Sankhyā : The Indian Journal of Statistics
2004, Volume 66, Part 2, pp 385-398

Optimal Scheduling of Intermediate Examination Times in Simple Illness-Death Model

Anup Dewanji and Atanu Biswas
Indian Statistical Institute, Kolkata, India

Abstract

Simple illness-death model arises in many examples of survival and reliability studies with industrial products. In a typical illness-death model, one is often interested in the event of occurrence of illness (D) which is assumed to be unobservable. This event is followed by the event of failure or death (F) which is observable in addition to the presence of illness. Failure (F) can also be observed before the illness occurs in which case the absence of disease is also recorded. In this work, we consider the problem of finding one or more intermediate examination times in order to make inference on D and/or also to safeguard the event of failure with illness before it is detected at an examination time. The standard likelihood based criteria are difficult to apply since calculation of the expected information matrix is not straightforward. We discuss two new optimality criteria which are based on some simple probability calculations and easy to apply.

AMS (2000) *subject classification.* 62K05, 62L05.
Keywords and phrases. Illness-death model, optimal designs, performance characteristics, cancer screening studies, risk, adaptive designs.

1 Introduction

Simple illness-death model arises in many examples of survival and reliability studies with industrial products. In a typical illness-death model, one is often interested in the event of occurrence of illness (D) which is assumed to be unobservable. This illness can be identified with a disease in the survival context and a fault in the reliability set-up. This event D is followed by the event of death or failure (F) which is observable in addition to the presence of disease/fault. Failure (F) can also be observed before the illness occurs in which case the absence of illness is also recorded. For notational convenience, let us also mean by D or F the corresponding time of event whenever the context is such. Applications can be found in animal carcinogenicity experiments, medical studies involving human subjects

(for example, in HIV/AIDS research), and in industrial applications with machine faults. See Dewanji and Biswas (2001) for more details.

Borgan et al. (1984) consider the comparison of several designs for an illness-death model with respect to their efficiencies relative to the design of continuous monitoring and conclude that the one with intermediate observation has highest efficiency. Since continuous monitoring is unrealistic in light of the cost and operational difficulties involved in such design, we consider here the design with K (≥ 1) intermediate examination times and address the problem of scheduling these times.

This design issue is important, for example, in cancer screening studies to schedule the visits of susceptible patients when they are to be examined for the presence of cancer (see Day and Walter, 1984). There are some ad-hoc approaches to schedule the visits, usually once in every year or so, although a more objective criterion to choose K time points for intermediate examination will be of interest (see Zelen, 1993). Similarly, in reliability context, it is important to schedule intermediate inspection(s) for detection of fault in a machine. It is desirable to guard against the possibility of both illness (disease or fault) and failure taking place before the illness can be detected in a subsequent examination time. This feature is particularly useful when some corrective measures can be started upon diagnosis/detection of illness. However, one may also be interested in the estimation of the distribution of D, the occurrence time of illness, in which case it is desirable to have as much information as possible on the distribution of D. This may be useful in some risk assessment issues concerning the occurrence of D. In this work, we consider finding a simple design for serving either or both of the purposes which is optimal in some sense.

The calculation of expected information matrix for a typical observation from the simple illness-death model seems to be difficult in general (see Dewanji and Biswas, 2001, and also the discussion in the following sections). Therefore, the task of obtaining an optimal design using the standard likelihood based criteria, which deal with the expected information matrix, is not straightforward. For this reason, we introduce some simple probability based optimality criteria which are very easy to deal with and have compelling intuitive appeal. Unlike Zelen's optimal scheduling, the approach is sequential in nature in the sense that the optimal examination times are chosen one at a time, the later choices depending on the earlier examination times and also the outcome of those examinations.

In Section 2, we introduce a model for the joint distribution of D and F, which is used for our description in the subsequent sections. Section 3 considers the design problem for choosing one examination time to intro-

duce the underlying principle. Section 4 generalizes the idea to choose more than one examination times. Section 5 discusses some Bayesian designs and Section 6 ends with some concluding remarks.

2 Preliminaries

We start with the simple parametric assumption of exponential model for D, with parameter α, having density $f(x) = \alpha e^{-\alpha x},\ \alpha > 0,\ x > 0$. The assumed conditional distribution of F, given $D = x$, is described in terms of its conditional density as follows:

$$g(y|x) = \begin{cases} \beta e^{-\beta y}, & \text{if } y < x \\ \gamma e^{-\beta x - \gamma(y-x)}, & \text{if } y \geq x. \end{cases} \tag{1}$$

That is, the conditional hazard β (before illness occurs) changes to γ once the illness occurs (see Freund, 1961). Although, in the following sections, we work with this model, the approach is simple and flexible enough for other general models, as noted in Section 5.

The optimal designs will depend on the model parameters which need to be estimated based on data from some other sources or from previous stages in a multistage framework. Estimation of the model parameters based on relevant data has been discussed elsewhere (see, for example, Dewanji and Biswas, 2001).

We now introduce four different quantities as performance characteristics associated with different purposes and study the performance of the proposed optimal designs of sections 3 and 4 with respect to these characteristics using model (1). See Biswas and Dewanji (2004) for details.

The first characteristic we consider is the average number of examinations (out of K), denoted by $E[N]$, where N is the number of examinations actually used. It may be noted here that, in the proposed designs, once the illness is detected at an examination time, the subsequent examinations are not necessary and hence not scheduled. Since each examination has a cost associated with it, the average number of examinations may seem to be a reasonable characteristic to consider. However, if cost is not much of concern, and noting that more number of examinations at optimally chosen time points lead to more information on the distribution of D and proper utilization of the K allowable examinations, larger values of $E[N]$ is desirable. For example, in the reliability context, routine periodic examinations for finding defects are usually not very expensive. For the model (1), $E[N]$

is found to be

$$E[N] = \sum_{j=1}^{K} \left\{ \frac{\alpha}{\alpha+\beta-\gamma} e^{-\gamma(t_j - t_{j-1})} + \frac{\beta-\gamma}{\alpha+\beta-\gamma} e^{-(\alpha+\beta)(t_j - t_{j-1})} \right\} e^{-(\alpha+\beta)t_{j-1}},$$

where $t_1 < \cdots < t_K$ denote the K optimal examination times.

In most practical situations, the aim is to reduce the time duration of illness remaining undetected so that, upon detection, a corrective step can be initiated. Let us, therefore, define 'loss' as the above time duration from the occurrence of illness till its detection at the subsequent examination time or failure, whichever is earlier. The risk (R), defined as the expected loss, is, for the model (1), given by

$$R = \sum_{j=1}^{K} (t_j A_{2j} - A_{1j} + A_{3j} - A_{4j}) + (A_5 - A_6),$$

where

$$\begin{aligned}
A_{1j} &= e^{-\gamma t_j} \left\{ \frac{\alpha}{\alpha+\beta-\gamma} \left[t_{j-1} e^{-(\alpha+\beta-\gamma)t_{j-1}} - t_j e^{-(\alpha+\beta-\gamma)t_j} \right] \right. \\
&\quad \left. + \frac{\alpha}{(\alpha+\beta-\gamma)^2} \left[e^{-(\alpha+\beta-\gamma)t_{j-1}} - e^{-(\alpha+\beta-\gamma)t_j} \right] \right\} \\
A_{2j} &= e^{-\gamma t_j} \frac{\alpha}{\alpha+\beta-\gamma} \left[e^{-(\alpha+\beta-\gamma)t_{j-1}} - e^{-(\alpha+\beta-\gamma)t_j} \right] \\
A_{3j} &= \frac{\alpha}{\alpha+\beta} \left[t_{j-1} e^{-(\alpha+\beta)t_{j-1}} - t_j e^{-(\alpha+\beta)t_j} \right] \\
&\quad + \left\{ \frac{\alpha}{(\alpha+\beta)^2} + \frac{\alpha}{\gamma(\alpha+\beta)} \right\} \left[e^{-(\alpha+\beta)t_{j-1}} - e^{-(\alpha+\beta)t_j} \right] \\
&\quad - (t_j + \gamma^{-1}) e^{-\gamma t_j} \frac{\alpha}{\alpha+\beta-\gamma} \left[e^{-(\alpha+\beta-\gamma)t_{j-1}} - e^{-(\alpha+\beta-\gamma)t_j} \right], \\
A_{4j} &= \frac{\alpha}{\alpha+\beta} \left[t_{j-1} e^{-(\alpha+\beta)t_{j-1}} - t_j e^{-(\alpha+\beta)t_j} \right] \\
&\quad + \frac{\alpha}{(\alpha+\beta)^2} \left[e^{-(\alpha+\beta)t_{j-1}} - e^{-(\alpha+\beta)t_j} \right] \\
&\quad - e^{-\gamma t_j} \frac{\alpha}{(\alpha+\beta-\gamma)^2} \left[e^{-(\alpha+\beta-\gamma)t_{j-1}} - e^{-(\alpha+\beta-\gamma)t_j} \right] \\
&\quad - e^{-\gamma t_j} \frac{\alpha}{\alpha+\beta-\gamma} \left[t_{j-1} e^{-(\alpha+\beta-\gamma)t_{j-1}} - t_j e^{-(\alpha+\beta-\gamma)t_j} \right],
\end{aligned}$$

and $A_5 = A_{3j}$ with $t_{j-1} = t_K$ and $t_j = \infty$, and $A_6 = A_{4j}$ with $t_{j-1} = t_K$ and $t_j = \infty$.

Another practical aim is to be able to detect illness before failure occurs with high probability. The probability of this event, as the third characteristic, is, for the model (1), given by

$$\begin{aligned} P &= \sum_{j=1}^{K} \left\{ \frac{\alpha}{\alpha+\beta} \left[e^{-(\alpha+\beta)t_{j-1}} - e^{-(\alpha+\beta)t_j} \right] \right. \\ &\quad \left. - \frac{\alpha e^{-\gamma t_j}}{\alpha+\beta-\gamma} \left[e^{-(\alpha+\beta-\gamma)t_{j-1}} - e^{-(\alpha+\beta-\gamma)t_j} \right] \right\}. \end{aligned}$$

In order to quantify the performance for estimating the distribution of D, without having to calculate the related information matrix, we introduce the following characteristic given by the sum of absolute differences between the K optimal examination times and the corresponding quantiles of the distribution of D. Formally, this characteristic Q is defined as

$$Q = \sum_{j=1}^{K} |t_j - q_j|,$$

where q_j, $j = 1, \cdots, K$, are the K quantiles of D given by $P[q_{j-1} < D < q_j] = \frac{1}{K+1}$, for $j = 1, \cdots, K$, which evenly cover its range. For the exponential distribution of D, given in (1), we have

$$Q = \sum_{j=1}^{K} \left| t_j + \frac{1}{\alpha} \log \left(1 - \frac{j}{K+1} \right) \right|.$$

Note that, these q_j's coincide with the d_j's defined later in Section 4 for exponential distribution of D.

3 One Examination Time

In this section, we discuss the design for choosing one examination time to introduce the idea, which then can be generalized, in the next section, to K examination times.

As D is the event of interest, we like to arrive at a design giving most 'information' on the event D so that the intermediate examination time, t_1 say, is not too early to miss most of the occurrences of illness. That is, D should occur before t_1 with highest probability. Without any further restriction, this will lead to the trivial optimal design of choosing $t_1 = \infty$. But we do not want t_1 to be too late to observe only the illnesses followed

by failures. Therefore, we consider maximizing the probability of the event $\{D < t_1 < F\}$ with respect to t_1. Intuitively, this also has the interpretation of trying to get t_1 as close to D as possible. It has also the natural appeal of incorporating the information on the associated failure process while focusing on the illness process. Let us denote this criterion by $\mathcal{C}_1$. In order to guard against the possibility of both illness and failure taking place before the examination time t_1, we may want to minimize the probability of the event $\{D < F < t_1\}$ as well. In order to achieve this, we consider maximizing the difference of the two probabilities, that is $P[D < t_1 < F] - P[D < F < t_1]$. Let us denote this criterion by $\mathcal{C}_2$.

Note that, for the model (1),

$$\begin{aligned} P[D < t_1 < F] &= \int_0^{t_1} \alpha e^{-(\alpha+\beta)x-\gamma(t_1-x)} dx \\ &= \frac{\alpha}{\alpha+\beta-\gamma}\left[e^{-\gamma t_1} - e^{-(\alpha+\beta)t_1}\right], \end{aligned}$$

assuming $\alpha + \beta \neq \gamma$. The value of t_1 which maximizes this probability is given by

$$t_1 = \frac{\log\left[(\alpha+\beta)/\gamma\right]}{\alpha+\beta-\gamma}. \tag{2}$$

For criterion $\mathcal{C}_2$, the difference $P[D < t_1 < F] - P[D < F < t_1]$ can be found as

$$\frac{2\alpha}{\alpha+\beta-\gamma}\left[e^{-\gamma t_1} - e^{-(\alpha+\beta)t_1}\right] - \frac{\alpha}{\alpha+\beta}\left[1 - e^{-(\alpha+\beta)t_1}\right],$$

assuming $\alpha + \beta \neq \gamma$. Maximizing this with respect to t_1, we get the optimal t_1 as

$$t_1 = \frac{1}{\alpha+\beta-\gamma}\log\left[\frac{(\alpha+\beta+\gamma)}{2\gamma}\right], \tag{3}$$

which can be seen to be less than the t_1 in (2). Note that both the optimal designs in (2) and (3) involve parameters related to both the illness and failure processes, as commented in the beginning of this section.

The optimal values of t_1 from (2) and (3), satisfying criteria $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively, have been reported in Table 4 of Dewanji and Biswas (2001) for different sets of parameter values. It is also demonstrated there that the asymptotic relative efficiency (ARE) of the two designs with respect to the traditional variance minimizing optimal design is very good. The ARE of $\mathcal{C}_1$ is almost 1, whereas that of $\mathcal{C}_2$ is above 90%.

The four performance characteristics of the previous section are calculated for the two designs and presented below in Table 1 along with the

optimal design points for some sets of parameters. The two entries in each cell correspond to the criteria $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively.

TABLE 1. OPTIMAL DESIGNS AND PERFORMANCE CHARACTERISTICS OF $\mathcal{C}_1$ AND $\mathcal{C}_2$ WITH $K = 1$.

Parameters				Performance characteristics			
α	β	γ	t_1	$E[N]$	R	P	Q
0.2	0.1	0.15	4.62	0.58	4.87	0.33	1.16
			2.70	0.74	6.05	0.30	0.76
0.1	0.1	0.15	5.75	0.53	5.33	0.21	1.18
			3.08	0.72	6.50	0.18	3.85
0.02	0.03	0.045	21.07	0.50	16.90	0.16	13.59
			10.81	0.71	20.37	0.13	23.84

As expected, $E[N]$ for $\mathcal{C}_1$ is lower than that for $\mathcal{C}_2$, since the value of t_1 for the latter is lower than that of the former. In terms of R and P, $\mathcal{C}_1$ performs better than $\mathcal{C}_2$. However, in terms of Q, the choice does not seem to be that clear.

4 Multiple Examination Times

As a natural generalization of the argument used in the previous section, for $\mathcal{C}_1$, D should occur by the first examination time, t_1; if not, then D should occur by the second examination time, t_2; and so on. One would, therefore, be tempted to choose the first examination time t_1 by maximizing $P[D < t_1 < F]$, and, if D does not occur by time t_1, then to choose the second examination time t_2 by maximizing $P[D < t_2 < F | t_1 < D, F]$, and so on. However, by this natural generalization, t_1 is same as the optimal time point when only one intermediate examination is allowed. This results in a large t_1 and the subsequent examinations, in many cases, may not be necessary at all. Therefore, instead of risking the information on possible early occurrence of D, it is reasonable to spread the K examination times evenly, in some sense, over the range of D, that is $(0, \infty)$. For this purpose, we choose K time points $d_1 < \cdots < d_K$, say, in the range of D. We first choose d_1 satisfying $P[D < d_1] = \frac{1}{K+1}$. Then, after finding t_{j-1}, the $(j-1)$th optimum examination time, d_j is chosen satisfying $P[t_{j-1} < D < d_j | D > t_{j-1}] = \frac{1}{K-j+2}$. That is, d_j is taken as the first of the $(K - j + 2)$ quantiles of the residual life time beyond t_{j-1} (with $t_0 = 0$). As a principle, we like the jth examination time not to exceed d_j, for $j = 1, \cdots, K - 1$.

Then, we may choose the optimal examination times as

$$\begin{aligned} t_j &= \min\{d_j,\ \arg\max_t P[D < t < F | t_{j-1} < D, F]\}, \quad 1 \le j \le K-1, \\ t_K &= \arg\max_t P[D < t < F | t_{K-1} < D, F]. \end{aligned}$$

Choosing t_K is equivalent to choosing one optimal inspection time from the residual life time after t_{K-1} conditional on $[t_{K-1} < D, F]$. This is criterion $\mathcal{C}_1$.

Similarly, the criterion $\mathcal{C}_2$ is generalized as follows.

$$\begin{aligned} t_j = \min\{d_j,\ \arg\max_t P[D < t < F | t_{j-1} < D, F] - P[D < F < t | t_{j-1} < D, F]\}, \\ 1 \le j \le K-1. \\ t_K = \arg\max_t P[D < t < F | t_{K-1} < D, F] - P[D < F < t | t_{K-1} < D, F]. \end{aligned}$$

Also, as in Zelen (1993), we can introduce a utility function given by

$$U = \sum_{j=1}^{K} \{P[t_{j-1} < D < t_j < F] - P[t_{j-1} < D < F < t_j]\}, \tag{4}$$

which is the sum of differences, at different examination times, between the probability of detecting illness D and the probability of failure with illness taking place before the examination. The optimal examination times can be obtained by maximizing U with respect to $t_1, \cdots, t_K$ simultaneously. Let us call this criterion $\mathcal{C}_3$. Note that, for $K = 1$, this is same as $\mathcal{C}_2$.

Simple probability calculation gives, for the model (1),

$$P[D < t < F | t_{j-1} < D, F] = \frac{\alpha}{\alpha + \beta - \gamma} \left[e^{-\gamma(t - t_{j-1})} - e^{-(\alpha+\beta)(t-t_{j-1})} \right].$$

Maximizing this with respect to t, it is easy to see that the optimal examination times, for the optimality criterion $\mathcal{C}_1$, are given by

$$\begin{aligned} t_j &= \min\left\{ d_j,\ t_{j-1} + \frac{\log\left(\frac{\alpha+\beta}{\gamma}\right)}{\alpha + \beta - \gamma} \right\}, \quad j = 1, \cdots, K-1, \\ \text{and} \quad t_K &= t_{K-1} + \frac{\log\left(\frac{\alpha+\beta}{\gamma}\right)}{\alpha + \beta - \gamma}. \end{aligned}$$

Similarly, for $\mathcal{C}_2$, one can calculate the difference $P[D < t < F | t_{j-1} < D, F] - P[D < F < t | t_{j-1} < D, F]$ as

$$\frac{2\alpha}{\alpha+\beta-\gamma}\left[e^{-\gamma(t-t_{j-1})}-e^{-(\alpha+\beta)(t-t_{j-1})}\right]-\frac{\alpha}{\alpha+\beta}\left[1-e^{-(\alpha+\beta)(t-t_{j-1})}\right].$$

Maximizing this with respect to t, the optimal examination times come out to be

$$t_j = \min\left\{d_j,\ t_{j-1}+\frac{1}{\alpha+\beta-\gamma}\log\left(\frac{\alpha+\beta+\gamma}{2\gamma}\right)\right\}, \quad j=1,\cdots,K-1,$$

$$\text{and}\quad t_K = t_{K-1}+\frac{1}{\alpha+\beta-\gamma}\log\left(\frac{\alpha+\beta+\gamma}{2\gamma}\right).$$

In order to use the criterion $\mathcal{C}_3$, the model (1) gives the expression for the utility function (4) as

$$U = \sum_{j=1}^{K}\left\{\frac{2\alpha e^{-\gamma t_j}}{\alpha+\beta-\gamma}\left[e^{-(\alpha+\beta-\gamma)t_{j-1}}-e^{-(\alpha+\beta-\gamma)t_j}\right]-\frac{\alpha}{\alpha+\beta}\left[e^{-(\alpha+\beta)t_{j-1}}-e^{-(\alpha+\beta)t_j}\right]\right\}.$$

Maximizing U with respect to the t_j's simultaneously, we obtain the following recursive relation for the optimal Δ_j's, where $\Delta_j = t_j - t_{j-1}$:

$$e^{-\gamma\Delta_{j+1}} = \frac{\alpha+\beta-\gamma e^{(\alpha+\beta-\gamma)\Delta_j}}{\alpha+\beta-\gamma}, \quad j=1,\cdots,K-1,$$

$$\Delta_K = \frac{1}{\alpha+\beta-\gamma}\log\left(\frac{\alpha+\beta+\gamma}{2\gamma}\right).$$

The optimal examination times $t_1,\cdots,t_K$, can now be obtained from the Δ_j's above. Unlike Zelen (1993), they are not equispaced.

Biswas and Dewanji (2004) present the K optimal examination time points satisfying criteria $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$, for different sets of parameter values, for $K = 3, 5$ and 10 (the last three entries in each cell of their Tables 1 and 2). These examination times are seen generally to be evenly located on either sides of $1/\alpha$, the expected value of D. In their Table 3, Biswas and Dewanji also present the values of the four performance characteristics for several design criteria. We present, in Table 2, the same for our three criteria. In order to save space, we denote the set of parameters (α, β, γ) by θ and the three parameter combinations of Table 1 above by θ_1, θ_2 and θ_3, respectively. The three entries in each cell correspond to the three criteria $\mathcal{C}_1$-$\mathcal{C}_3$, respectively.

TABLE 2. PERFORMANCE CHARACTERISTICS OF $\mathcal{C}_1$, $\mathcal{C}_2$ AND $\mathcal{C}_3$ WITH $K = 3$, 5 AND 10.

	$K=3$				$K=5$				$K=10$			
θ	$E[N]$	R	P	Q	$E[N]$	R	P	Q	$E[N]$	R	P	Q
	1.62	2.89	0.49	0.00	2.47	1.91	0.55	0.00	4.51	1.00	0.60	0.76
θ_1	1.64	3.23	0.48	0.76	2.48	2.14	0.54	0.76	4.52	1.12	0.60	0.76
	1.64	3.26	0.48	0.93	2.38	2.09	0.54	1.36	4.11	1.00	0.60	4.69
	1.23	2.47	0.34	1.18	1.92	1.51	0.39	1.18	3.62	0.71	0.44	1.18
θ_2	1.36	3.44	0.33	5.79	1.98	2.09	0.38	5.79	3.64	0.96	0.44	5.79
	1.67	4.07	0.32	10.31	2.45	2.80	0.38	15.31	4.23	1.44	0.43	24.29
	0.97	7.11	0.25	13.59	1.55	4.12	0.29	13.59	3.00	1.90	0.34	13.59
θ_3	1.37	11.65	0.25	53.48	1.71	6.62	0.29	54.85	3.04	2.71	0.34	54.85
	1.69	13.58	0.24	70.18	2.48	9.69	0.29	111.83	4.29	5.20	0.34	202.42

As expected, $E[N]$ for $\mathcal{C}_1$ is lower than that for $\mathcal{C}_2$, since the examination time points for the latter are lower than those of the former. This quantity is generally higher for $\mathcal{C}_3$. In terms of R, $\mathcal{C}_1$ performs by far the best with $\mathcal{C}_3$ being the worst. With respect to P, all the three seem to perform equally well. However, in terms of Q, $\mathcal{C}_1$ stands out ahead of others and $\mathcal{C}_3$ seems to perform badly. Therefore, $\mathcal{C}_1$ seems to be the best criterion, in general, for serving both the purposes.

Although the K optimal examination times $t_1, \cdots, t_K$ satisfying any of the criteria $\mathcal{C}_1$–$\mathcal{C}_3$ can be fixed before the actual study begins, for a particular patient, in practice, not all of them may be necessary. By construction, they are in increasing order to maintain the chronological timing. However, as the study progresses and accumulates information on the disease process up to, say, the $(j-1)$th examination time, use of this information for choosing the optimal jth examination time becomes important. One can improve upon the initial optimal choice of t_j, by using the information accumulated up to t_{j-1}, thus making the optimal choice of t_j adaptive. We propose to use this information to update or improve the estimate of $\theta = (\alpha, \beta, \gamma)$ using the current accumulated data. This can be done by using the EM algorithm (Dempster et al., 1977) as described in Dewanji and Biswas (2001). Then, the improved estimate can be used to choose the next optimal examination time. For $\mathcal{C}_3$, however, since the optimal choice of the t_j's is simultaneous, this adaptive method is purely algebraic lacking any criterion of optimality. Note that in each case, t_j has the expression as t_{j-1} *plus* a positive quantity, which is being evaluated at the current estimate. Thus, the adaptive optimal scheduling also leads to naturally (increasing) ordered examination times, at least for the model (1).

5 Bayesian Designs

The examination times obtained in the earlier sections are functions of the parameters. For example, for the model (1), the examination times are functions of $\theta = (\alpha, \beta, \gamma)$, the model parameters, which are typically unknown. Therefore, to implement the design, we need some idea (or estimates) of these parameters. A reasonable approach is to employ the adaptive method as described in the last paragraph of the previous section. Another alternative is to use some Bayesian design where a prior $\pi(\theta)$ for θ is chosen according to the belief and knowledge of the experimenter. Let Θ be the domain of θ. We can go on updating our knowledge about these parameters with the accumulated data (posterior distribution of θ) at any point of time. For the sake of simplicity and administrative convenience, suppose we find the posterior of θ only at the successive examination times.

The Bayesian approach can be used in two ways. First, the successive examination times t_j's can be the expected value of the same, as derived in Section 4 using $\mathcal{C}_1 - \mathcal{C}_3$, with respect to the posterior obtained at the current time (see Berger, 1985, Section 4.3.4). In order to obtain the first examination time t_1, we find the corresponding expected value with respect to $\pi(\theta)$. For the optimality criterion $\mathcal{C}_1$ and model (1), for example, this time is

$$t_1 = \int_\Theta \min \left\{ d_1(\alpha), \frac{\log\left(\frac{\alpha+\beta}{\gamma}\right)}{\alpha+\beta-\gamma} \right\} \pi(\theta) d\theta,$$

where $d_1 = d_1(\alpha)$ is such that $P[D < d_1] = \frac{1}{K+1}$, the dependence on α made explicit. The above integration can be numerically carried out.

Let the likelihood of the accumulated data E_j till the examination time t_j (for $j = 1, \cdots, K-1$) from all the subjects be denoted by $L_j(\theta)$. Note that E_j consists of all information on D and/or F till time t_j from all the subjects (see Dewanji and Biswas, 2001). Then, the posterior density of θ at time t_j is given by

$$\pi_j(\theta|E_j) = \frac{L_j(\theta)\pi_{j-1}(\theta|E_{j-1})}{\int_\Theta L_j(\theta)\pi_{j-1}(\theta|E_{j-1})d\theta},$$

with E_0 being empty and $\pi_0(\cdot) = \pi(\cdot)$. Then, the $(j+1)$st examination time (for $j = 1, \cdots, K-2$), from Section 4 for $\mathcal{C}_1$, and model (1), is

$$t_{j+1} = \int_\Theta \min \left\{ d_{j+1}(\alpha), t_j + \frac{\log\left(\frac{\alpha+\beta}{\gamma}\right)}{\alpha+\beta-\gamma} \right\} \pi_j(\theta|E_j) d\theta,$$

where $d_{j+1} = d_{j+1}(\alpha)$ depends on α.

Although this approach seems straightforward, the numerical derivation of examination times becomes a formidable task as we deviate from model (1). As commented in the last section, besides the expectation with respect to the current posterior, the computation of t_j for given θ itself will require numerical integration and maximization. An alternative approach is to take the expectation of the objective function with respect to the current posterior before optimizing it. For example, for $\mathcal{C}_1$ and model (1), the first examination time t_1 is chosen as

$$t_1 = \min\left\{\int_0^\infty d_1(\alpha)\pi(\alpha)d\alpha, \arg\ \max_t \int_\Theta P[D < t < F]\pi(\theta)d\theta\right\}.$$

The subsequent examination times can be written in similar fashion using the current posterior. However, as we deviate from model (1), the numerical task again becomes formidable. One can also think of employing Monte Carlo integration to calculate the successive examination times as they have expectation-like expression. Besides the Bayesian approach, there are few other strategies to deal with the problem of parameter-dependent design (see Pukelsheim, 1993, Ch. 11) based on mixture of competing models, constrained optimization, etc.

6 Concluding Remarks

In order to arrive at an optimal choice of the examination times (the t_j's), at least in non-Bayesian framework, the following strategy may be implemented. We first make a guess for their values from past experience or prior knowledge, if any. Start the study with n individuals with these examination times. Based on the n observations, estimate the model parameters and obtain the optimal choice of t_j's using these estimates. This idea can be extended to suggest strategies involving multiple stages.

Note that the design by $\mathcal{C}_1 - \mathcal{C}_3$ can be obtained in simple closed form at least for model (1), whereas the ones using expected information are difficult. The efficiency of $\mathcal{C}_1$ or $\mathcal{C}_2$ also seem to be quite high (at least for $K = 1$). These criteria, therefore, seem like a useful and appealing alternative to the traditional ones, specially when the optimal design is to be obtained by strategies involving multiple stages, as commented above. Since the criteria $\mathcal{C}_1$ and $\mathcal{C}_2$ optimize some probability terms only and do not depend on the likelihood, the corresponding designs do not change due to minor changes in the secondary aspects of data. For example, when there is possibility of censoring or some missing mechanism inherent in the

data leading to different forms of likelihood, criteria $\mathcal{C}_1$ and $\mathcal{C}_2$ give the same optimal designs given in sections 3 and 4; whereas, for finding an optimal design by the likelihood based criteria, the calculation of expected information matrix requires knowledge on the censoring distribution or the missing mechanism which is usually unknown. This is the major advantage of our criteria.

Another advantage of the probability based criteria $\mathcal{C}_1$ and $\mathcal{C}_2$ is easy incorporation of one or more covariates, denoted by $Z = z$, say, in the optimal design. For example, if the distribution of D happens to depend on $Z = z$ via the exponential parameter $\alpha = \alpha(z) = \alpha_0 e^{\alpha_1 z}$ (say), the optimal design for the t_j's by $\mathcal{C}_1$ or $\mathcal{C}_2$, can be readily obtained simply by replacing α by $\alpha(z)$. Therefore, the optimal design for individuals with different Z values will be different (making it more realistic) but can be obtained by using one formula. In order to achieve this by a criterion based on expected information, one has to do the extensive computation again and again for different values of Z.

Although we demonstrated the results based on the simple model (1) of Section 2, the criteria $\mathcal{C}_1$ and $\mathcal{C}_2$ can be easily employed for more general models. The calculation of the probabilities $P[D < t < F | t' < D, F]$ and $P[D < F < t | t' < D, F]$, in general, involves numerical integration, and so, the maximization required for $\mathcal{C}_1$ and $\mathcal{C}_2$ may not have closed form solutions as in sections 3 and 4. However, numerically finding the optimal t_j's requires much less computation than having to calculate the expected information matrix. Since $\mathcal{C}_3$ involves simultaneous maximization with respect to K variables, this may be more computer intensive. See Biswas and Dewanji (2003) for details.

Note that the optimal choice of examination times implicitly depends on the knowledge that a maximum of K examinations are allowed. However, if this knowledge is not available or there is no such restriction on the number of examinations (which is the likely scenario in many cases), all the criteria of this paper fail. One possible solution is to start with a K, and as the examinations are held, vary the value of K, depending on the current budget and information, while choosing the successive d_j's. This also makes the choice of t_j's adaptive in some sense. Since $\mathcal{C}_3$ makes simultaneous choice of the t_j's, the above remedy does not apply to it.

Acknowledgment. The authors wish to thank the referee and the Editor for constructive suggestions which led to improvement over an earlier version of the paper.

References

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, New York.

Biswas A, Dewanji A. (2004). On optimal examination times in simple illness-death model. *J. Statist. Plann. Inference*, to appear.

Borgan, Ø., Liestøl, K. and Ebbesen, P. (1984). Efficiencies of experimental designs for an illness-death model. *Biometrics* **40**, 627-638.

Day, N.E. and Walter, S.D. (1984). Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. *Biometrics* **40**, 1-13.

Dewanji, A. and Biswas, A. (2001). An optimal design for simple illness-death model. *J. Statist. Plann. Inference*, **96**, 289-300.

Freund, J.E. (1961). A bivariate extension of the exponential distribution. *J. Amer. Statist. Assoc.,* **56**, 971-977.

Pukelsheim, F. (1993). *Optimal Design of Experiments.* Wiley, New York.

Zelen, M. (1993). Optimal scheduling of examinations for the early detection of disease. *Biometrika,* **80**, 279-293.

Anup Dewanji and Atanu Biswas
Applied Statistics Unit
Indian Statistical Institute
203, B.T. Road
Kolkata 700 108, INDIA
E-mail: dewanjia@isical.ac.in
atanu@isical.ac.in

Paper received: May 2002; revised March 2004.