A MODEL FOR SELECTING ONE OF TWO MEDICAL TREATMENTS*

THEODORE COLTON
Harvard Medical School

A simple cost function approach is proposed for designing an optimal clinical trial when a total of N patients with a disease are to be treated with one of two medical treatments. The cost function is constructed with but one cost, the consequences of treating a patient with the superior or inferior of the two treatments. Fixed sample size and sequential trials are considered. Minimax, maximin, and Bayesian approaches are used for determining the optimal size of a fixed sample trial and the optimal position of the boundaries of a sequential trial. Comparisons of the different approaches are made as well as comparisons of the results for the fixed and sequential plans.

1. INTRODUCTION

IN THE planning of medical experiments to assess the therapeutic efficacy of new drugs or treatments, a most important question is how large to make the trial. On the one hand one wants as few patients as possible to participate so that the number of patients receiving the inferior treatment during the trial is minimized, the trial is brought to as speedy a conclusion as possible, and the results are quickly made available to aid in the treatment of other patients with the disease in question. On the other hand, enough patients must participate so that one can be reasonably certain that the truly superior treatment is selected and its subsequent use is appropriate.

The classical approach, of course, is to require the experimenter to arbitrarily assign errors of the first and second kind and to then determine the sample size accordingly for a fixed size trial or to determine the location of the boundaries for a sequential trial, be it an "open" type Wald plan [10] or one of the more recent "closed" sequential plans that Armitage [1] has suggested.

The difficulty with this approach is the arbitrariness of error levels and corresponding values of the population parameters. In many cases it is indeed difficult for the medical experimenter to select a difference and to state with what probability he wants to detect this difference.

As an alternative, it seems reasonable to approach the problem from the point of view of the consequences of decisions made, i.e., to use a cost function, decision theory approach to this problem. This paper is an attempt in this direction.

Clearly, this approach is not without its difficulties. Chief among these is it is often impossible to specify the costs involved. Asking the experimenter to assign the relative values to the elements comprising a complete cost formulation of the problem might be more difficult than asking him to select the necessary values for the classical approach. This difficulty is avoided in this research by proposing that as a first approximation there is but a single cost involved, the

* Paper read at the 33rd session of the International Statistical Institute, Paris, September, 1961. This research was undertaken while on a National Science Foundation Postdoctoral Fellowship.

consequence of treating a patient with the superior or the inferior of the two treatments, and all other costs may be disregarded.

There are other difficulties with this approach and these will become evident as the problem is formulated in the following section. Nevertheless, this approach is convenient, not too unreasonable in its assumptions, and a useful way to attack the problem.

2. FORMULATION OF THE PROBLEM

The problem, in its statistical formulation, is to determine the optimal procedure for choosing between two populations. In recent years there have been a number of papers on this subject [2, 3, 4, 5, 6, 9]. In constructing mathematical models and cost functions these authors have considered both the consequences of right and wrong decisions and the costs of experimentation. With a clinical trial it is difficult, often impossible, to have both these elements in the same model, since the two types of costs are largely incommensurate. From an ethical point of view the consequences of right and wrong decisions are the principal concern and, as we stated in the preceding section, we disregard all other costs and concentrate solely on the consequences of treating a patient with the superior and inferior of the two treatments.

Such an approach based solely on consequences of wrong decisions has been considered by Maurice [7], though not in a medical setting. We consider further exploration of this formulation with the design of medical trials as the objective.

It is assumed that N patients with a disease are to receive one of two treatments whose effects are unknown. For example, N might be the number of patients with a particular type of cancer in a particular time period. We assume N to be large but finite. The finite N is reasonable in that the decision between the two treatments is not everlasting. We shall use the selected treatment at least until some new drug has reached the stage of testing in a clinical trial, or until some new form of therapy has been developed. Under stable conditions we should know how many new cases arise per unit time period, and we should have some rough idea of the time elapsing before another treatment was suitable for testing; hence we could obtain at least some idea of N.

A clinical trial will be performed on $2n$ of the N individuals, n on each treatment. The remaining $N-2n$ patients will receive the treatment selected as the better at the conclusion of the trial. (We first consider $2n$ fixed and later consider the situation where it is not fixed but a sequential procedure is followed.)

It is assumed that we obtain a quantitative measure of response for each individual. The individual responses to each treatment are assumed to be normally distributed with unknown mean (μ_A for treatment A and μ_B for treatment B) and known variance σ^2 (the same for the two treatments). We assume that higher response is associated with better effect. Then, letting $\delta = \mu_A - \mu_B$, we should like the trial to select Treatment A if δ is positive and Treatment B if δ is negative.

In formulating the consequences of right and wrong decisions, we consider two approaches which differ in location of a base line. One is the "Cost" or

"Loss," approach (also referred to as "Regret"). For this we assume that for each individual receiving the inferior treatment we incur a loss or cost directly proportional to δ, the true difference between the two treatments. (Without loss of generality we take δ as positive.) We then obtain the expected loss for all N individuals,

$$E\ Loss = C\delta[n + (N - 2n)\ \Pr(\text{Select inferior})] \tag{1}$$

where C is a proportionality factor. Having the cost linearly related to δ seems a reasonable procedure for this first over-all approximation. Surely, more complicated functions of δ can be used, though at the price of increased difficulty of the subsequent algebra.

The other approach is the "Net Gain" approach. For this we assume that each time an individual receives the superior treatment we gain in direct proportion to δ, while if receiving the inferior treatment we lose in direct proportion to δ (i.e., we have a negative gain). We then get the expected net gain for all N individuals,

$$E\ Net\ Gain = G\delta(N - 2n)[\Pr(\text{Select superior}) - \Pr(\text{Select inferior})] \tag{2}$$

where G is a proportionality factor. It is easily seen that

$$E\ Net\ Gain = G\delta N - \frac{2G}{C} E\ \text{Loss}. \tag{3}$$

The problem is to determine n so that we minimize expected loss in the first case and maximize expected net gain in the second case. In both cases we have functions involving n and δ. One way to proceed is by a minimax method with expected loss (the procedure used by Maurice [7]) and a maximin method with expected net gain. These lead to different solutions.

The other method for handling δ is to assume an *a priori* distribution and then to integrate out for δ in expected loss and expected net gain. In this case the loss and net gain approaches lead to an identical result.

For a sequential trial analogous expected loss and expected net gain functions are constructed. The minimax method with expected loss, maximin method with expected net gain, and an *a priori* distribution for δ are considered.

Finally, a comparison is made of the optimal fixed sample size trials to optimal sequential trials.

3. THE OPTIMAL FIXED SAMPLE SIZE CLINICAL TRIAL

3.1 Minimax Method with Expected Loss

With samples of n observations from each of two normal populations with unknown means, it is well known that the optimal procedure for choosing the population with the larger mean is to compute the difference in sample means, $\bar{d} = \bar{x}_A - \bar{x}_B$, and to select the population corresponding to A if $\bar{d}$ is positive and that corresponding to B if $\bar{d}$ is negative. Since $\bar{d}$ is normally distributed with mean δ and variance $2\sigma^2/n$, then

$$\Pr(\text{Select inferior}) = F\left(\frac{-\delta\sqrt{n}}{\sigma\sqrt{2}}\right)$$

where

$$F(x) = \int_{-\infty}^{x} f(t)dt \quad \text{and} \quad f(t) = (2\pi)^{-1/2}\exp(-t^2/2).$$

The expected loss becomes

$$E\,Loss = C\delta\left[n + (N - 2n)F\left(\frac{-\delta\sqrt{n}}{\sigma\sqrt{2}}\right)\right].$$

Letting $p=n/N$ it is more convenient to consider

$$E\,Loss/N = C\delta\left[p + (1 - 2p)F\left(\frac{-\delta\sqrt{Np}}{\sigma\sqrt{2}}\right)\right] \tag{4}$$

and to determine the optimum p.

R. Maurice [7] has applied the minimax method to (4). The difficulty, as she notes, is that expected loss cannot be minimized for the most unfavorable value of δ since the loss is infinite when δ is infinite. It is possible though to determine a local minimax. Differentiating (4) with respect to δ and p and setting the derivatives equal to zero, gives the two equations

$$(1 - p)/(1 - 2p) = F(x) + xf(x) \tag{5}$$

$$(1 - 2p)/(2p) = [2F(x) - 1]/[xf(x)] \tag{6}$$

where

$$x = \frac{\delta\sqrt{Np}}{\sigma\sqrt{2}}.$$

Solving numerically gives $p=.10225$, $x=1.3729$, or $\delta=4.262(\sigma\sqrt{2})/\sqrt{N}$. Examination of second derivatives verifies location of a local minimax.

Suppose we know an upper bound for δ. We should have to know the value of $E\,Loss/N$ at the point which locates the local minimax as well as the value of $E\,Loss/N$ at the point where δ is equal to its upper bound and where we minimize $E\,Loss/N$ with respect to p. Then from these two points we choose a minimax solution. R. Maurice [7] indicates the procedure for doing this. The important point is that there is not an unrestricted minimax solution. A solution exists only if we have some prior information regarding δ, i.e. an upper bound for δ.

3.2 Maximin Method with Expected Net Gain

Differentiating $E\,Net\,Gain/N$ with respect to δ and setting the derivative equal to zero gives

$$1/2 = F(x) + xf(x). \tag{7}$$

It is easy to see that $x=0$ (i.e. $\delta=0$) is the unique solution of (7). (This result is obvious since, of course, the expected net gain is zero when $\delta=0$.)

From (3) we see that differentiating *E Net Gain/N* with respect to p is essentially the same as differentiating *E Loss/N* with respect to p, resulting in (6) when the derivative is set equal to zero. Using $x=0$ [the solution of (7)] the right-hand side of (6) is the indeterminate form 0/0. Applying L'Hopital's rule this expression approaches 2 as δ approaches zero. Hence the maximin solution for p is obtained by solving

$$(1 - 2p)/(2p) = 2,$$

giving $p=\frac{1}{6}$. A maximin is verified by considering the signs of second derivatives at $\delta=0$, $p=\frac{1}{6}$.

We note the maximin solution of net gain to be different from the local minimax solution of loss. The maximin solution with net gain is unrestricted and does not require any prior information regarding δ.

There is another point that is worth investigating. It is not difficult to show that the right-hand side of (6) increases as x increases (for positive x). Thus, the smallest value of the right-hand side of (6) is at $x=0$ which we have seen to be 2. This implies then that the solution of (6) for $x\neq 0$ results in a p that is less than one-sixth. In other words, an optimal expected loss or expected net gain procedure never requires more than one-sixth of the N patients on each treatment, or, always requires total trial participation by less than one-third of the N patients.

3.3 An A Priori Distribution for δ

Using (3) we see that integrating over a distribution for δ gives over-all expected net gain and loss that differ only by a constant.

Consider an *a priori* distribution for δ which is normal with zero mean and variance σ_0^2. Letting $\delta/\sigma_0=u$ and letting $\overline{E\ Net\ Gain}$ denote the result of integrating *E Net Gain* over the probability distribution for δ, we obtain

$$\overline{E\ Net\ Gain}/N = G\sigma_0(1 - 2p)\int_{-\infty}^{\infty} u[1 - 2F(-\sqrt{Rp}\,u)]f(u)du$$

$$= G\sigma_0(1 - 2p)\left[\int_{-\infty}^{\infty} uf(u)du - 2\int_{-\infty}^{\infty} uF(-\sqrt{Rp}\,u)f(u)du\right]$$

where

$$R = \frac{N\sigma_0^2}{2\sigma^2}.$$

The first term in brackets is Eu which is, of course, zero, and integration by parts in the second term gives

$$\overline{E\ Net\ Gain}/N = \frac{2G\sigma_0}{(2\pi)^{1/2}}(1 - 2p)\left(\frac{Rp}{1 + Rp}\right)^{1/2}. \qquad (8)$$

To determine the optimum p we differentiate (8) with respect to p and set the derivative equal to zero. This gives the quadratic

$$4Rp^2 + 6p - 1 = 0.$$

Solving for p gives

$$p^* = (-3 + \sqrt{9 + 4R})/(4R)$$

which is more conveniently written as

$$p^* = 1/(3 + \sqrt{9 + 4R}). \tag{9}$$

(The notation p^* is used to denote the optimal p). The sign of the second derivative with respect to p at $p=p^*$ is negative, verifying location of a maximum.

Examining (9) we see that for $R=0$, $p^*=\frac{1}{6}$, the maximum solution of the preceding section. (This is because $R=0$ means net gain is maximized at the point $\delta=0$, which is precisely what was done with the maximin method.) As R increases p^* decreases and approaches zero as R approaches infinity. Values of p^* for varying R are shown in Table 1.

Substituting the p^* of (9) into the expression for $\overline{E\ Net\ Gain}/N$ in (8) gives

$$[\overline{E\ Net\ Gain}/N]_2{}^* = \frac{2G\sigma_0}{(2\pi)^{1/2}}\left(\frac{3 + 2R - \sqrt{9 + 4R}}{9 + 2R + 3\sqrt{9 + 4R}}\right)^{1/2}. \tag{10}$$

Suppose now that one adhered to a rule of always using $p=\frac{1}{6}$, no matter what the circumstances. How inefficient is this rule compared to using the optimal p given by (9)? To answer this we first note that

$$[\overline{E\ Net\ Gain}/N]_{1/6} = \frac{2G\sigma_0}{(2\pi)^{1/2}}\,\frac{2}{3}\left(\frac{R}{R+6}\right)^{1/2}$$

so that

$$\begin{aligned}\text{Rel. Eff. of one-sixth} &= [\overline{E\ Net\ Gain}]_{1/6}/[\overline{E\ Net\ Gain}]_{p^*} \\ &= \frac{2}{3}\left(\frac{R}{R+6}\,\frac{9 + 2R + 3\sqrt{9 + 4R}}{3 + 2R - \sqrt{9 + 4R}}\right)^{1/2}.\end{aligned}$$

The relative efficiency of the one-sixth rule is tabulated in the last column of Table 1. Of course, at $R=0$ the relative efficiency is unity. As R increases the relative efficiency decreases and as R approaches infinity the relative efficiency approaches $\frac{2}{3}$. Thus, adhering to a rule of always using one-sixth, the net gain is never less than two-thirds of the gain using the optimal p^*. We recall that the preceding statement is still based on the *a priori* assumption that the true difference in mean effects is normally distributed with zero mean.

4. THE OPTIMAL SEQUENTIAL CLINICAL TRIAL

4.1 Minimax Method with Expected Loss

We now suppose that the trial no longer calls for a fixed number of participants, but the trial is performed sequentially on a pair of patients at a time,

TABLE 1. THE OPTIMAL FIXED SAMPLE SIZE CLINICAL TRIAL

R	p^*	Efficiency of maximum solution
0	.167	100.0
0.5	.158	99.9
1	.151	99.7
2	.140	99.0
4	.125	97.4
5	.119	96.6
10	.100	93.2
20	.080	88.7
50	.057	82.6
100	.043	78.7
∞	0	66.7

one member of the pair on Treatment A and the other on Treatment B. After the results from each pair are available a decision, based on the cumulative evidence at hand, is made to select one of the two treatments as the better and use it on all remaining patients or to continue the trial by having an additional pair participate.

Maurice [6] reviews the use of the sequential probability ratio test for the problem of ranking two normal populations with unknown means and common known variance and shows that this leads to a test based on the cumulative difference between pairs,

$$\sum_1^m d = \sum_1^m (x_A - x_B).$$

After the m-th pair one

$$\text{Selects Treatment } A \text{ if } \sum_1^m d > k\sigma^2$$

$$\text{Selects Treatment } B \text{ if } \sum_1^m d < -k\sigma^2.$$

Continues with another pair if

$$-k\sigma^2 \leq \sum_1^m d \leq k\sigma^2.$$

The boundaries consist of straight lines parallel to the x-axis. The problem is to determine k, the position of the boundary, so that expected loss is minimized.

Assuming N and $k\sigma$ to be large so that the formulae for unrestricted sequential sampling are reasonable approximations, the expected loss function becomes

$$E\ Loss = C\delta[En + (N - 2En)\ \Pr(\text{Select inferior})]$$

where En denotes the A.S.N. (average number of pairs) of the sequential trial.

It can be shown that application of the approximate formulae for unrestricted sequential sampling gives

$$En = k\sigma^2 \frac{e^{k\delta} - 1}{\delta(e^{k\delta} + 1)} \quad \Pr(\text{Select inferior}) = \frac{1}{e^{k\delta} + 1}. \tag{11}$$

Substituting into *E Loss* we obtain, after simplification,

$$E\,Loss/N = C\left[\frac{\delta}{e^{k\delta} + 1} + k\frac{\sigma^2}{N}\left(\frac{e^{k\delta} - 1}{e^{k\delta} + 1}\right)^2\right]. \tag{12}$$

Maurice [7] has applied the minimax method to this expression. Differentiating with respect to δ and k give

$$\frac{\partial}{\partial\delta}(E\,Loss/N) = C\left[\frac{e^{k\delta} + 1 - k\delta e^{k\delta}}{(e^{k\delta} + 1)^2} + 4\frac{\sigma^2}{N}\frac{k^2e^{k\delta}(e^{k\delta} - 1)}{(e^{k\delta} + 1)^3}\right] \tag{13}$$

$$\frac{\partial}{\partial k}(E\,Loss/N) = C\left[-\frac{\delta^2e^{k\delta}}{(e^{k\delta} + 1)^2} + \frac{\sigma^2}{N}\frac{(e^{k\delta} - 1)(e^{2k\delta} - 1 + 4k\delta e^{k\delta})}{(e^{k\delta} + 1)^3}\right]. \tag{14}$$

Setting these derivatives equal to zero and solving for k and δ, Maurice [7] has shown that

$$k = .8262\sqrt{N}/(\sigma\sqrt{2}) \qquad \delta = 2.668(\sigma\sqrt{2})/\sqrt{N}$$

is a unique solution and locates a minimax. Unlike the fixed sample size situation the minimax here is unrestricted. Substitution into (11) gives the expected sampling proportion at the minimax,

$$E\frac{n}{N} = .1241.$$

4.2 Maximin Method with Expected Net Gain

Use of a sequential plan does not alter the relation between expected net gain and expected loss given by (3). Thus, differentiating expected net gain with respect to δ gives

$$\frac{\partial}{\partial\delta}E\,Net\,Gain/N = G - 2\frac{G}{C}\frac{\partial}{\partial\delta}(E\,Loss/N)$$

$$= \frac{G}{(1 + e^{k\delta})^2}\left[e^{2k\delta} - 1 + 2k\delta e^{k\delta} - 8k^2\frac{\sigma^2}{N}\frac{e^{k\delta}(e^{k\delta} - 1)}{e^{k\delta} + 1}\right].$$

At $\delta = 0$ this derivative vanishes. It can be shown without difficulty that this is the only value of δ for which this is true. (Again, this result is obvious since the least value of *E Net Gain* is, of course, zero when $\delta = 0$.)

Differentiating *E Net Gain/N* with respect to k gives the same result as differentiating *E Loss/N* with respect to k in the preceding section, i.e., equation (14) (except for a multiplicative constant). Setting this derivative equal to zero gives

$$\frac{N}{\sigma^2} = \frac{(e^{k\delta} - 1)(e^{2k\delta} - 1 + 4ke^{k\delta})}{\delta^2 e^{k\delta}(1 + e^{k\delta})}.$$

When $\delta = 0$ the right-hand side is the indeterminate form 0/0. Applying L'Hopital's Rule twice, as δ approaches zero, the right-hand side approaches $3k^2$. Thus,

$$\frac{N}{\sigma^2} = 3k^2 \quad \text{or} \quad k = \sqrt{\frac{2}{3}}\,\frac{\sqrt{N}}{\sigma\sqrt{2}}$$

is the maximin solution. Examination of second derivatives verifies location of a maximin.

Substituting in *En* (equation (11)) and making use of L'Hopital's Rule, the expected sampling proportion at the maximin is $\frac{1}{6}$. We note that this $\frac{1}{6}$ is the maximin solution of the fixed sample size clinical trial.

4.3 An A Priori Normal Distribution for δ

As before, we assume an *a priori* distribution for δ which is normal with zero mean and variance σ_0^2. Again, it matters not whether loss or net gain is used. Integrating over the distribution for δ in *E Net Gain*, and substituting

$$\delta/\sigma_0 = x, \qquad k\sigma_0 = a, \qquad \frac{N\sigma_0^2}{2\sigma^2} = R$$

gives

$$\overline{E\ Net\ Gain}/N = G\sigma_0\left[\int_{-\infty}^{\infty} \frac{e^{ax} - 1}{e^{ax} + 1} xf(x)dx - \frac{a}{R}\int_{-\infty}^{\infty}\left(\frac{e^{ax} - 1}{e^{ax} + 1}\right)^2 f(x)dx\right].$$

The integrals are symmetric and noting that

$$\int_0^{\infty}\left(\frac{e^{ax} - 1}{e^{ax} + 1}\right)^2 f(x)dx = \frac{1}{2} - \frac{2}{a}\int_0^{\infty} \frac{e^{ax} - 1}{e^{ax} + 1} xf(x)dx$$

we obtain

$$\overline{E\ Net\ Gain}/N = 2G\sigma_0\left[\left(1 + \frac{2}{R}\right)\int_0^{\infty} \frac{e^{ax} - 1}{e^{ax} + 1} xf(x)dx - \frac{a}{2R}\right].$$

Integrating by parts gives

$$\overline{E\ Net\ Gain}/N = 2G\sigma_0\left[2\left(1 + \frac{2}{R}\right)\int_0^{\infty} \frac{ae^{ax}}{(1 + e^{ax})^2} f(x)dx - \frac{a}{2R}\right]. \tag{15}$$

We want to determine the a which maximizes (15). Differentiating with respect to a and setting the derivative equal to zero gives

$$\frac{1}{4} = (R + 2)\frac{d}{da}\int_0^{\infty} \frac{ae^{ax}}{(1 - e^{ax})^2} f(x)dx,$$

an equation involving a and R. Analytical solution of this equation for a in terms of R is not feasible. Solving for R in terms of a gives

$$R = \frac{1}{4\dfrac{d}{da}\displaystyle\int_0^\infty ae^{ax}(1 + e^{ax})^{-2}f(x)dx} - 2. \tag{16}$$

Thus for arbitrary choice of a, (16) gives the R such that the chosen a is optimal.

Using (16) we express (15) in terms of a only, giving

$$\overline{E\ Net\ Gain}/N = 4G\sigma_0 \frac{\displaystyle\int_0^\infty ae^{ax}(1 + e^{ax})^{-2}f(x)dx - a\frac{d}{da}\int_0^\infty ae^{ax}(1 + e^{ax})^{-2}f(x)dx}{1 - 8\dfrac{d}{da}\displaystyle\int_0^\infty ae^{ax}(1 + e^{ax})^{-2}f(x)dx}. \tag{17}$$

For arbitrary choice of a, (17) gives the value of $\overline{E\ Net\ Gain}/N$ when the chosen a is optimal.

Of course, to obtain numerical results the integrals must be evaluated. Formal integration does not appear feasible. What is done is to express $(1+e^{-ax})^{-2}$ in an infinite series involving terms e^{-jax} and to then integrate term by term. The results are

$$\int_0^\infty \frac{ae^{ax}}{(1 + e^{ax})^2}f(x)dx = \frac{1}{(2\pi)^{1/2}}\left\{\sum_{j=1}^\infty (-1)^{j+1}jaG(ja)\right\}$$

where

$$G(u) = F(-u)/f(u) \quad \text{(Mill's Ratio)},$$

and by straightforward differentiation,

$$a\frac{d}{da}\int_0^\infty \frac{ae^{ax}}{(1 + e^{ax})^2}f(x)dx = \frac{1}{(2\pi)^{1/2}}\sum_{j=1}^\infty (-1)^{j+1}ja[G(ja) + jaG'(ja)].$$

Substitution into (16) and (17) gives

$$R = \frac{\sqrt{2\pi}}{4}\frac{a}{\displaystyle\sum_{j=1}^\infty (-1)^{j+1}ja[G(ja) + jaG'(ja)]} - 2 \tag{18}$$

and

$$\overline{E\ Net\ Gain}/N = \frac{4G\sigma_0}{(2\pi)^{1/2}}a\frac{\displaystyle\sum_{j=1}^\infty (-1)^j(ja)^2G'(ja)}{a - \dfrac{8}{(2\pi)^{1/2}}\displaystyle\sum_{j=1}^\infty (-1)^{j+1}ja[G(ja) + jaG'(ja)]}. \tag{19}$$

TABLE 2. THE OPTIMAL SEQUENTIAL CLINICAL TRIAL

a^*	R	$\frac{\sqrt{2\pi}}{2G\sigma_0}(\overline{E \text{ Net Gain}}/N)$	Efficiency of maximin solution
0.5	.37	.201	100
1.0	1.47	.370	100
1.5	3.37	.501	100
2.0	6.26	.601	99
2.5	10.37	.678	98
3.0	15.96	.736	97
3.5	23.29	.781	96
4.0	32.63	.817	96
4.5	44.28	.845	95
5.0	58.49	.867	95

We can now consider numerical evaluation. The functions $G(u)$ and $G'(u)$ have been tabulated by Sheppard [8] for $u=0(.01)10$. For $u>10$ a series expansion can be used.

Arbitrarily, values of a of .5(.5)5.0 were selected. Using (18), R was determined such that the chosen a is optimal. The resulting values of R are shown in the second column of Table 2. The value of $\overline{E \text{ Net Gain}}/N$ when the chosen a is optimal is calculated from (19) and appears as the third column of Table 2.

Analogous to the fixed sample size situation we ask what would happen if we adhered to the rule of always employing the maximin solution, i.e., using

$$k = \sqrt{\frac{2}{3}}\,\frac{\sqrt{N}}{\sigma\sqrt{2}}$$

which in terms of a and R is

$$a = \sqrt{\frac{2}{3}}\,R.$$

The relative efficiency of the maximin solution is then

$$\text{Rel. efficiency of maximin} = [\overline{E \text{ Net Gain}}/N] - \sqrt{2R/3}/[\overline{E \text{ Net Gain}}/N]_a{}^*$$

where a^* denotes the optimal value of a for given R.

The final column of Table 2 gives the relative efficiency of the maximin solution. This last column should be considered a function of R (the second column of the table). The results indicate that little is lost by adhering to the rule of always using the maximin solution. The largest tabulated value of R, 58.5, shows the maximin solution has a net gain which is still 95 per cent of the optimal net gain.

5. COMPARISON OF OPTIMAL FIXED AND OPTIMAL SEQUENTIAL TRIAL

Finally, we compare the optimal fixed sample size clinical trial to the optimal sequential clinical trial. Intuitively, one expects better results with a sequential clinical trial than with a fixed, but how much better is the sequential?

Comparisons are made when there is an *a priori* distribution for δ which is normal with mean zero and variance σ_0^2. The $\overline{E\ Net\ Gain}/N$ for a fixed sample size trial is compared to $\overline{E\ Net\ Gain}/N$ of a sequential trial. Since $\overline{E\ Net\ Gain}/N$ depends on R then the comparison also depends on R.

Table 3 gives some numerical results of this comparison. The first two columns of Table 3 are a transcription from Table 2. They give the optimal value of the parameter a (denoted by a^*) of a sequential trial for given R. The third column of the table gives the optimal value of the parameter p (denoted by p^*) of a fixed sample size trial with the same R (using (9)). The over-all expected net gain of the optimal sequential trial is obtained from Table 2, and (10) gives the over-all expected net gain of the optimal fixed sample size trial for the same R. The last column of Table 3 gives the ratio of the over-all expected net gains.

The results show the optimal sequential trial has its greatest advantage over

TABLE 3. COMPARISON OF OPTIMAL FIXED AND OPTIMAL SEQUENTIAL TRIALS

R	a^*	p^*	Per Cent Additional Gain of Sequential Over Fixed
0		.167	25.3
0.37	0.5	.160	25.2
1.47	1.0	.146	24.3
3.37	1.5	.129	22.7
6.26	2.0	.113	20.7
10.37	2.5	.099	18.7
15.96	3.0	.087	16.9
23.29	3.5	.076	15.3
32.63	4.0	.068	13.8
44.28	4.5	.060	12.6
58.50	5.0	.054	11.5

the optimal fixed sample size trial at $R=0$, i.e. at the maximin point. Here the over-all expected net gain of the optimal sequential trial is 25.3 per cent more than that of the optimal fixed sample size clinical trial. As R increases the relative advantage of the sequential over the fixed sample size trial decreases. For R as high as 58.5 the optimal sequential trial has an over-all expected net gain which is only 11.5 per cent more than that of the optimal fixed sample size trial.

ACKNOWLEDGMENT

I would like to thank Professor P. Armitage for suggesting this problem and for his assistance and criticisms during the course of this research.

REFERENCES

[1] Armitage, P., *Sequential Medical Trials*. Oxford, England: Blackwell Scientific Publications, 1961.

[2] Bross, I., "Two choice selection," *Journal of the American Statistical Association*, 45 (1950), 530–40.

[3] Dunnett, C. W., "On selecting the largest of k normal population means," *Journal of the Royal Statistical Society, Series B*, 22 (1960), 1–40.
[4] Grundy, P. M., Healy, M. J. R., and Rees, D. H., "Economic choice of the amount of experimentation," *Journal of the Royal Statistical Society, Series B*, 18 (1956), 32–49.
[5] Grundy, P. M., Healy, M. J. R., and Rees, D. H., "Decision between two alternatives —how many experiments?" *Biometrics*, 10 (1954), 317–23.
[6] Maurice, R. J., "A minimax procedure for choosing between two populations using sequential sampling," *Journal of the Royal Statistical Society, Series B*, 19 (1957), 255–61.
[7] Maurice, R. J., "A different loss function for the choice between two populations," *Journal of the Royal Statistical Society, Series B*, 21 (1959), 203–13.
[8] Sheppard, W. F., *The Probability Integral*. British Association Mathematical Tables, Volume 7. Cambridge, England: Cambridge University Press, 1939.
[8] Somerville, P. N., "Some problems of optimum sampling," *Biometrika*, 41 (1954), 420–9.
[10] Wald, A., *Sequential Analysis*. New York: John Wiley and Sons, 1947.