is $1{,}000(2 \times 0.7143 - 1) = 428.6$. Since his expected utility for the decision to refuse the bet is 0, he should accept the bet if he wants to maximize expected utility. Nonetheless, there is an opportunity loss associated with deciding to accept the bet. In general, the opportunity loss associated with a decision is the utility of the best decision *minus* the utility of the decision made. The opportunity loss of accepting the bet depends on whether the coin comes up heads or tails. If it comes up heads there is no opportunity loss because, in that case, accepting the bet is the best decision. If the coin comes up tails, the best decision would have been to refuse the bet. The utility of refusing the bet is zero, but the utility of accepting the bet when it comes up tails is −$1,000. Thus, the opportunity loss of accepting the bet when it comes up tails is the utility of refusing the bet *minus* the utility of accepting the bet, i.e., $0 - (-1{,}000) = \$1{,}000$. Consequently, Karl's opportunity loss function is $1{,}000 \times I(\text{coin comes up tails})$, where $I(\cdot)$ is the indicator function. Therefore, Karl's *expected* opportunity loss based on the current information (EOL_0) is given by

$$\text{EOL}_0 = 1{,}000 \times \Pr(\theta < 0.5) = 1{,}000 \times F_B(0.5; 10, 4). \text{ That is,}$$

$$\text{EOL}_0 = 1{,}000 \sum_{j=10}^{13} \frac{13!}{j!(13-j)!} 0.5^{13} = 46.14.$$

Therefore, based on current information, Karl faces an expected opportunity loss of $46.14 associated with the decision to accept the bet which is his expected utility-maximizing course of action.

Suppose Karl is given the opportunity to pay $20.00 to toss the coin 12 more times. The question is: Is the additional information worth $20.00? In decision theory that question is interpreted as: Will the additional information provided by 12 more tosses reduce the expected opportunity loss by more than $20.00? Suppose Karl tosses the coin 12 more times and observes r heads. The posterior distribution is $\text{Beta}(a_1, b_1)$, where $a_1 = a_0 + r = 10 + r$ and $b_1 = b_0 + (12 - r) = 16 - r$. The posterior expected opportunity loss if Karl observes r heads in 12 tosses is

$$\text{EOL}_1 = 1{,}000 \times \Pr(\theta < 0.5) = 1{,}000 \times F_B(0.5; 10 + r, 16 - r).$$

Since the expected opportunity loss is a function of the number of heads observed, the expected value of the expected opportunity loss must be taken with respect to the random variable *number of heads observed*, denoted Y. Thus, the expected opportunity loss including the new information provided by the 12 coin tosses (EOL_1) is given by

$$\begin{aligned}\text{EOL}_1 &= \sum_{r=0}^{12}\{1{,}000 \times F_B(0.5; 10+r, 16-r) \times \Pr(Y=r)\} \\ &= 1{,}000\sum_{r=0}^{12}\left\{\sum_{j=10+r}^{25}\frac{25!}{j!(25-j)!}0.5^{25}\frac{3!}{r!(3-r)!}\theta_0^r(1-\theta_0)^{12-r}\right\} \\ &= 29.87.\end{aligned}$$

where $\theta_0 = a_0/(a_0 + b_0) = 0.7143$, Karl's current mean of θ. Therefore, the expected value of sample information provided by the 12 coin tosses is $\text{EOL}_0 - \text{EOL}_1 = 46.14 - 29.87 = \$16.27$, which is less that the offered cost of \$20.00. Therefore, Karl's optimal action is to accept the bet based the information from the initial 12 tosses. To emphasize here, Karl's optimal decision is to accept the bet without paying for the additional information because the cost of the additional information exceeds the amount by which it would reduce the expected opportunity loss. As illustrated in later sections, a similar situation arises in evaluating evidence from a clinical trial. Because of the uncertainty inherent in the evidence, the decision to adopt the utility-maximizing intervention will be associated with an expected opportunity loss. Additional evidence should be sought only if the cost of attaining the evidence is the less than the amount by which it reduces the expected opportunity loss.

Suppose now that Karl was offered the opportunity, prior to deciding whether or not to accept the bet, to make as many tosses as he wished at \$0.50 a toss. If he took 12 tosses the \$6.00 cost would be less than the expected value of information of \$16.27. The question now is: What is the optimal number of tosses? The answer is: It is the number of tosses that maximizes the difference between the expected value of sample information and the cost of making the tosses. The difference between the expected value of information and the cost is referred to as the expected net gain (ENG). If we let m be the number of tosses taken, then the posterior expected opportunity loss is

$$\begin{aligned}\text{EOL}_1(m) &= \sum_{r=0}^{m}\{1{,}000 \times F_B(0.5; 10+r, 4+m-r) \times \Pr(Y=r)\} \\ &= 1{,}000 \times 0.5^{16} \times \sum_{r=0}^{m}\left\{\sum_{j=10+r}^{13+m}\frac{(13+m)!m!\theta_0^r(1-\theta_0)^{m-r}}{j!(13+m-j)!r!(m-r)!}\right\},\end{aligned}$$

where $\theta_0 = a_0/(a_0 + b_0) = 0.7143$, Karl's current mean of θ. Plots of the EVSI (i.e., $\text{EOL}_0 - \text{EOL}_1(m)$) and total cost (i.e., 0.5 m), as functions of m, are given in Fig. 2.2. By inspection, the difference between the expected value of information and total cost is maximized at 34 tosses, where the expected value of sample

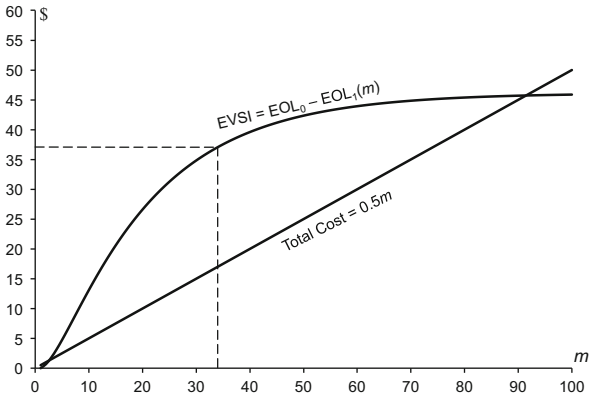

Fig. 2.2 The expected value of information and total cost for the coin toss example

information = \$37.06 and the total cost is \$17.00, yielding an expected net gain of \$20.06. As illustrated in later sections, an analogous situation arises when assessing the evidence from an RCT. One must decide to adopt the utility-maximizing intervention or, if given the chance, perform another trial which should be performed only if the maximum amount by which the expected opportunity loss is reduced (i.e., EVSI) exceeds the cost of the trial. That is, only if the maximum ENG is positive.

It may seem odd that the expected opportunity loss based on the initial 12 tosses is only \$46.14, given that Karl has a $1 - \theta_0 = 0.2957$ probability of losing \$1,000. But the expected opportunity loss relates to the uncertainty regarding θ, not its actual value. That is, if Karl knew for certain that the probability of heads is 0.55, his expected opportunity loss is zero, even though there is a 0.45 probability that he will lose \$1,000. The value of perfect information, when you have it, is zero. Karl accepts the bet if the expected value of the utility (i.e., $1{,}000(2\theta_0 - 1)$) is greater than zero because we have assumed Karl is risk-neutral, that is, a dollar lost has the same value as a dollar won. Being risk-neutral makes most sense if the bet can be accepted or refused numerous times, thus spreading the risk of any single bet over many others. Based on the information Karl has from the initial 12 tosses, the probability that he will lose money on a single toss is $1 - \theta_0 = 0.2857$. However on k tosses he will lose money only if less than a half of them are heads, that is, with probability

$$\sum_{r=0}^{k^*} \frac{k!}{r!(k-r)!} \theta_0^r (1-\theta_0)^{k-r},$$

where k^* is the largest integer less than $k/2$. So, for 10 tosses the probability of losing money is 0.03764 and for 20 tosses it is 0.01171.

2.3.3 The Expected Value of Sample Information

Consider the problem of determining the sample size for a randomized clinical trial designed to examine the cost-effectiveness of *Treatment* in comparison to *Standard*. The trial is conducted with the purpose of adopting *Treatment* if it is found to be cost-effective. *Treatment* is cost-effective if the INB is greater than zero. Recall from Sect. 2.2 that the INB is defined as $b \equiv \Delta_e \lambda - \Delta_c$, where λ is the threshold value for a unit of health outcome (effectiveness); $\Delta_e = e_T - e_S$, where e_j, for $j = T, S$, is the mean effectiveness for intervention j; and $\Delta_c = c_T - c_S$, where c_j, for $j = T, S$, is the mean cost for intervention j. Recall that $b \equiv e_T\lambda - c_T - (e_S\lambda - c_S)$, so that INB $= \text{NB}_T - \text{NB}_S$, where $\text{NB}_j (\equiv e_j\lambda - c_j)$ is the net benefit for intervention j.

In the following, the threshold value is initially considered fixed for ease of notation but can be allowed to vary, as demonstrated later when examining robustness. Let the current information regarding incremental net benefit be characterized by a normal prior *pdf* with mean b_0 and variance v_0, where $b_0 > 0$ and $v_0 > 0$. Since the prior mean INB (b_0) is positive, adopting *Treatment*, rather than retaining *Standard* maximizes the expected net benefit for future patients. However, since the prior variance of INB (v_0) is positive, adopting *Treatment* is not necessarily the optimum decision facing a decision maker. Consideration must be given to collecting more information, i.e., conducting another trial. Decision uncertainty resulting from a positive v_0 implies that a decision maker faces an opportunity loss when adopting *Treatment*, even though doing so is the decision that maximizes expected net benefit for future patients. The opportunity loss per patient associated with the decision to adopt *Treatment* is defined as the utility of the best decision *minus* the utility of adopting *Treatment*. Since, in this context, utility equals net benefit, the opportunity loss becomes the maximum of (NB_T, NB_S) *minus* NB_T. The maximum of (NB_T, NB_S) depends on b, the INB. If b is positive, then $\text{NB}_T > \text{NB}_S$, and NB_T is the maximum. On the other hand, if b is not positive, then $\text{NB}_T \leq \text{NB}_S$, and NB_S is the maximum. Thus the opportunity loss per patient associated with adopting *Treatment* (OLpp_T), as a function of INB, is given by:

$$\text{OLpp}_T(b) = \left\{ \begin{array}{ll} \text{Max}(\text{NB}_T, \text{NB}_S) - \text{NB}_T = \text{NB}_S - \text{NB}_T = -b & : b \leq 0 \\ \text{Max}(\text{NB}_T, \text{NB}_S) - \text{NB}_T = \text{NB}_T - \text{NB}_T = 0 & : b > 0 \end{array} \right\}.$$

When INB is positive there is no opportunity loss associated with adopting *Treatment* since future patients would receive the net benefit-maximizing intervention. However, if *Treatment* is adopted when incremental net benefit is negative, future patients would not receive the net benefit-maximizing intervention and each patient would experience a reduction in net benefit equal to the absolute value of INB. A plot of $\text{OLpp}_T(b)$ is given in Fig. 2.3.

Taking the expected value of $\text{OLpp}_T(b)$ with respect to the current information regarding incremental net benefit which, as assumed above, is characterized by a normal prior *pdf* with mean b_0 and variance v_0, yields the prior expected opportunity

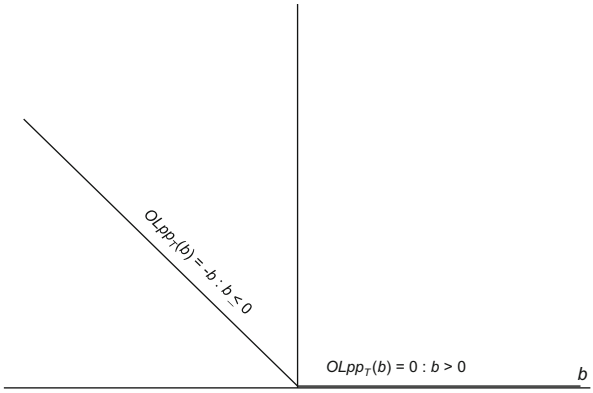

Fig. 2.3 The opportunity loss function per patient of adopting *Treatment*

loss per patient (EOLpp_{T0}). Letting $f_N(x;\mu,v)$ be the *pdf* for normal random variable with mean μ and variance v, then

$$\text{EOLpp}_{T0} = \int_{-\infty}^{\infty} \text{OLpp}_T(b)\, f_N(b;b_0,v_0)\text{d}b = \int_{-\infty}^{0} -b f_N(b;b_0,v_0)\text{d}b = \mathscr{D}(b_0,v_0)\,,$$

where

$$\mathscr{D}(\mu,v) = [v/(2\pi)]^{\frac{1}{2}} \exp\left[-\mu^2/(2v)\right] - \mu\left[\Phi(-\mu/v^{\frac{1}{2}}) - I(\mu \le 0)\right]; \qquad (2.1)$$

where $\Phi(\cdot)$ is the *cdf* for the standard normal random variable; and, $I(\cdot)$ is the indicator function, see Willan and Pinto [43] for details. The expected opportunity loss per patient, multiplied by the number of future patients, is the total expected opportunity loss and is also known as the expected value of perfect information, since if the decision maker had perfect information (i.e., $v_0 = 0$), the opportunity loss could be avoided by adopting *Treatment* if b_0 is positive and retaining *Standard*, otherwise. Applying decision theory, as illustrated in Sect. 2.3.2, the expected value of sample information (EVSI) of a new trial is the amount by which the information from the new trial reduces the total expected opportunity loss.

Suppose a new trial of n patients per arm is conducted where $\hat{\Delta}_e$ and $\hat{\Delta}_c$ are the respective estimators of Δ_e and Δ_c from the trial data. Thus, the estimate of INB based on the trial data is $\hat{b} = \hat{\Delta}_e\lambda - \hat{\Delta}_c$ and relying on the central limit theorem regarding the distribution of $\hat{b}$ the posterior mean and variance for incremental net benefit are given by:

$$b_1 = v_1\left(\frac{b_0}{v_0} + \frac{n\hat{b}}{\sigma_+^2}\right) \quad \text{and} \quad v_1 = \left(\frac{1}{v_0} + \frac{n}{\sigma_+^2}\right)^{-1},$$

where σ_+^2 is the sums over treatment arm of the between-patient variances of net benefit, and is assumed known or determinable from prior data. Details for

estimating σ_+^2 for the CADET-Hp trial are given in the Appendix. The posterior (i.e., post-trial) expected opportunity cost per patient is given by $\text{EOLpp}_1 = \mathscr{D}(b_1, v_1)$. EOLpp_1 is a function of the random variable $\hat{b}$ and to determine the expected reduction in per-patient opportunity loss, with the purpose of identifying the optimal sample size, the expectation of EOLpp_1 must be taken with respect to $\hat{b}$. Applying the central limit theorem, the predictive distribution for $\hat{b}$ is $N(b_0, v_{\hat{b}})$, where $v_{\hat{b}} = v_0 + \sigma_+^2/n$, and the expected value of EOLpp_1 with respect to $v_{\hat{b}}$ becomes, see Willan and Pinto [43],

$$\text{E}_{\hat{b}}\text{EOLpp}_1 = \text{E}_{\hat{b}}\mathscr{D}(b_1, v_1) = \int_{-\infty}^{\infty} \mathscr{D}(b_1, v_1) f(\hat{b}; b_0, v_{\hat{b}}) \text{d}\hat{b} = I_1 + I_2 + I_3, \text{ where}$$

$$I_1 = \sqrt{v_0/(2\pi)}\sigma_+^2 \exp\left(-b_0^2/2v_0\right)/(nv_{\hat{b}}),$$

$$I_2 = -b_0\Phi\left(-b_0/\sqrt{v_0}\right) + v_0^{3/2} \exp\left(-b_0^2/2v_0\right) \Big/ \left(v_{\hat{b}}\sqrt{2\pi}\right), \text{ and}$$

$$I_3 = b_0\Phi\left(-b_0\sqrt{v_{\hat{b}}}/v_0\right) - v_0 \exp\left(-b_0^2 v_{\hat{b}}/(2v_0^2)\right) \Big/ \sqrt{2\pi v_{\hat{b}}}.$$

Thus, the expected value of sample information of a trial of n patients per arm is given by

$$\text{EVSI}(n) = B(n)\left\{\mathscr{D}(b_0, v_0) - \text{E}_{\hat{b}}\mathscr{D}(b_1, v_1)\right\},$$

where $B(n)$ refers to the post-trial patient horizon, defined as the number of patients who could potentially receive the new intervention following the trial and therefore can benefit from a reduction in the opportunity loss. For an incidence rate of k patients per year, a time horizon of h years and a trial duration of $t(n)$ years, $B(n) = k\{h - t(n)\}$. The time horizon is the duration for which the decision to either adopt *Treatment* or perform another trial is relevant. Although there is no software packages for determining EVSI, it components can be calculated directly from the formulae using a spreadsheet.

2.3.4 Expected Total Cost

The cost of a trial is assumed to have two components, one financial and the other reflecting opportunity costs. Let C_f be the fixed financial cost of setting up a trial and let C_v be the variable financial cost per patient. Then the total financial cost of a trial with n patients per arm is $C_f + 2nC_v$. The assumption is made that since b_0 is positive, if the trial is not performed, all future patients would receive *Treatment*. This is referred to as the assumption of perfect implementation. It is also assumed that while the trial is performed, all patients outside the trial and half the patient within the trial will receive *Standard*. All patients who receive *Standard* while the

trial is performed, denoted as $D(n)$, pay an expected opportunity cost equal to b_0. The decision to perform the trial means that these patients have an expected reduction in net benefit equal to b_0 because they will receive *Standard* rather than *Treatment*. Therefore $D(n) = kt(n) - n$. That is, the number of patients who receive *Standard* because of the trial are all the patient who are incident while the trial is performed, *minus* the n patients who receive *Treatment* in the trial. Therefore, the expected total cost (ETC) of delaying the decision and performing the trial is $\text{ETC}(n) = C_f + 2nC_v + D(n)b_0$.

The function $t(n)$, an important part of the functions of $B(n)$ and $D(n)$, will depend on what assumptions are made regarding the proportion of patients that are recruited into the trial and the duration between when the last patient is randomized and when the trial results are available. These assumptions and their implications are discussed in Sect. 2.4 using the CADET-Hp trial as an example.

2.3.5 The Expected Net Gain and Optimal Sample Size

Given b_0, v_0, σ_+^2, h and k, the EVSI is a function of the sample size n, given as $\text{EVSI}(n) = B(n)\left\{\mathscr{D}(b_0, v_0) - \text{E}_{\hat{b}}\mathscr{D}(b_1, v_1)\right\}$. Likewise, given b_0, C_f and C_v, the expected total cost is a function of the sample size n, given as $\text{ETC}(n) = C_f + 2nC_v + D(n)b_0$. The expected net gain is defined as $\text{ENG}(n) \equiv \text{EVSI}(n) - \text{ETC}(n)$. Considering the trial in isolation, and being free of budget constraints, let n^* be that value of n that maximizes the expected net gain. That is, $\text{ENG}(n^*) \geq \text{ENG}(n)$ for all positive integers n. If $\text{ENG}(n^*) \leq 0$ then optimal sample size is zero and the current information, i.e., b_0 and v_0, is sufficient for decision making. In this case no trial is necessary, since the expected value of the information from the trial is less than the expected total cost, regardless of the sample size. On the other hand, if $\text{ENG}(n^*) > 0$, the decision maker is in a state of equipoise and the optimal decision is to delay adopting *Treatment*, even though $b_0 > 0$, and perform a trial with n^* patients per arm.

2.4 Applying VOI Methods: The CADET-Hp Trial

Suppose, for sake of illustration, the threshold value of a treatment success is \$250, i.e., $\lambda = 250$. Assuming no prior information, the current mean and variance for incremental net benefit is given by:

$$b_0 = 0.1371 \times 250 - (-53.01) = 87.285,$$

$$v_0 = 0.003356 \times 250^2 + 4{,}792 - 2 \times (-0.7129) \times 250 = 5{,}344.7,$$

and, invoking the central limit theorem, the prior *pdf* for INB for the planning of a new trial is $N(87.285, 5{,}344.7)$. The probability that *Treatment* is cost-effective for $\lambda = 250$ is 0.88, see Fig. 2.1. Since $b_0 > 0$ the net benefit maximizing decision, based on current evidence, is to adopt *Treatment* (i.e., add the antibiotics) for future patients. However, since $v_0 > 0$, the decision to adopt *Treatment* is associated with an expected per-patient opportunity loss of $\mathscr{D}(87.285, 5{,}344.6) = 4.1528$ (from Eq. 2.1), and the optimal decision might be to delay the adoption of *Treatment* and perform another trial. Performing another trial would be optimal if the reduction in total expected opportunity loss (i.e., the expected value of sample information) is greater than the expected total cost, that is, if the expected net gain is greater than zero.

2.4.1 Simplifying Assumptions

If we make the simplifying assumptions that all patients in the jurisdiction of interest are recruited into the trial and that the results of the trial are available immediately after the last patient is randomized, then duration of the trial will equal total sample size *divided* by the incidence (i.e., $t(n) = 2n/k$). Under the same assumptions the number of patients that can benefit from the new information (B) will equal the total patient horizon (i.e., kh) minus the $2n$ patients in the trial (i.e., $B(n) = kh - 2n$). The time horizon of a decision is the duration over which the decision is considered relevant. Assuming an incidence of 80,000 per year and a time horizon of 20 years, the plots of EVSI, ETC and ENG as functions of n are given in Fig. 2.4. The fixed (C_f) and variable (C_v) financial cost of the trial were assumed to be \$800,000 and \$2,000, respectively. The optimal sample size is 463 patients per arm, yielding an optimum ENG of \$1,329,020 with an ETC of \$2,692,413 for a return on investment of 49 %.

Willan et al. [51] demonstrates that plotting the combinations of the threshold value (λ) and horizon (h) for which the ENG is zero provides a sensitivity analysis for those variables, see Fig. 2.5. For combinations of λ and h above the curve the

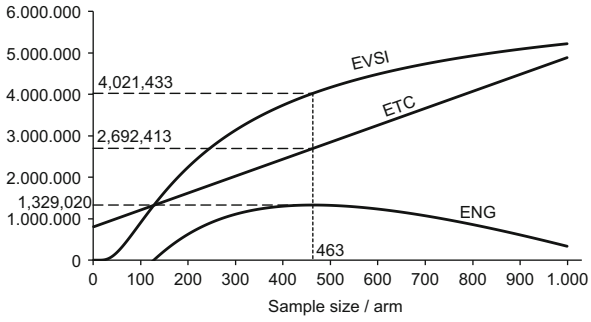

Fig. 2.4 EVSI, ETC and ENG for the CADET-Hp example using the unrealistic assumptions

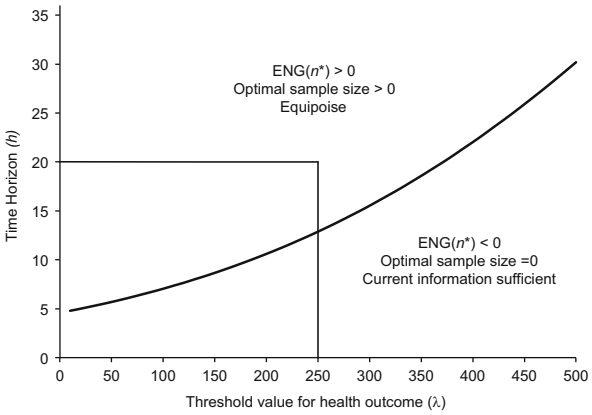

Fig. 2.5 The values of the threshold value for health outcome and horizon for which the ENG is zero for the CADET-Hp example using the unrealistic assumptions

ENG is positive (i.e., a state of equipoise exists) and a new trial is the optimal decision. Whereas, for combinations below the curve the ENG is negative and the current evidence is sufficient for decision making. Note that the combination of $\lambda = 250$ and $h = 20$ lies above the line.

2.4.2 More Realistic Assumptions

In Sect. 2.4.1 it was assumed that all patients in the jurisdiction of interest are recruited into the trial and that the trial results are available immediately after the last patient is randomized. These assumptions almost never hold. Usually only a small fraction of the eligible patients are recruited, and patients need to be followed to observe outcomes. Further, time is required for data entry, cleaning and analysis. If we let the annual accrual rate be denoted by a and the number of years between when the last patient is randomized and the data is analysed be denoted by τ, the trial duration becomes $t(n) = \tau + 2n/a$. Consequently, the number of patients who will benefit from the trial results (B) and the number of patient incurring an opportunity cost (D) are given by:

$$B(n) = k\{h - (2n/a + \tau)\}$$
$$D(n) = k(2n/a + \tau) - n$$

For the CADET-Hp example, if we assume an accrual fraction of 1 % (i.e., $a = 800$ per year), and allow for 1 year of follow-up (necessary to observe the measure of effectiveness) with 3 months for data entry, cleaning and analysis (i.e., $t = 1.25$), the optimal sample size is zero. A plot of the expected total cost and expected value of sample information is given in Fig. 2.6, where it can be seen that costs exceeds value for all sample sizes. The expected total cost have been

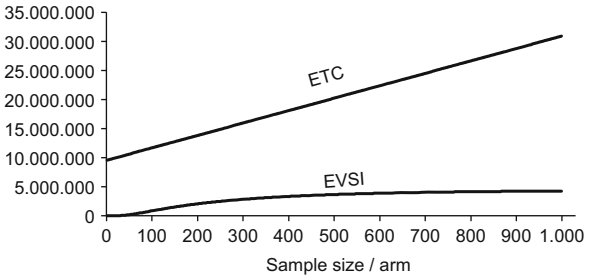

Fig. 2.6 EVSI and ETC for the CADET-Hp example using the realistic assumptions

driven up by the very high expected opportunity cost for the patients who receive *Standard* while the trial is conducted. These patients consist of 99 % of incident cases while the trial is recruiting patients and 100 % during the follow-up period. The expected opportunity cost for the 1.25 years of follow-up alone is $8,728,500. Further, because the trial takes longer to perform the number of patients that can benefit from the new information is reduced, which in turn reduces the EVSI. For details, the reader is referred to Eckermann and Willan [15, 16].

2.4.3 Relaxing the Assumption of Perfect Implementation

For the solution above and in Sect. 2.4.1 the assumption has been made that if the current mean INB is positive (i.e., $b_0 > 0$) that all future patients would receive *Treatment* in the absence of a new trial. Referred to as perfect implementation this assumptions is unlikely to hold. To examine the effect of allowing imperfect implementation Willan and Eckermann [48] assume that the probability that a future patient that would receive *Treatment* if no additional evidence is forthcoming, is a non-decreasing function of the strength of the evidence as measured by the z-statistic, defined as $z_i = b_i / \sqrt{v_i}$, $i = 0, 1$. To demonstrate the dramatic effect that this more realistic assumption has on the solution, the authors use a *sliding step function*, where if $z_i \leq \gamma$, the probability that a future patient would receive *Treatment* is 0, and if $z_i \geq \beta$, the probability that a future patient would receive *Treatment* is 1, where $\gamma \leq \beta$. For $\gamma < z_i < \beta$, a linear function is assumed, where the probability that a future patient receives *Treatment* is $(z_i - \gamma)/(\beta - \gamma)$. For perfect implementation, $\gamma = \beta = 0$.

Relaxing the assumption of perfect implementation can have a dramatic effect on the value of information solution. Firstly, additional information, apart from reducing the expected opportunity cost as before, now has value in increasing the expected proportion of future patients receiving the net benefit maximizing intervention (i.e., $\mathrm{E}(z_1) > z_0$). Secondly, the expected opportunity cost of delaying the decision to adopt *Treatment* and performing a future trial is far less since only a portion of the patients would receive *Treatment* in the absence of the future trial and therefore incur an expected opportunity cost.

To demonstrate the effect on the solution using the CADET-Hp example suppose γ and β are chosen to correspond with values of the probability of *Treatment* being cost-effective of 75 and 99 %, respectively, that is, $\gamma = \Phi^{-1}(0.75) = 0.675$ and $\beta = \Phi^{-1}(0.99) = 2.326$. (Recall that for normally distributed incremental net benefit, the CEAC is given by $\Phi(z\text{-statistic})$.) Based on the evidence from the existing trial

$$z_0 = b_0 \big/ \sqrt{v_0} = 87.28 \big/ \sqrt{5{,}345} = 1.194,$$

and the probability of a future patient receiving *Treatment* in the absence of a new trial is

$$(z_0 - \gamma)/(\beta - \gamma) = (1.194 - 0.6745)/(2.326 - 0.6745) = 0.3135.$$

Consequently, the expected opportunity cost of performing new trial is less than a third of what it is under the assumption of perfect implementation.

Figure 2.7 contains a plot of the expected value of information and expected total cost as a function of sample size assuming imperfect implementation as characterized by the value of γ and β given above and the same assumption regarding accrual and follow-up given in Sect. 2.4.2. The optimal sample size is 486, with an expected net gain of $64,299,751. Compared to Fig. 2.6 a dramatic increase in the expected value of information is observed. This is because the new information, apart from reducing the total expected opportunity cost, is expected to increase the proportion of future patients receiving *Treatment*. For the optimal sample size of 486 the post-trial expected probability that a future patient receives Treatment is 0.7. Also observed is a dramatic decrease in the expected total cost, which is a result of the reduction in expected opportunity cost as noted above.

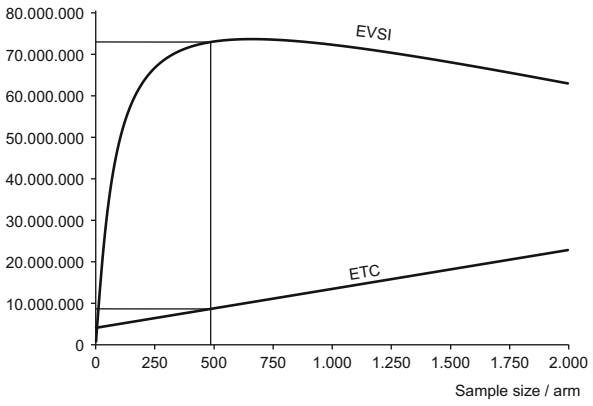

Fig. 2.7 EVSI and ETC for the CADET-Hp example using the realistic assumptions with imperfect implementation

2.5 Discussion

In this chapter the application of decision theory and associated value of information (VOI) methods for the design and analysis of RCTs have been proposed as an alternative to the standard hypothesis testing approach with its reliance on arbitrarily chosen Type I and II error probabilities and smallest clinically important differences. VOI methods allow for the explicit incorporation of important factors, such as the value of health outcomes, incidence and accrual rates, time horizon, current information, follow-up times and trial costs. They also can be used to identify those situations where the evidence is sufficient for decision making and, where evidence is insufficient (equipose), the optimal size of a future trial. Using VOI methods to assess the evidence from a clinical trial or the meta-analysis of several trials provides a more rational alternative to the standard methods since they maximize the expected net benefit for future patients and optimize the allocation of research funding, while providing an operational definition for equipoise. In addition, since the EVSI increases with incidence, interventions for rarer diseases need less evidence for adoption. Thus VOI methods help address the obvious difficulty of patient recruitment in rare diseases.

The use of VOI methods raises a number of issues. Perhaps the most subtle is that of jurisdiction. The assumption is made that trial financial costs are borne by society through government or private donation-based or philanthropic agencies. This raises an issue for research funding agencies. On whose behalf is it acting? The answer to this question has a huge impact on VOI methods since it determines the incidence, which is an important determinant of EVSI. Agencies acting on behalf of small jurisdictions, such as provincial/state governments or health insurers, are less likely to find the funding of additional trials attractive, since the optimal sample size will be zero with greater frequency. However, for federal governments or private donation-based or philanthropic agencies, which may take a more global view, funding additional trials may be more attractive.

Typically VOI methods are based on the assumptions that if a new trial is carried out, the definitive decision regarding the adoption of *Treatment* will be made at the end of the trial. However, the truly optimal procedure would be to repeat the VOI process at the end of the new trial to determine if the updated evidence is sufficient. Relaxing this assumption leads to multi-stage designs as discussed in Willan and Kowgier [46].

The limitations of VOI methods have mostly to do with specifying values for the required parameters. The parameter incidence should be available from the literature, and is generally required to establish the burden of the health condition under study regardless of what methods are used to determine the sample size. Similarly, regardless of the methodology used, the financial cost and accrual rate are needed for planning and budgetary reasons. The parameters that could be considered specific to VOI methods are the threshold value for a unit of health outcome and the time horizon. Various threshold values for a quality-adjusted life-year have been applied in cost-utility analyses, however threshold values for other health outcomes are less well established. The time horizon for a new health care intervention varies

depending on the type of intervention (e.g., pharmacological, surgical) and the health condition under study. Time horizons of 20–25 years are often used because they correspond to infinite time horizon with discount rates for future benefits of around 4 or 5 %. It is worth noting that the advantage of VOI methods is that they make the assumptions regarding threshold value of health outcome and time horizon explicit, and although both parameters may be associated with uncertainty, a sensitivity analysis can be performed, as illustrated in the example.

In conclusion, decision theoretic/VOI methods can be used to identify those situations where the evidence is sufficient for decision making, and where evidence is insufficient, the optimal size of a future trial.

Appendix

Let e_{ji} and c_{ji} be the respective observations of effectiveness and cost for patient i receiving intervention j, where $j = T, S$; $i = 1, 2 \ldots n_j$; and n_j is the number of patients on intervention j. For the CADET-Hp trial $e_{ji} = 1$ if the patient is a success, 0 otherwise.

$$\text{Let} \quad \bar{e}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} e_{ji} \quad \text{and} \quad \bar{c}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} c_{ji}$$

Then

$$\begin{aligned}
\hat{\Delta}_e &= \bar{e}_T - \bar{e}_S \\
\hat{\Delta}_c &= \bar{c}_T - \bar{c}_S \\
\hat{V}(\hat{\Delta}_e) &= \hat{V}(\bar{e}_T) + \hat{V}(\bar{e}_S) = \frac{\bar{e}_T(1-\bar{e}_T)}{n_T} + \frac{\bar{e}_S(1-\bar{e}_S)}{n_S} \\
\hat{V}(\hat{\Delta}_c) &= \hat{V}(\bar{c}_T) + \hat{V}(\bar{c}_S) = \frac{\sum_{i=1}^{n_T}(c_{Ti} - \bar{c}_T)}{n_T(n_T-1)} + \frac{\sum_{i=1}^{n_S}(c_{Si} - \bar{c}_S)}{n_S(n_S-1)} \\
\hat{C}(\hat{\Delta}_e, \hat{\Delta}_c) &= \hat{C}(\bar{e}_T, \bar{c}_T) + \hat{C}(\bar{e}_S, \bar{c}_S) \\
&= \frac{\left(\sum_{i=1}^{n_T} e_{Ti}c_{Ti}\right) - n_T\bar{e}_T\bar{c}_T}{n_T(n_T-1)} + \frac{\left(\sum_{i=1}^{n_S} e_{Si}c_{Si}\right) - n_S\bar{e}_S\bar{c}_S}{n_S(n_S-1)}
\end{aligned}$$

and

$$\hat{\sigma}^2_+ = \sum_{j=T,S} n_j \left\{\lambda^2 \hat{V}(\bar{e}_j) + \hat{V}(\bar{c}_j) - 2\lambda \hat{C}(\bar{e}_j, \bar{c}_j)\right\},$$

where $\hat{\sigma}^2_+$ is the sum of treatment arms of the between-patient variance of incremental net benefit.

References

1. Adcock, C.J.: Sample size determination: a review. *The Statistician* **46**, 261–283 (1997)
2. Ament, A., Baltussen, R.: The interpretation of results of economic evaluation: Explicating the value of health. *Health Economics* **6**, 625–635 (1997)
3. Berry, D.A., Ho, C-H.: One-sided sequential stopping boundaries for clinical trials: A decision-theoretic approach. *Biometrics* **44**, 219–227 (1988)
4. Briggs, A.H., Wonderling, D.E., Mooney, C.Z.: Pulling cost-effectiveness analysis up by its bootstraps; a non-parametric approach to confidence interval estimation. *Health Economics* **6**, 327–340 (1997)
5. Briggs, A.H., Fenn, P.: Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Economics* **7**, 723–740 (1998)
6. Briggs, A.H., Mooney, C.Z., Wonderling, D.E.: Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using monte carlo simulation. *Statistics in Medicine* **18**, 3245–3262 (1999)
7. Briggs, A.H.: A Bayesian approach to stochastic cost-effectiveness analysis. *Health Economics* **8**, 257–261 (1999)
8. Chaudhary, M.A., Stearns, S.C.: Confidence intervals for cost-effectiveness ratios: An example from a randomized trial. *Statistics in Medicine* **15**, 1447–1458 (1996)
9. Cheng, Y., Su, F., Berry, D.A.: Choosing sample size for a clinical trial using decision analysis. *Biometrika* **90**, 923–936 (2003)
10. Chiba, N., van Zanten, S.J., Sinclair, P., Ferguson, R.A., Escobedo, S., Grace, E.: Treating Helicobacter pylori infection in primary care patients with uninvestigated dyspepsia: the Canadian adult dyspepsia empiric treatment-Helicobacter pylori positive (CADET-Hp) randomised controlled trial. *British Medical Journal* **324**, 1012–1016 (2002)
11. Claxton, K., Posnett, J.: An economic approach to clinical trial design and research priority setting. *Health Economics* **5**, 513–524 (1996)
12. Claxton, K.: The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics* **18**, 341–364 (1999)
13. Claxton, K., Thompson, K.M.: A dynamic programming approach to the efficient design of clinical trials. *Journal of Health Economics* **20**, 797–822 (2001)
14. Eckermann, S., Willan, A.R.: Expected value of information and decision making in HTA. *Health Economics***16**, 195–209 (2007)
15. Eckermann, S., Willan, A.R.: Time and EVSI wait for no patient. *Value in Health* **11**, 522–6 (2008)
16. Eckermann, S., Willan, A.R.: The option value of delay in health technology assessment. *Medical Decision Making* **28**, 300–305 (2008)
17. Fenwick, E., O'Brien, B., Briggs, A.: Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions. *Health Economics* **13**, 405–415 (2004)
18. Gittins, J.: Quantitative methods in the planning of pharmaceutical research. *Drug Information Journal* **30**, 479–487 (1996)
19. Gittins, J., Pezeshk, H.: How large should a trial be? *The Statistician* **49**, 177–197 (2000)
20. Gittins, J., Pezeshk, H.: A behavioral Bayes method for determining the size of a clinical trial. *Drug Information Journal* **34**, 355–363 (2000)
21. Grundy, P.M., Healy, M.J.R., Rees, D.H.: Economic choice of the amount of experimentation. *Journal of the Royal Statistical Society: Series B* **18**, 32–48 (1956)
22. Halpern, J., Brown, Jr B.W., Hornberger, J.: The sample size for a clinical trial: a Bayesian-decision theoretic approach. *Statistics in Medicine* **20**, 841–858 (2001)
23. Hannah, M.E., Hannah, W.J., Hewson, S.H., Hodnett, E.D., Saigal, S., Willan, A.R.: Term Breech Trial: a multicentre international randomised controlled trial of planned caesarean section and planned vaginal birth for breech presentation at term. *The Lancet* **356**, 1375–1383 (2000)
24. Heitjan, D.F.: Fieller's method and net health benefit. *Health Economics* **9**, 327–335 (2000)

25. Hornberger, J.C., Brown, Jr B.W., Halpern, J.: Designing a cost-effective clinical trial. *Statistics in Medicine* **14**, 2249–2259 (1995)
26. Hornberger, J., Eghtesady, P.: The cost-benefit of a randomized trial to a health care organization. *Controlled Clinical Trials* **1p**, 198–211 (1998)
27. Laska, E.M., Meisner, M., Siegel, C.: Statistical inference for cost-effectiveness ratios. *Health Economics* **6**, 229–242 (1997)
28. Lindley, D.V.: The choice of sample size. *The Statistician* **46**, 129–138 (1997)
29. Lothgren, M., Zethraeus, N.: Definition, interpretation and calculation of cost-effectiveness acceptability curves. *Health Economics* **9**, 623–630 (2000)
30. Manning, W.G., Fryback, D.G., Weinstein, M.C. (1996) Reflecting uncertainty in cost effectiveness analysis. In Gold MR, Siegel JE, Russell LB, Weinstein MC (eds) Cost Effectiveness in Health and Medicine, *Oxford University Press*, New York
31. Mullahy, J., Manning, W. (1994) Statistical issues of cost-effectiveness analysis. In Sloan F (ed) Valuing Health Care, *Cambridge University Press*, Cambridge
32. O'Brien, B.J., Drummond, M.F., Labelle, R.J., Willan, A.R.: In search of power and significance: Issues in the design and analysis of stochastic cost-effectiveness studies in health care. *Medical Care* **32**, 150–163 (1994)
33. O'Hagan, A., Stevens, J.W.: Bayesian assessment of sample size for clinical trials of cost effectiveness. *Medical Decision Making* **21**, 219–230 (2001)
34. Pezeshk, H., Gittins, J.: A fully Bayesian approach to calculating sample sizes for clinical trials with binary response. *Drug Information Journal* **36**, 143–150 (2002)
35. Pezeshk, H.: Bayesian techniques for sample size determination in clinical trials: a short review. *Statistical Methodology in Medical Research* **12**, 489–504 (2003)
36. Phelps, C.E. and Mushlin, A.I.: On the (near) equivalence of cost-effectiveness and cost-benefit analysis. *International Journal of Technology Assessment in Health Care* **7**, 12–21 (1991)
37. Polsky, D., Glick, H.A., Willke, R., Schulman, K.: Confidence intervals for cost-effectiveness ratios: A comparison of four methods. *Health Economics* **6**, 243–252 (1997)
38. Stinnett, A.A., Mallahy, J.: Net health benefits: A new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical Decision Making* **18**(Suppl), S68–S80 (1998)
39. Tambour, M., Zethraeus, N., Johannesson, M.: A note on confidence intervals in cost-effectiveness analysis. *International Journal of Technology Assessment* **14**, 467–471 (1998)
40. van Hout, B.A., Al, M.J., Gordon, G.S., Rutten, F.F.H..: Costs, effects and C/E ratios alongside a clinical trial. *Health Economics* **3**, 309–319 (1994)
41. Wakker, P., Klaassen, M.P.: Confidence intervals for cost/effectiveness ratios. *Health Economics* **4**, 373–381 (1995)
42. Willan, A.R., Cruess, A.F., Ballantyne, M.: Argon green vs krypton red laser photocoagulation of extrafoveal choroidal neovascular lesions: Three-year results in age-related macular degeneration. *Canadian Journal of Ophthalmology* **31**, 11–7 (1996)
43. Willan, A.R., Pinto, E.M.: The expected value of information and optimal clinical trial design. *Statistics in Medicine* **24**, 1791–1806 (2005). Correction: Statistics in Medicine 2006;25:720
44. Willan, A.R.: Clinical decision making and the expected value of information. *Clinical Trials* **4**, 279–285 (2007)
45. Willan, A.R.: Optimal sample size determinations from an industry perspective based on the expected value of information. *Clinical Trials* **5**, 587–594 (2008)
46. Willan, A.R., Kowgier, M.E.: Determining optimal sample sizes for multi-stage randomized clinical trials using value of information methods. *Clinical Trials* **5**, 289–300 (2008)
47. Willan, A.R., Eckermann, S.: Optimal clinical trial design using value of information methods with imperfect implementation. *Health Economics* **19**, 549–561 (2010)
48. Willan, A.R., Eckermann, S.: Accounting for between-study variation in incremental net benefit in value of information methodology. *Health Economics* **21**, 1183–1195 (2012)
49. Willan, A.R., Eckermann, S.: Value of information and pricing new health care interventions. *PharmacoEconomics* **30**, 447–459 (2012)

50. Willan, A.R., O'Brien, B.J.: Confidence intervals for cost-effectiveness ratios: An application of Fieller's theorem. *Health Economics* **5**, 297–305 (1996)
51. Willan, A.R.: Incremental net benefit in the analysis of economic data from clinical trials with application to the CADET-Hp Trial. *European Journal of Gastroenterology and Hepatology* **16**, 543–549 (2004)
52. Willan, A.R., Goeree, R., Boutis, K. (2012) Value of Information Methods for Planning and Analyzing Clinical Studies Optimize Decision Making and Research Planning. *Journal of Clinical Epidemiology* doi: http://dx.doi.org/10.1016/j.jclinepi.2012.01.017
53. Willan, A.R.: Analysis, sample size and power for estimating incremental net health benefit from clinical trial data. *Controlled Clinical Trials* **22**, 228–237 (2001)
54. Willan, A.R., Lin, D.Y.: Incremental net benefit in randomized clinical trials. *Statistics in Medicine* **20**, 1563–1574 (2001)
55. Willan, A.R., Lin, D.Y., Cook, R.J., Chen, E.B.: Using inverse-weighting in cost-effectiveness analysis with censored data. *Statistical Methods in Medical Research* **11**, 539–551 (2002)
56. Willan, A.R., Chen, E.B., Cook, R.J., Lin, D.Y.: Incremental net benefit in clinical trials with quality-adjusted survival. *Statistics in Medicine* **22**, 353–362 (2003)
57. Willan, A.R., Briggs, A.H. (2006) The Statistical Analysis of Cost-effectiveness Data. Wiley, Chichester UK