TEST SELECTION FOR A MASS SCREEEENING PROGRAM*

John A. Voelker

Argonne National Laboratory
Argonne, Illinois

William P. Pierskalla

University of Pennsylvania
Philadelphia, Pennsylvania

ABSTRACT

Periodic mass screening is the scheduled application of a test to all members of a population to provide early detection of a randomly occurring defect or disease. This paper considers periodic mass screening with particular reference to the imperfect capacity of the test to detect an existing defect and the associated problem of selecting the kind of test to use. Alternative kinds of tests differ with respect to their reliability characteristics and their cost per application.

Two kinds of imperfect test reliability are considered. In the first case, the probability that the test will detect an existing defect is constant over all values of elapsed time since the incidence of the defect. In the second case, the test will detect the defect if, and only if, the lapsed time since incidence exceeds a critical threshold T which characterizes the test.

The cost of delayed detection is an arbitrary increasing function (the "disutility function") of the duration of the delay. Expressions for the long-run expected disutility per unit time are derived for the above two cases along with results concerning the best choice of type of test (where the decision rules make reference to characteristics of the disutility function).

INTRODUCTION

Mass screening is the process of inspecting all members of a large population for defects. If the early detection of a defect provides benefits, it may be advantageous to employ a test capable of revealing the defect's existence in its earlier stages. (Throughout this paper, the words "defect" and "unit" or "individual" will refer to defect, disorder, or disease and to a member of the population, respectively).

Defects may arrive in a seemingly random fashion such as many types of machine failure, the incidence of certain types of cancer, diabetes, glaucoma, heart disease, etc.; or they may arrive as the result of some contagion such as smallpox, polio, etc. It is the former type of arrival process, random and independent arrivals, which is studied in this paper.

*This research was supported by the Office of Naval Research under Contracts N00014-67-0356-0030 and N00014-75-C-0797.

Continuous monitoring would provide the most immediate detection of a defect, but considerations of expense and practicality will frequently rule out continuous monitoring so that a schedule of periodic testing—a periodic screening program—may be the most practical means of achieving early detection of the defect. In general terms, the question then becomes one of how best to trade off the expense of testing which increases both with the frequency of test applications and with the cost of the type of test used against the benefits to be achieved from detecting the defect in an earlier state of development.

The benefits of early detection depend upon the application considered. For example, in a human population being screened for some chronic disease, the benefits of early detection might include an improved probability of ultimate cure, diminished time period of disability, discomfort, and loss of earnings; and reduced treatment costs. If the population being screened consists of machines engaged in some kind of production, the benefits of early detection might include a less costly ultimate repair and a reduction in the time period during which a faulty product is being unknowingly produced. If the population being screened consists of machines held in readiness to meet some emergency situation, an early detection of a defect would reduce the time the machine was not serving its protective function.

The expense of testing includes easily quantifiable economic costs such as those of the labor and materials needed to administer the testing. However, there can also be other important cost components which are more difficult to quantify. For example, in the case of a human population subject to medical screening, the cost of testing includes the inconvenience and possible discomfort necessitated by the test, the cost of false positives which entails both emotional distress and the need to do unnecessary follow-up testing, and even the risk of physical harm to the testee; e.g., from the cumulative effect of x-ray exposure.

The design of a mass screening program must address two important questions: How frequently to test and what kind of test to use. Optimal testing frequency has been investigated as a function of the defect incidence rate and other factors by Derman [3], Roeloffs [8,9], Barlow, et al. [1], Keller [4], Kirch and Klein [5], and Lincoln and Weiss [6].

The second question follows from the fact that more than one kind of test may be available for use in a mass screening situation. The alternate tests can be entirely different procedures; or they can be the same procedure with different criteria for what constitutes a positive outcome, e.g., alternate levels at which a recording of systolic blood pressure would induce follow-up testing.

Alternative tests will generally differ both in their reliability characteristics and in their cost of application. How to select which test to use is a question which, to our knowledge, has not previously been examined in the context of a general model. This paper will examine this question. For two different ways of modeling test reliability, we develop a framework for test selection and present certain specific decision rules.

Test reliability is assumed to be a function, $p(t)$, of elapsed time, t, since defect incidence. Define $\delta_S(t) = \begin{cases} 1 \text{ if } t \in S \\ 0 \text{ otherwise} \end{cases}$. Usually S will be an interval, e.g., $[T, \infty)$. The two cases $p(t) = p$ and $p(t) = \delta_{[T,\infty)}(t)$ are considered. $p(t) = p$ indicates that test reliability is independent of defect age. For $p(t) = \delta_{[T,\infty)}(t)$, the test will detect a defect if, and only if, the defect has existed for at least T units of time. In a sense, these two classes of $p(t)$ represent polar extremes in the responsiveness of test reliability to defect maturity. The test choice decision is posed within each of these two reliability classes.

A crucial feature in any optimization model of mass screening is the characterization of the cost due to detection delay. Detection delay is the gap between the time of detection and the time of defect incidence (or the time the defect becomes potentially detectable by a screening test). The mapping between detection delay and the resultant cost we call the disutility function, $D(\cdot)$. Obviously, the shape of the disutility depends upon the particular application considered. Section 1 gives examples.

Pierskalla and Voelker [7] demonstrated that the shape of $D(\cdot)$ impacts the optimal allocation of a screening budget among segments of the client population characterized by differing defect incidence rates. Results in Sections 3 and 4 below demonstrate the role $D(\cdot)$ plays in test choice decision rules.

Previous research in the area of optimal mass screening which utilized closed form expressions for expected cost (or disutility) placed restrictive assumptions upon the shape of $D(\cdot)$. Early work assumed $D(t) = ct$ (Barlow, et al. [1]). In Kirch and Klein [5], $D(t) = \min(t, T)$ where t is detection delay and T is the (possibly random) delay between defect incidence and the time when the defect would be discovered in the absence of a screening program. Keller [4] restricts the generality of $D(\cdot)$ by requiring that $D(\cdot)$ and the test frequency (density) $r(t)$ be such that $r(t)\int_0^{1/r(t)} D(s)\,ds$ be well approximated by $D(1/2r(t))$. (A density is employed to represent the schedule of test times so that the calculus of variations could serve as the optimization tool.)

Lincoln and Weiss [6] derive two kinds of optimal testing schedules. Both schedules maximize the time between successive tests subject to, respectively, a bound on the mean detection delay and a bound on the probability that detection delay will exceed a fixed threshold. Neither version is equivalent to using a general disutility function.

A few authors, Schwartz and Galliher [10], Thompson and Disney [11], and Voelker [12] let both the reliability of the test and the disutility (or utility) of detection be a function of the defect's state rather than of time since the defect's incidence. Although such models are more general and do utilize a general concept of disutility, they have not been amenable to closed form evaluation of expected disutility.

To incorporate random defect arrivals into their models, previous researchers focus upon an individual who will incur the defect. They use the density function for the age when that individual incurs the defect as a fundamental element of their model. Since the density function reflects age-specific incidence rates, a "life time" testing schedule can, thereby, be developed to tailor testing frequency at each age to the probability that the defect will occur at that age.

Our way of modeling the randomness of defect arrivals reflects a somewhat different perspective on the mass screening problem. We look through the eyes of a decision maker charged with intelligently allocating a fixed budget. The time frame over which the allocation must be made is often short compared to a typical life time of a member of the client population. Therefore, the decision maker does not plan lifetime screening schedules for particular individuals. Instead, he tries to maximize the benefit that can be derived from his available budget over a much shorter planning horizon. (For modeling purposes the objective of minimizing expected long run cost per unit time is not unreasonable with the problem viewed in this way since steady state conditions should approximately obtain after the initial screening. This is especially the case when an existing and ongoing screening program is being optimized by the decision maker. Also, lacking information to the contrary, the decision maker has no reason to anticipate abrupt changes in the screening policy at the end of the planning horizon.)

With the problem viewed in this perspective, the random nature of defect arrivals is most naturally modeled as a Poisson process with its parameter determined by the incidence rate of the defect and the size of the population. This approach has proved particularly useful in the following context: If different segments of the client population exhibit different incidence rates, subpopulations can be defined with defect incidence within each modeled as a Poisson process with its respective parameter. Then the budget can be so allocated among the subpopulation to permit appropriate relative testing frequencies (cf. Pierskalla and Voelker [7]). In this way, age-specific incidence rates can be incorporated into the notion of Poisson defect arrivals. Moreover, factors other than age which affect defect incidence rates (family history, smoking habits, work environment, etc.) can also be incorporated into the model.

Although this paper does not follow Pierskalla and Voelker [7] in studying the case of heterogeneous populations, we use the same Poisson model of random defect arrivals. Section 1 presents examples of disutility functions. Section 2 sketches the basic structure of the model and represents the expected long run disutility per unit time for general reliability and disutility functions. Section 3 considers the class of tests with $p(t) = p$. Section 4 does so for $p(t) = \delta_{[T,\infty)}(t)$. In both Sections 3 and 4, results regarding and test choice criteria are presented. Proofs are deferred to the APPENDIX.

SECTION 1: SOME EXAMPLES OF DISUTILITY FUNCTIONS

Suppose a production process is subject to a randomly occurring defect. Although production appears to proceed normally after the incidence of the defect, the product produced is, thereafter, defective to an extent which remains constant until the production process is returned to its proper mode of operation. The only way to learn if the production process is in this degraded state is to perform a costly test. Now, if a test detects the existence of the degraded mode of production t units of time after its incidence, the harm done will be proportional to the amount of defective product (unknowingly) produced which, in turn, is proportional to t. Hence, $D(t) = at$ for some $a > 0$.

Another example where a linear $D(\cdot)$ function may be appropriate would be for the periodic inspection of an inactive device (such as a missile) stored for possible use in an emergency. If t is the time between the incidence of the disorder and its detection, the disutility incurred is proportional to the probability that the device would be needed in that time interval. If such "emergencies" arise according to a Poisson process with rate μ, then the probability of an emergency in a time interval of length t is $1 - e^{-\mu t}$, which, for μ small, is approximately μt. Hence, if b is the cost incurred should there be an emergency while the device is defective, and if μ is the (small) arrival rate of emergencies, then $D(t) = b\mu t$.

A quadratic disutility could arise in the following situation. Suppose the magnitude of a randomly occurring defect increases linearly with time since the occurrence of the defect. For example, the magnitude of the defect might be the size of a small leak in a storage container for a fluid, and as fluid escapes, the leak gets larger. Further suppose that the harm done accumulates at a rate proportional to the magnitude of the defect. Hence, the quantity of fluid lost (at least initially) increases the longer the defect exists, and the rate of fluid loss is proportional to the size of the leak.

Let the size of the leak (as measured by rate of fluid loss), at time s since the leak's incidence, be cs. Then, if the defect is detected at time t since incidence, the disutility incurred (fluid lost) is $D(t) = \int_0^t \text{csds} = 1/2(ct)^2$.

SECTION 2: EXPECTED DISUTILITY WITH A GENERAL RELIABILITY FUNCTION

We assume that the times of defect arrivals in the screened population form a stationary Poisson process. Since there is a certain intuitive appeal to considering the defect arrival rate proportional to both the size of the population, N, and the intrinsic incidence rate, λ; let $N\lambda$ designate the parameter of the above arrival process. It is not necessary to know the value of $N\lambda$ in order to apply the decision rules of test selection developed in this paper.

Let $p_l(t)$ be the probability that a test of type l will detect a defect which has been present t units of time. $p_l(t) = 0$ for $t < 0$. Let $\bar{S}^k_{r,l}$ be a random variable denoting the time at which the k^{th} defect is detected. $\bar{S}^k_{r,l}$ depends upon the arrival time of the defect, δ_k; the type of test used, l; and the testing frequency, r.

Given the application of test type l at the times $\{1/r, 2/r, \ldots\}$, the disutility incurred by the k^{th} defect is $D(\bar{S}^k_{r,l} - S^k)$. The total disutility incurred due to defects which occur before time A is

$$\sum_{k=1}^{\infty} D(\bar{S}^k_{r,l} - S^k)\ \delta_{[O,A)}(S^k).$$

In Pierskalla and Voelker [7], the long run expected disutility per unit time under the above screening program $r,p(\ ))$ was shown to be

$$rN\lambda \sum_{n=1}^{\infty} \int_{n-1/r}^{n/r} D(u)\ p_l(u) \prod_{m-1}^{n-1} [1 - p_l(u - m/r)]\ du. \tag{2.1}$$

This result will serve as our starting piont for the technical results of this paper. Lincoln and Weiss established essentially the same result based upon a different probabilistic model of defect arrival.

SECTION 3: CONSTANT TEST RELIABILITY

The mass screening model yields interesting results for a test which has a fixed probability p of detecting the disorder if it is present in an individual. Such a model would arise if the unreliability of the test is entirely intrinsic to the test procedure rather than partially dependent upon the state or age of the defect. An example of this is the administration of a Mantoux test for tuberculosis in, say, a population of grade school children. The test has a small but relatively constant level of false negatives. There are other medical tests with similar characteristics.

A Quality Control Example

To see how another type of situation with constant test reliability could arise, consider a production system which is subject to a randomly occurring defect which degrades the system's performance. Once the defect occurs, the level of degradation of the process remains constant until the defect is discovered. Suppose the defect is such that each item produced has probability δ of being defective and that the system without the defect never produces defective items.

The only way to discover the existence of the defect in the production system is to examine an item produced which is itself defective. Now, if the examination of an item is expensive (e.g., the item is destroyed as a result of the inspection) and if the capacity to examine a sequence of items involves a set-up cost (say a), the following strategy might be called for: At specified times $1/r$, $2/r$, ..., set up the capacity to examine a sequence of items and examine, say, l items at each of those times. The times $1/r$, $2/r$, ..., are then the times of testing and sample size l specifies the test type.

Assume that if a defective item is examined, the defect is always observed and the production process is, thereby, discovered to be in the degraded state. Hence, a degraded state of the production process will go undetected at the testing occasion k/r if, and only if, each of the l items sampled at time k/r is, by chance, not defective. But the probability of that event is $(1-\delta)^l$. Note that the elapsed time t between the entry of the production process into the degraded state and the test time k/r does not affect this probability. Hence, $p_l = p_l(t) = 1 - (1-\delta)^l$ which represents the probability that a test (the inspection of l items) will detect a degraded production process. Note that the choice of l affects both the test's reliability and cost.

Expected Disutility

Designate the expected long run disutility per unit time per member of a population screened with frequency r using a test of constant reliability ($p(t) = p$, $t > 0$) by $C(r,p)$. From Eq. (2.1),

$$C(r,p) = N\lambda rp \sum_{n=0}^{\infty} (1-p)^n \int_{n/r}^{(n+1)/r} D(u)\,du.$$

The question examined next is how do changes in r and p affect $C(r,p)$. After that, explicit solutions are given for $C(r,p)$ when $D(\cdot)$ takes certain simple forms. And lastly, some general rules are indicated for selecting between a particular kind of test and a more expensive but more reliable alternative test when $D(\cdot)$ takes certain forms.

PROPOSITION 1: If $D(\cdot)$ is a strictly increasing function, $[\partial C(r,p)]/\partial p < 0$ and $[\partial C(r,p)]/\partial r < 0$.

Note that for $D(\cdot)$ nondecreasing, the above inequalities still hold, but not strictly.

From this proposition, as anticipated, when p increases, the expected disutility decreases. Similarly as r increases, the interval $1/r$ between tests decreases and the expected disutility decreases. Consequently, as better test types are used or the tests are more frequently applied, the value of such changes in terms of reduced disutility versus the costs of the changes can, in principle, be assessed and the tradeoffs evaluated.

It is easy to compute the Hessian for $C(r,p)$ when $D(\cdot)$ is differentiable:

$$\frac{\partial^2 C}{\partial r^2} = N\lambda p \sum_{n=0}^{\infty} q^n r^{-3}\left[(n+1)^2 D'\left(\frac{(n+1)}{r}\right) - n^2 D'\left(\frac{n}{r}\right)\right],$$

$$\frac{\partial^2 C}{\partial r^2} = N\lambda r \sum_{n=2}^{\infty} n(n-1)\, q^{n-2}\, [W(n,r) - W(n-1,r)],$$

$$\frac{\partial^2 C}{\partial p\,\partial r} = N\lambda \sum_{n=0}^{\infty} [1-(n+1)p]\, q^{n-1}\, [W(n,r) - B(n+1,r) + B(n,r)],$$

where

$$q = 1 - p, \; W(n,r) = \int_{n/r}^{(n+1)/r} D(s)\,ds, \text{ and } B(n,r) = \frac{n}{r} D(n/r).$$

Note that if $D'(\cdot)$ is increasing, then $(\partial^2 C/\partial r^2) \geqslant 0$ and $(\partial^2 C/\partial p^2) \geqslant 0$. Hence, along coordinate directions $C(r,p)$ is convex.

Simple expressions for $C(r,p)$ can be given when $D(\cdot)$ is specialized to a polynomial or an exponential function. Since these two types of functions are reasonably general, they can be quite useful as realistic approximations in applications.

PROPOSITION 2: If $D(t) = \sum_{i=1}^{I} a_i t^{m_i}$, then

$$C(r,p) = N\lambda p \sum_{i=1}^{I} [a_i/r^{m_i}(m_i + 1)] \sum_{n=1}^{\infty} n^{m_i+1} pq^{n-1}. \tag{3.1}$$

If the m_i; $i = 1, \ldots, N$ are positive integers, the inner summation of Eq. (3.1) is simply the $(m_i + 1)$ moment of a geometric random variable. Hence, using Laplace transforms, Eq. (3.1) becomes

$$C(r,p) = N\lambda p \sum_{i=1}^{I} [a_i/r^{m_i}(m_i + 1)] \, \psi^{m_i+1} (0{:}p) \tag{3.2}$$

where $\psi(t{:}p) = pe^t/(1 - qe^t)$ and $\psi^{(m)}(0{:}p) = d^m \psi(t{:}p)/dt^m \,|_{t=0}$. For example, when $m = 1$, $C(r,p) = (aN\lambda/2r)[(2 - p)/p]$ and for $m = 2$, $C(r,p) = [(aN\lambda/3r^2)]\,[1 + (6 - 6p)/p^2]$.

PROPOSITION 3: If $D(t) = \beta e^{at}$ for $a, \beta > 0$, then $C(r,p) = \beta N\lambda pr(e^{a/r} - 1)/a(1 - qe^{a/r})$, for $r > -a/\log(q)$.

Test Selection

Propositions 2 and 3 can provide a means to select between two alternative kinds of tests which differ with respect to reliability of detection and cost per application. Let test No. 1 have cost per application c_1 and reliability p_1. The corresponding parameters for test No. 2 are c_2 and p_2. If test No. 1 is administered with frequency rc_2/c_1 and test No. 2 administered with frequency r, both testing regimes will consume equal quantities of the budgeted resource; *vis.*, Nrc_2 per unit time. If $C(rc_2/c_1,p_1) \leqslant C(r,p_2)$ for all $r \geqslant 0$, then the expected disutility per unit time will be less with test No. 1 at all levels of budget Nrc_2. That is, if test No. 2 is being used with frequency r, the expected disutility can be decreased without any additional allotment of budget, simply by switching to test No. 1 and testing as frequently as the budget permits.

Suppose, for example, that $D(t) = at^m$ for m a positive integer. Then $C(rc_2/c_1,p_1) \leqslant C(r,p_2)$ is equivalent by Eq. (3.2) to

$$(c_1/c_2)^m \leqslant \frac{p_2\, \psi^{(m+1)}(0{:}p_2)}{p_1\, \psi^{(m+1)}(0{:}p_1)}. \tag{3.3}$$

Therefore, test No. 1 is preferred over test No. 2 if, and only if, Eq. (3.3) obtains.

SECTION 4: THRESHOLD TEST RELIABILITY

In the previous section, the reliability of the test depended only on factors intrinsic to the test itself and did not depend at all on the elapsed time since incidence at the time of the test. In this section, a special form for $p(t)$ is considered which is very different from the case of constant test reliability. Here the test reliability is zero if the elapsed time since the defect's incidence is less than T; otherwise, the reliability is one. That is, $p(t) = \delta_{[T,\infty)}(t)$ where the number T is a characteristic of the type of test chosen.

For a screening program in which a test with the above reliability characteristics is applied with frequency r, let $A(r,T)$ represent the expected long run disutility per unit time.

Of course, the "blind period" of the test for 0 to T does not, in a mass screening situation, delay detection of each arriving defect exactly T units of time. The amount of delay depends on the interplay among the time of arrival of the defect, the testing schedules $\{1/r, 2/r, \ldots\}$, and the magnitude of T.

The primary results in this section are a simple characterization of $A(r,T)$ and rules which, in some cases, will permit selection betwseen two tests which differ in their reliability (i.e., in their detection threshold T) and in their cost per application.

Long Run Disutility

PROPOSITION 4: If $p(\cdot) = \delta_{[T,\infty)}(\cdot)$ for some $T \geqslant 0$, then

$$A(r,T) = N\lambda r \int_0^r D(T+u)\, du.$$

Suppose, for example, that $D(t) = \exp(at)$ for $a > 0$. Then, $A(r,T) = N\lambda r \exp(aT)[\exp(a/r) - 1]/a$.

Test Selection

Suppose the decision maker has two kinds of tests available and he must choose one of them for implementation in a mass screening program. Suppose the first kind of test—call it test No. 1— has sensitivity characterized by the "time-until-detectability" threshold T_1. Let $c_1 > 0$ be the cost per application (to an individual unit) of this kind of test. For the second kind of test under consideration, test No. 2, let T_2 and c_2 be the corresponding parameters.

Assume test No. 1 is better in the sense that $T_1 < T_2$. To avoid triviality, assume $c_1 > c_2$.

If the exact shape of the function $D(\cdot)$ is known, Proposition 4 can be used to decide which test to use for each possible level of budget. Let b be the budget per unit time per individual in the population. Then the use of test No. 1 will permit a testing frequency of b/c_1 and the use of test No. 2 permits frequency b/c_2. To decide which test to use, compare the expected disutilities per unit time assuming a fully allocated budget, i.e., compare $A(b/c_1, T_1)$ and $A(b/c_2, T_2)$. With $D(\cdot)$ known, these quantities can be evaluated explicity by Proposition 4 and compared.

It is clear that the entire budget should be allocated because, when $D(\cdot)$ is an increasing function,

$$\frac{\partial}{\partial r} A(r,T) = N\lambda\left[\int_0^{1/r} D(T+u) - (1/r)\, D(T+\frac{1}{r})\, du\right] < 0.$$

When the exact form of the disutility function is not known, Proposition 4 does not suffice to select between tests No. 1 and 2. However, the two following theorems will permit such a determination at least for certain relative configurations of budget, relative test sensitivity $T_2 - T_1$, and cost differential $c_2 - c_1$ of the tests.

Specifically, Proposition 5 will show that for any (increasing) disutility function, test No. 1 is indicated if the budget (per unit population) exceeds $(c_1 - c_2)/(T_2 - T_1)$. On the other hand, Proposition 6 shows that for a convex increasing disutility function, test No. 2 is better if the budget is less than $(c_1 - c_2)/[2(T_2 - T_1)]$.

A decision rule for the case where the budget falls between $(c_1 - c_2)/(T_2 - T_1)$ and $(c_1 - c_2)/[2(T_2 - T_1)]$ has not been found for general disutility functions.

Just how the statement of Propositions 5 and 6 are translated into the above decision rules is explained after the statements of the respective theorems.

PROPOSITION 5: Given $D(\cdot)$ a strictly increasing function, $T_1 < T_2$ and $c_1 > c_2$, then $T_2 - T_1 > (c_1 - c_2)/rc_2$ implies

$$A(rc_2/c_1, T_1) < A(r, T_2), \tag{4.1}$$

making test No. 1 preferable at the per unit population budget level of rc_2.

To apply test No. 1 with frequency rc_2/c_1 versus test No. 2 applied with frequency r (actions reflected, respectively, in the left- and right-hand sides of Eq. (4.1)) would require the same budget, $b = rc_2$, per unit population. The hypothesis of theorem 2 implies $b = rc_2 > (c_1 - c_2)/(T_2 - T_1)$. With the hypothesis in this form, the theorem provides a lower bound on the budget which is a sufficient condition for test No. 1 to entail lower expected disutility per unit time vis-a-vis test No. 2, were the two tests scheduled at their maximal (subject to budget) frequencies rc_2/c_1 and r, respectively.

The following lemma is needed for the proof of Proposition 6 and is recorded here for general interest.

LEMMA 1: If f is a convex function, the $(1/y)\int_t^{t+y} f(s)\,ds \leqslant (1/2)[f(t) + f(t+y)]$.

PROPOSITION 6: If $D(\cdot)$ is convex and increasing, $T_1 < T_2$ and $c_1 > c_2$, then $T_2 - T_1 \leqslant (c_1 - c_2)/2rc_2$ implies $A(rc_2/c_1, T_1) \geqslant A(r, T_2)$, making test No. 2 preferable at the per unit population budget level of $b = rc_2$.

The hypothesis of this proposition implies $b = rc_2 \leqslant (c_1 - c_2)/2(T_2 - T_1)$. Hence, Proposition 6 indicates the superiority of test No. 2 when the budget per unit population is less than $(c_1 - c_2)/2(T_2 - T_1)$.

REFERENCES

[1] Barlow, R., L. Hunter, and F. Proschan, "Optimum Checking Procedures," Journal SIAM, Vol. *11*, No. 4, pp. 1078-95 (1963).

[2] Butler, David A., "A Hazardous Inspection Model," Technical Report No. 187, Department of Operations Research and Department of Statistics, Stanford University (1977).
[3] Derman, C., "On Minimax Surveillance Schedules," Naval Research Logistics Quarterly, Vol. *8*, pp. 415-9 (1961).
[4] Keller, J., "Optimum Checking Schedules for Systems Subject to Random Failure," Management Science, Vol. *21*, pp. 256-60 (1974).
[5] Kirch, R. and M. Klein, "Surveillance Schedules for Medical Examinations," Management Science, Vol. *20*, pp. 1403-9 (1974).
[6] Lincoln, T. and G. H. Weiss, "A Statistical Evaluation of Recurrent Medical Examinations," Operations Research, Vol. *12*, pp. 187,205 (1964).
[7] Pierskalla, W.P. and J.A. Voelker, "A Model for Optimal Mass Screening and the Case of Perfect Test Reliability," Technical Report #3, Department of Industrial Engineering and Management Science, Northwestern University (1976).
[8] Roeloffs, R., "Minimax Surveillance Schedules with Partial Information," Naval Research Logistics Quarterly, Vol. *10*, pp. 307-22 (1963).
[9] Roeloffs, R., "Minimax Surveillance Schedules for Replacement Units," Naval Research Logistics Quarterly, Vol. *14*, pp. 461-71 (1967).
[10] Shwartz, M., and H. Galliher, "Analysis of Serial Screening in an Asymptotic Individual to Detect Breast Cancer," Tech. Report, Dept. of Ind. and Operations Engr., College of Engineering, University of Michigan (1975).
[11] Thompson, D., and R. Disney, "A Mathematical Model of Progressive Diseases and Screening," presented at the November 1976 ORSA TIMS Meeting.
[12] Voelker, J.A., "Contributions to the Theory of Mass Screening," Ph.D. Dissertation, Northwestern University (1976).

APPENDIX

PROPOSITION 1: With $q = 1 - p$,

$$\frac{\partial}{\partial p} C(n,p) = N\lambda r \sum_{n=0}^{\infty} q^n W(n) - N\lambda rp \sum_{n=0}^{\infty} n\, q^{n-1}\, W(n)$$

where

$$W(n) = \int_{n/r}^{(n+1)/r} D(s)\, ds.$$

Now let $V_0 = W_0$; $V_n = W_n - W_{n-1}$, $n = 1, 2, \ldots$. Note that $D(\cdot)$ non-negative increasing implies $V_j > 0$ for $j = 0, 1, 2, \ldots$. Then $W_n = \sum_{j=0}^{n} V_j$ and

$$\begin{aligned}
\frac{\partial}{\partial p} C(r,p) &= N\lambda r \sum_{n=0}^{\infty} q^n \sum_{j=0}^{n} V_j - N\lambda rp \sum_{n=0}^{\infty} n\, q^{n-1} \sum_{j=0}^{n} V_j \\
&= N\lambda r \sum_{j=0}^{\infty} V_j \Big[\sum_{n=j}^{\infty} q^n - p \sum_{n=j}^{\infty} n\, q^{n-1}\Big] \\
&= -N\lambda r \sum_{j=0}^{\infty} j\, q^{j-1}\, V_j < 0.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial r} C(r,p) &= N\lambda p \sum_{n=0}^{\infty} q^n \int_{n/r}^{(n+1)/r} D(s)\, ds \\
&\quad - N\lambda p \sum_{n=0}^{\infty} q^n \left[\frac{n+1}{r} D\left(\frac{n+1}{r}\right) - \frac{n}{r} D\left(\frac{n}{r}\right)\right]
\end{aligned}$$

$$= N\lambda p \sum_{n=0}^{\infty} q^n \left[\int_{n/r}^{(n+1)/r} D(s)ds - \frac{n}{r}\left\{ D\left(\frac{n+1}{r}\right) - D\left(\frac{n}{r}\right)\right\} - \frac{1}{r} D\left(\frac{n+1}{r}\right)\right] < 0.$$

This inequality follows from $D(\cdot)$ increasing through the relations $D[(n+1)/r] - D(n/r) > 0$ and

$$\int_{n/r}^{(n+1)/r} D(s)ds - \frac{1}{r} D\left(\frac{n+1}{r}\right) < \int_{n/r}^{(n+1)/r} D\left(\frac{n+1}{r}\right)ds - \frac{1}{r} D\left(\frac{n+1}{r}\right) = 0.$$

Q.E.D.

PROOF OF PROPOSITION 2: $C(r,p) = N\lambda rp \sum_{n=0}^{\infty} q^n \int_{n/r}^{(n+1)/r} \sum_{i=1}^{l} a_i t^{m_i} dt$

$$= N\lambda rp \sum_{i=1}^{l} \sum_{n=0}^{\infty} a_i q^n/(m_i+1)\left[\left(\frac{n+1}{r}\right)^{m_i+1} - \left(\frac{n}{r}\right)^{m_i+1}\right]$$

$$= N\lambda p \sum_{i=1}^{l} a_i/r^{m_i}(m_i+1) \left[\sum_{n=1}^{\infty} q^{n-1} n^{m_i+1} - \sum_{n=1}^{\infty} q^n m^{m_i+1}\right]$$

$$= N\lambda p \sum_{i=1}^{l} a_i/r^{m_i}(m_i+1) \sum_{n=1}^{\infty} n^{m_i+1} p\, q^{n-1}$$

Q.E.D.

PROOF OF PROPOSITION 3: $C(r,p) = N\lambda p \sum_{k=0}^{\infty} q^k \frac{r}{a} e^{ak/r}(e^{a/r}-1)$

$$= \frac{N\lambda pr}{a}(e^{a/r}-1)\sum_{k=0}^{\infty}(q\, e^{a/r})^k.$$

The geometric series $\sum_{k=0}^{\infty}(q\, e^{a/r})^k$ converges if, and only if, $r > -a/\log q$. Therefore, for $r > -a/\log q$,

$$C(r,p) = \frac{N\lambda pr}{a}(e^{a/r}-1)\frac{1}{1-q\, e^{a/r}}.$$

Q.E.D.

PROOF OF PROPOSITION 4: Note that

$$\prod_{m=1}^{n-1}\left[1 - \delta_{[T,\infty)}\left(u - \frac{m}{r}\right)\right] = 1 - \delta_{[T,\infty)}\left(u - \frac{1}{r}\right).$$

By Eq. (2.1),

$$A(r,T) = N\lambda r \sum_{n=1}^{\infty} \int_{(n-1)/r}^{n/r} D(u)\delta_{[T,\infty)}(u)\prod_{m=1}^{n-1}\left[1 - \delta_{[T,\infty)}\left(u - \frac{m}{r}\right)\right]du$$

$$= N r \sum_{n=1}^{\infty} \int_{(n-1)/r}^{n/r} D(u)\delta_{[T,\infty)}(u)\left[1 - \delta_{[T,\infty)}\left(u - \frac{1}{r}\right)\right]du.$$

Now $u < T + (1/r) \Leftrightarrow u - (1/r) < T \Leftrightarrow 1 - \delta_{[T,\infty)}[u - (1/r)] = 1$. Therefore,

$$A(r,p) = N\lambda r \int_0^\infty D(u)\delta_{[T,\infty)}(u)\ \delta_{[-\infty,T+1/r)}(u)\ du$$
$$= N\lambda r \int_T^{T+1/r} D(u)\,du$$

Q.E.D.

PROOF OF PROPOSITION 5: By Proposition 4, Eq. (4.1) is equivalent to

$$(rc_2/c_1) \int_0^{c_1/rc_2} D(T_1 + s)\ ds < r \int_0^{1/r} D(T_2 + s)\ ds.$$

By hypothesis, $T_2 - T_1 > (c_1 - c_2)/(rc_2)$ which implies $T_2 - T_1 > u(c_1 - c_2)/c_2$ for $u \in [0, 1/r]$ or $T_2 - T_1 + u > uc_1/c_2$ for $u \in [0, 1/r]$. Since $D(\cdot)$ is increasing, this inequality implies

$$\frac{rc_2}{c_1} \int_0^{c_1/(rc_2)} D(T_1 + s)\ ds = r\int_0^{1/r} D(T_1 + uc_1/c_2)\ du$$
$$< r \int_0^{1/r} D(T_1 + (T_2 - T_1 + u))\ du = r \int_0^{1/r} D(T_2 + u)\ du.$$

Q.E.D.

PROOF OF LEMMA 1: Let $h(t + s) = f(t) + (s/y)[f(t + y) - f(t)],\ s \in [0,y]$. Since f is convex, $f(t + s) \leqslant h(t + s),\ s \in [0,y]$. Hence, $\frac{1}{y} \int_t^{t+y} f(s)\,ds \leqslant \frac{1}{y} \int_t^{t+y} h(s)\,ds$ $= \frac{1}{y} \int_0^{y} h(t + s)\,ds = \frac{1}{2}[f(t + y) + f(t)]$.

Q.E.D.

PROOF OF PROPOSITION 6: By Proposition 4, it suffices to show

$$(rc_2/c_1) \int_0^{c_1/rc_2} D(T_1 + u)\ du \geqslant r \int_0^{1/r} D(T_2 + u)\ du.$$

Letting $f(t) = D(T_1 + t)$ and $T = T_2 - T_1$, this becomes

$$(rc_2/c_1) \int_0^{c_1/rc_2} f(s)\,ds \geqslant r \int_0^{1/r} f(T + s)\,ds \tag{A.1}$$

Note that $D(\cdot)$ being an arbitrary convex increasing function implies $f(\cdot)$ is also convex increasing.

Let $x = c_1/rc_2$ and let $a = c_2/c_1 < 1$. Then Eq. (A.1) becomes

$$(1/ax) \int_T^{T+ax} f(s)\,ds \leqslant (1/x) \int_0^{x} f(s)\,ds$$

which will follow from

$$\frac{1}{ax} \int_T^{T+ax} f(s)\,ds \leqslant \frac{1}{2T + ax} \int_0^{2T+ax} f(s)\,ds \leqslant \frac{1}{x} \int_0^{x} f(s)\,ds \tag{A.2}$$

Note that the above expressions are all average values of $f(s)$ over their respective intervals of integration. The right-hand inequality follows from $f(\cdot)$ increasing and $2T + ax \leqslant x$. To prove the theorem, it only remains to estabilsih the left-hand inequality in Eq. (A.2).

$$\frac{1}{2T + ax} \int_0^{2T+ax} f(s)\,ds - \frac{1}{ax} \int_T^{T+ax} f(s)\,ds \tag{A.3}$$
$$= \int_0^{T} \frac{d}{d\delta} \left[\frac{1}{2\delta + ax} \int_{T-\delta}^{T+ax+\delta} f(s)\,ds\right] d\delta$$

$$\frac{d}{d\delta}\left[\frac{1}{ax+2\delta}\int_{T-\delta}^{T+ax+\delta} f(s)\,ds\right]$$
$$= \frac{2}{ax+2\delta}\left[\frac{1}{2}\left(f(T-\delta)+f(T+ax+\delta)\right) - \frac{1}{ax+2\delta}\int_{T-\delta}^{T+ax+\delta} f(s)\,ds\right]$$

which is non-negative for $\delta \in [0,T]$ by Lemma 1. Therefore, (A.3) is non-negative and

$$\frac{1}{2T+ax}\int_{0}^{2T+ax} f(s)\,ds \geqslant \frac{1}{ax}\int_{T}^{T+ax} f(s)\,ds.$$

Q.E.D.