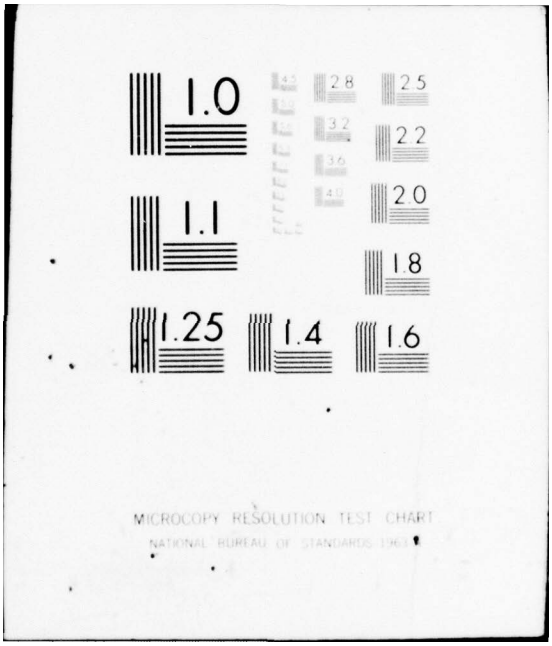
AD-A038 257 NORTHWESTERN UNIV EVANSTON ILL DEPT OF INDUSTRIAL E--ETC F/G 12/1
A MODEL FOR OPTIMAL MASS SCREENING AND THE CASE OF PERFECT TEST--ETC(U)
DEC 76 W P PIERSKALLA, J A VOELKER N00014-67-A-0356-0030
UNCLASSIFIED TR-3 NL

1 OF 1
AD A038257

END

DATE FILMED
4-77


1.0
1.1
1.25
1.4
1.6
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963

(12)

(6)

A MODEL FOR OPTIMAL MASS SCREENING AND THE CASE OF PERFECT TEST RELIABILITY.

by

(10) William P. Pierskalla*
and
John A. Voelker**

(9) Technical Report, No. 3 (14) TR-3

(11) 31 December 1976

(12) 44p.

(15) Prepared under Office of Naval Research
Contracts N00014-67-A-0356-0030 (Task NR042-322)
and N00014-75-C-0797 (Task NR042-322)
W. P. Pierskalla, Project Director

*Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois 60201

**Corporate Research Department, Montgomery Ward Company, Chicago, Illinois 60680

ABSTRACT

This paper considers a population (composed, for example, of people, machines, or livestock) subject to a randomly occurring defect or disease. If there exist testing procedures capable of detecting the defect before it would otherwise become known, and if such early detection provides benefit, the periodic administration of such a test procedure to the members of the population, i.e., a mass screening program, may be advisable. An analytical model of a mass screening program is developed and analyzed. In the development of the model, the defect arrival process is Poisson and the ability of the test to recognize a defect is not, in general, assumed perfect.

The model focuses on the time since incidence of a defect as the factor which determines the reliability of a test administration and the value of a detection. The objective is to minimize an arbitrary increasing function of the detection delay which represents the disutility from the occurrence of the defect and its detection. This disutility is also represented as a function of the type of test applied, the probability of success of the test, the testing intervals and the arrival rates of the defects over Q subpopulations. A relatively simple expression is derived for the objective function and a comprehensive mathematical program is presented. The case of perfect test reliability is then considered and, from the class of "cyclic" schedules, the optimum schedules are shown to be equally spaced. For a polynomial disutility function, properties of the optimal schedules are presented. Finally, a method for determining the disutility function from experimental data is suggested.

A MODEL FOR OPTIMAL MASS SCREENING AND THE CASE OF PERFECT TEST RELIABILITY

There are many situations where a defect can occur randomly among the members of a population--either a population of human beings, inanimate objects, or perhaps livestock--and once present, exist and develop, at least for a time, without any manifest symptoms. If the early detection of such a defect provides benefit, it may be worthwhile to employ a test capable of revealing the defect's existence in its earlier stages. (A defect, disorder, or disease will generally be considered simply as a defect; and the word unit or individual will refer to a member of the population.)

Of course, continuous monitoring would provide the most immediate such revelation. But considerations of expense and practicality will frequently rule out continuous monitoring so that a schedule of periodic testing--a 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 stage of development.

The benefits of early detection often depend upon the application considered. For example, in a human population being screened for some chronic disease (cancer, diabetes, glaucoma, heart disease, etc.) 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 cost. 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. This process of inspecting a sizable population for defects is called mass screening.

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 rationale for constructing a mathematical model of mass screening is to provide a conceptual framework within which a mass screening program might best be designed and its worth evaluated. Such a design must determine which kind of testing technology will be used. Several candidate technologies may be available, each with different reliability characteristics and costs. In addition, the frequency of testing must be decided. If the target population can be partitioned into subpopulations according to susceptibility to the defect (e.g., by age, family background, time since last overhaul (for a machine), etc.), then the best allocation of

the testing budget among the subpopulations must be determined. Lastly, a decision must be made, for the population and type of defect under consideration, whether a mass screening program is justified at all. It is felt that the above determinations are best carried out within the conceptual framework of a cogent model of mass screening.

It is the purpose of this paper to present a reasonably comprehensive model for decision making with regard to mass screening. The model utilizes a time-based approach. But rather than seeking to analyze or minimize detection delay, as in some of the literature on this topic, the objective function is an arbitrary increasing function of detection delay. (Detection delay is the time between the incidence of the defect and its detection, regardless of whether that detection is the result of a screening test or of the defect becoming self-evident.) The reason for choosing a general function is that the disutility experienced upon the delayed detection of a defect may well vary in a highly nonlinear way with the length of the delay.

Another objective sometimes used for inspection models is to maximize the lead time where the lead time is the time difference between the time of detection via a screening test and the time detection would otherwise have occurred had not a screening program been in existence. However, in the case in which the test being used is perfectly reliable then it is possible to consider the lead-time criterion in terms of the following simple function of detection delay. Suppose that at the (possibly random) age T (measured from the time of incidence of the defect) the defect, in the normal course of its development, would become manifest even without a screening test. Then, if the defect is detected at age t, the lead time

gained is $(T - t)^+$. Therefore, if sup T is finite, the function $D(t) = \sup T - E(T - t)^+$ may be used as a disutility function which, when minimized, will maximize the expected lead time.

Consequently, the results presented in this paper, based on an arbitrary disutility function, generalize as well as extend results in earlier work. In addition, new results are presented.

In the next section a brief review of the literature is given. Since the work on inspection models is very large, only the most relevant papers are discussed. The interested reader may refer to the surveys by McCall [1965] and Pierskalla and Voelker [1976] for a more comprehensive review.

In the third section the model is formally stated and an expression derived for the expected disutility per unit time incurred under a regime of uniformly spaced test applications. In section 4, the model is specialized to the case of a perfect test; i.e., when the test is administered to an individual with the defect, the defect will be detected with certainty. Under this assumption, a regime of uniform test intervals is proved optimal within a wider class of "cyclic" testing schedules for a single subpopulation. The problem of allocating a screening budget among various subpopulations is also analyzed. As part of this analysis, the explicit screening schedule $(r_1, r_2, \ldots, r_Q)$, where r_j designates the optimal testing frequency for members of the jth subpopulation, is obtained in terms of the subpopulation specific incidence rates $N_j\lambda_j$ and the budget constraint for the disutility function $D(t) = at^m$. For $D(\cdot)$ convex, the above solution (when $m = 1$) is shown to provide a bound on the ratios r_i/r_j.

The last section provides a technique to estimate the shape of the disutility function from empirical data in the case of a perfect test. The Appendix contains the proofs of all results.

Finally, it should be mentioned that the population (or each subpopulation, in the instances where a heterogeneous population is considered) is assumed to be of fixed size, N, and the defect to arrive according to a stationary Poisson process with rate $N\lambda$ (N could be a very large but finite number). It may be somewhat more realistic to set the defect arrival rate proportional to the number of defect-free units, rather than to the total number of units in the population. However, it is also assumed that no defect can remain undetected longer than T^*, even without any screening tests being given. Hence, $T^*N\lambda$ is an upper bound on the expected number of undetected defects in the population. If it is the case that once a defect is detected, the afflicted unit is replaced in the population with a healthy unit, then $T^*N\lambda$ represents a bound on the expected difference between the number of healthy units and the total number of units in the population. For λ small, as would be the case for a relatively infrequently occurring disorder, this should represent no difficulty. Consequently, it is assumed that λ is small relative to N, which is consistent with the examples of potential applications which have been or will be mentioned.

2. Literature Review

Some of the early papers which have a bearing on time dependent models of mass screening considered a problem formulated as follows: Suppose a single unit is subject to failure and failure may be detected only through an inspection. Each inspection has an associated cost and there is another cost proportional to the duration of undetected failure. For this problem Derman [1961] found a minimax testing schedule when the failure distribution

is completely unknown and there is a finite time horizon. The reliability of the test is a constant, p, not necessarily equal to one.

Roeloffs [1963, 1967] addresses essentially the same problem with $p = 1$. However, the decision maker knows a single percentile for the time-to-failure distribution of the unit.

Barlow, Hunter, and Proschan [1963] derive explicit rules for calculating the optimal inspection schedule for a broad class of failure distributions (now assumed completely known to the decision maker). The test is assumed perfectly reliable, i.e., $p = 1$.

Keller [1974], again assuming $p = 1$, developed an alternative computational technique utilizing the calculus of variations to characterize the optimal inspection schedule. Again there is a fixed cost for each inspection. And there is a cost which is a function, not necessarily linear, of the detection delay. In order to use the calculus of variations, Keller assumed that the intervals between inspections are sufficiently small (compared, say, to the mean time-to-failure) so that the inspection schedule may be expressed as a density function mapping the unit's age into an associated testing frequency.

Kirch and Klein [1974], whose paper is addressed explicitly to a mass screening application, seek an inspection schedule which will minimize expected detection delay subject to a constraint on the expected number of examinations an individual would incur over a lifetime. The test is assumed perfectly reliable. Kuhn-Tucker methods are utilized in solving for the optimal screening schedule. The point of view adopted in this paper is similar to that of Kirch and Klein and Theorem 3 is related (although it is somewhat more general) to their main result. One

respect in which the approach here differs from that of Kirch and Klein is that this paper considers several subpopulations each with its own characteristic incidence rate for the disorder. The idea is then to allocate optimally a fixed screening budget among the subpopulations. Kirch and Klein instead take a longitudinal view. An individual through his lifetime is subject to different probabilities of incurring the defect and a screening schedule is optimized subject to a constraint on the expected number of examinations over a lifetime.

McCall [1969] considered the problem of scheduling dental examinations under the assumption that the time between the incidence of a cavity and the scheduled dental examination controls whether the cavity results in a filling or an extraction. Cavities are assumed to occur according to a Poisson process. As a generalization, he permits the time required for a cavity to become *beyond repair* (by a filling) to be a random variable.

Lincoln and Weiss [1964] studied the statistical characteristics of detection delay under the assumption that the times of examinations form a renewal process and that the probability of detecting the defect, $p(t)$, is a function of the defect's age, t. They derive open form equations, similar to renewal type equations, which relate the density functions for the following entities: the probability of detection at a test application ($p(t)$), the time until the defect becomes potentially detectable and from this time the forward recurrence time to the first test, the probability of the event that at a particular time a test occurs and all prior tests had failed to detect the defect, and the detection delay. For the two special cases where $p(t)$ is a constant and where $p(t)$ is exponential,

the moments for the detection delay are derived in closed form. For uniform testing intervals (and general $p(\cdot)$), the distribution and moments of the detection delay are computed. For $p(t) = 1$, they solve for that testing schedule which maximizes the time between tests subject to a constraint on the performance of the screening program relative to detection delay. Two such constraints are considered. The first bounds the probability of detection delay exceeding some threshold T. The second bounds the mean detection delay.

Theorem 1 in the next section, which gives the expected disutility in terms of the test interval and test type, is similar to the objective function given by Lincoln and Weiss although their approach, as outlined above, is different.

Several recent papers have studied the statistical characteristics of the lead time provided by a screening program--either a one-shot screen of the population or a periodic screening program. One of the earliest such papers was that of Hutchison and Shapiro [1968].

Zelen and Feinleib [1969] point out an important use for the expected lead time in comparing survival time data (measured from the time of detection) gathered from a screened population versus that gathered from a control population. The expected lead time must be subtracted from the expected survival time of the screened population before a valid comparison can be made against the control group's expected survival time. In a later paper, Zelen [1971] derived expressions for mean lead time in the case of two successive screens.

Prorok [1973] has established properties for the lead time in the case of n successive screening tests conducted at uniform intervals. He

also derives the proportion of preclinical cases detected, both at a given test and over all n tests. The test is assumed to have a constant reliability not necessarily equal to one. In addition, he derives an estimator for the mean lead time.

3. A Model for Mass Screening

Although most of the definitions are stated prior to their use in each section, listed below are some notation and conventions used at various times throughout the paper.

(i) $1_{\frac{m}{r}}(t) = 1_{[\frac{m}{r}, \frac{m+1}{r}]}(t) = \begin{cases} 1 & \text{if } \frac{m}{r} \le t \le \frac{m+1}{r} \\ 0 & \text{otherwise} \end{cases}$

(ii) $\prod_{i=1}^{n} x_i = 1 \qquad \text{for } n < 1$

(iii) $\sum_{i=1}^{n} x_i = 0 \qquad \text{for } n < 1$

(iv) $[x]$ is the largest integer not exceeding x.

(v) The phrase "increasing function" shall mean a strictly increasing function.

As mentioned previously, the objective function will not be to maximize expected utility, but rather to minimize expected disutility. The disutility associated with a particular detection will depend only on the elapsed time since the defect's incidence. The notation D(t) will express the disutility incurred if detection occurs t units of time after incidence. $D(\cdot)$ is assumed throughout this paper to be a nonnegative increasing function.

In those cases where $\sup_{s \geq 0} D(s) < \infty$, there is a natural relation between the notion of a utility function and $D(\cdot)$; namely,

$$U(t) = \sup_{s \geq 0} D(s) - D(t) .$$

$U(t)$ is a decreasing function expressing the disutility avoided by detecting a defect t units of time after its incidence.

Initially, only a single susceptibility class of size N will be considered. The results will then be generalized to an arbitrary number of subclasses. The times of incidence for the defect in the population are assumed to form a Poisson process with parameter $N\lambda$ and is designated by the sequence $\{S^k\}$, $k = 1, 2, \ldots$ This assumption is motivated by the fact that any arrival process with the following characteristics is a Poisson process: at the time of an arrival there is almost surely (i.e., with probability one) only one arrival, that the number of arrivals in a time interval does not depend on past arrivals, and that the number of arrivals in intervals of equal length are identically distributed (Çinlar [1975]). For many applications such as the arrival of diseases like cancer, heart trouble, etc., this assumption is quite reasonable.

Were a screening test of type ℓ to be administered to the individual with the kth defect (i.e., kth in the order of defect incidence) at time $S^k + t$, then the test outcome random variable is:

Definition:

$$Y_\ell^k(t) = \begin{cases} 0 & \text{if kth defect is not detected at time t after incidence} \\ 1 & \text{if kth defect is detected at time t after incidence.} \end{cases} \qquad (3.1)$$

Since a defect cannot be detected before its incidence, $Y_\ell^k(t) = 0$ for $t < 0$. Notice that the argument of $Y_\ell^k(\cdot)$ refers to time relative to the incidence of the defect. If the population is screened at time s, where s is absolute time, then the kth defect will be detected if and only if $S^k \le s$ and $Y_\ell^k(s - S^k) = 1$.

It is assumed $Y_\ell^k(t)$, $Y_\ell^j(s)$ are independent except when $k = j$ and $t = s$, and that $P\left(Y_\ell^k(t) = 1\right)$ depends only on ℓ and t. This latter assumption makes possible the

$$\underline{\text{Definition}}: \quad p_\ell(t) = P\left(Y_\ell^1(t) = 1\right). \tag{3.2}$$

The function $p_\ell(\cdot)$ describes the reliability characteristics of the type ℓ screening test, i.e., how likely such a test is to detect a defect as a function of the defect's age. Note that $p_\ell(t) = 0$ for $t < 0$.

In this section the screening test is assumed to be administered to the entire population at the times $1/r, 2/r, 3/r, \ldots$ The testing frequency r is a control variable. In the next section, upon assuming perfect test reliability, it is shown that the above schedule of uniform testing intervals is optimal within a wider class of "cyclic" schedules.

The random variable $\bar{S}_{r,\ell}^k$ denotes the time at which the kth defect is detected. $\bar{S}_{r,\ell}^k$ depends on the arrival time of the defect (S^k), the type of test used (ℓ), and the testing frequency (r).

$$\underline{\text{Definition}}: \quad \bar{S}_{r,\ell}^k = \min_{n=1,2,\ldots} \left\{ n/r \,|\, Y_\ell^k\left(\frac{n}{r} - S^k\right) = 1 \right\} \tag{3.3}$$

Given the application of test type ℓ at the times $\{1/r, 2/r, \ldots\}$, the disutility incurred by the kth defect is $D\left(\bar{S}_{r,\ell}^k - S^k\right)$. To represent

the total disutility incurred due to those defects which occurred in the interval $[j/r, (j+1)/r)$, it is useful to define:

Definition: $B_{r,\ell,j} = \sum_{k=1}^{\infty} D\left(\bar{S}^k_{r,\ell} - S^k\right) 1_{j/r}(S^k)$.

The following theorem provides an expression for $E[B_{r,\ell,j}]$ in terms only of the disutility due to detection delay (as expressed by $D(\cdot)$) and of the reliability of test type ℓ (as expressed by $p_\ell(\cdot)$). In addition, the theorem shows that $E[B_{r,\ell,j}] = E[B_{r,\ell,0}]$ for $j = 0, 1, 2,...$

Theorem 1

$$E[B_{r,\ell,j}] = N\lambda \sum_{n=1}^{\infty} \int_{\frac{n-1}{r}}^{\frac{n}{r}} D(u)p_\ell(u) \prod_{m=1}^{n-1} [1 - p_\ell(u - \tfrac{m}{r})]du$$

for $j = 0, 1, 2,...$

This theorem yields a relatively simple expression for the expected total disutility in any interval of length $1/r$ and using test type ℓ with probability of detection $p_\ell(\cdot)$. This expectation is used in the objective function of a mathematical program to determine the optimal testing frequency for a mass screening program for a heterogeneous population. In order to develop this mathematical program, the following definitions are useful.

Definition: $\bar{B}_{r,\ell} = \lim_{n\to\infty} \frac{r}{n} \sum_{j=0}^{n-1} E[B_{r,\ell,j}]$.

$\bar{B}_{r,\ell}$ is the long-run expected disutility per unit time. (The factor r enters the definition to convert disutility per unit testing-interval into

per unit time.) By Theorem 1, $E[B_{r,\ell,j}] = E[B_{r,\ell,0}]$ for $j = 0, 1, 2, \ldots$ Therefore,

$$\bar{B}_{r,\ell} = rE[B_{r,\ell,0}]$$

$$= rN\lambda \sum_{n=1}^{\infty} \int_{\frac{n-1}{r}}^{\frac{n}{r}} D(u)p_{\ell}(u) \prod_{m=1}^{n-1} [1 - p_{\ell}(u - \tfrac{m}{r})]du . \qquad (3.4)$$

Notice that if test type ℓ provides perfect reliability, i.e., $p_{\ell}(t) = 1$ for $t \geq 0$, then (3.4) gives

$$\bar{B}_{r,\ell} = rN\lambda \int_{0}^{\frac{1}{r}} D(u)du . \qquad (3.5)$$

Now consider the problem of selecting testing frequencies and test-types for each of Q different susceptibility classes which together comprise the whole population. These classes may differ from one another in the number of units they contain, in their defect incidence intensity, and in the cost per test application to an individual for a particular type of test. The subpopulations, however, are assumed to share a common $D(\cdot)$ function.

The solution to (3.4) for the expected long-run disutility per unit time in the context of the prior portion of this section may be regarded as applying to subpopulation j where $j = 1, 2, \ldots, Q$ by writing

$$\bar{B}_{j,r(j),\ell(j)} = r(j)N_j\lambda(j) \sum_{i=1}^{\infty} \int_{\frac{i-1}{r(j)}}^{\frac{i}{r(j)}} D(u)p_{\ell(j)}(u) \prod_{m=1}^{i-1} [1 - p_{\ell(j)}(u - \tfrac{m}{r(j)})]du .$$

This permits a multi-subpopulation screening problem subject to a budget constraint to be formulated as follows:

$$\underset{(r(1),\ldots,r(Q),\ell(1),\ldots,\ell(Q))}{\text{Minimize}} \quad \sum_{j=1}^{Q} \bar{B}_{j,r(j),\ell(j)} \tag{3.6}$$

such that

$$\sum_{j=1}^{Q} N_j c_{j,\ell(j)} r_j \leq b \tag{3.7}$$

$$r_j > 0 \qquad j = 1,\ldots, Q \tag{3.8}$$

$$\ell(j) \in \mathcal{L} \qquad j = 1,\ldots, Q \tag{3.9}$$

where

$c_{j,\ell(j)}$ = cost per application of a test of type $\ell(j)$ to an individual of subpopulation j,

N_j = number of units (or individuals) in subpopulation j,

b = budget per unit time.

In order to make this mathematical program even more comprehensive, it is possible to add constraints on the amount of testing labor available and on the capacity of the testing facilities in terms of the number of arrivals, the frequency of testing and the type of tests used. In addition, the cost of false positives can be included as a part of the test costs in inequality (3.7).

For example, a constraint on the total labor available is:

$$\sum_{j=1}^{Q} N_j \delta_{j,\ell(j)} r_j \leq L_\ell \tag{3.10}$$

and on the total testing facilities available is:

$$\sum_{j=1}^{Q} N_j f_{j,\ell(j)} r_j \leq F_\ell \tag{3.11}$$

for each type of test $\ell(\cdot)$ used over the subpopulations being tested, where

$\delta_{j,\ell(j)}$ = amount of labor needed to administer test type $\ell(j)$ to an individual of subpopulation j,

$f_{j,\ell(j)}$ = amount of testing facility time needed to administer test type $\ell(j)$ to an individual of subpopulation j,

L_ℓ = total amount of labor available to administer test type $\ell(\cdot)$ per unit time,

F_ℓ = total amount of facility time available to administer test type $\ell(\cdot)$ per unit time.

4. The Case of Perfect Test Reliability

In this section, it is assumed that $p_\ell(t) = 1$ for all $t \geq 0$; that is, there is perfect test reliability. If a defect is present at the time the test is administered, that defect will be detected with certainty.

Unlike the previous section, uniform testing intervals are not assumed. Rather, such uniform intervals (within a subpopulation) will be proven optimal within a larger class of "cyclic" testing schedules. Once this is done, the problem of allocating a screening budget for a target population among Q subpopulations which comprise the target population is analyzed. That part of the budget allocated to a particular subpopulation controls how frequently the screening test can be administered to that subpopulation. This budget allocation, in turn, is determined by the relative incidence rates of the disorder and the relative costs of testing which characterize the different subpopulations.

Until now, r was used to designate the testing frequency and 1/r the interval between tests. In order to consider more general testing schedules,

a different notation is needed. Suppose that the jth subpopulation is tested according to an m(j)-cycle. To be specific, the intervals between tests are successively $x_1^j, x_2^j, \ldots, x_{m(j)}^j, x_1^j, \ldots, x_{m(j)}^j, \ldots$ Letting $y_i^j = \sum_{\nu=1}^{i} x_\nu^j$, the successive times of test administration for subpopulation j are $y_1^j, y_2^j, \ldots, y_{m(j)}^j, y_{m(j)}^j + y_1^j, \ldots$

Let $\bar{D}\left(x_1^j, \ldots, x_{m(j)}^j\right)$ be the long-run expected disutility per unit time for the jth subpopulation under the above testing schedule. The problem then becomes the following:

$$\left.\begin{array}{ll} \underset{\substack{(m(1),\ldots,m(Q)) \\ (x_1^1,\ldots,x_{m(1)}^1) \\ \vdots \\ (x_1^Q,\ldots,x_{m(Q)}^Q)}}{\text{Minimize}} & \sum_{j=1}^{Q} \bar{D}\left(x_1^j, \ldots, x_{m(j)}^j\right) \\ \text{such that} & \sum_{j=1}^{Q} N_j c_j m(j)/y_{m(j)} \leq b \\ & x_i^j > 0\,; \quad i = 1, \ldots, m(j); \quad j = 1, \ldots, Q \end{array}\right\} \qquad (4.1)$$

where

N_j = the number of individuals or units in the jth subpopulation,

c_j = cost per test application per individual in the jth subpopulation,

b = bound on the average cost per unit time for all tests administered under the screening program.

For the moment consider a homogeneous population, i.e., let Q = 1, and suppress the index j in the above cyclic testing schedule. Under the

assumption of a perfectly reliable test, if the kth defect occurs before time y_m, it will be detected at test time y_i if and only if $y_{i-1} < S^k \leq y_i$.

Therefore, the total disutility incurred due to all defects arising in the interval $(0, y_m]$ is

$$\sum_{i=1}^{m} \sum_{k=1}^{\infty} D(y_i - S^k) 1_{(y_{i-1}, y_i]}(S^k)$$

where $y_0 = 0$, although a test is not made at time 0.

Similarly, the total disutility due to defects arising in the interval $(\nu y_m, \nu y_m + y_m)$ for $\nu = 0, 1, 2, \ldots$ is

$$\sum_{i=1}^{m} \sum_{k=1}^{\infty} D(\nu y_m + y_i - S^k) 1_{(\nu y_m + y_{i-1}, \nu y_m + y_i]}(S^k) . \tag{4.2}$$

Taking expectations and using Lemma A given in the Appendix, (4.1) becomes

$$\sum_{i=1}^{m} \lambda N \int_{\nu y_m + y_{i-1}}^{\nu y_m + y_i} D(\nu y_m + y_i - s) ds$$

$$= \sum_{i=1}^{m} \lambda N \int_{y_{i-1}}^{y_i} D(y_i - s) ds = \sum_{i=1}^{m} \lambda N \int_{0}^{x_i} D(t) dt .$$

Hence, the expected disutility due to defects arising in the interval $[\nu y_m, \nu y_m + y_m)$ for $\nu = 0, 1, 2, \ldots$ does not depend on ν and

$$\overline{D}(x_1, \ldots, x_m) = \frac{\lambda N}{y_m} \sum_{i=1}^{m} \int_{0}^{x_i} D(t) dt . \tag{4.3}$$

It will now be shown that for any schedule such that $x_i \neq x_j$ for some $i, j \in \{1, 2, \ldots, m\}$, the alternate schedule $\{\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_m\}$ where

$\bar{x}_1 = \bar{x}_2 = \ldots = \bar{x}_m = \sum_{j=1}^{m} x_j/m$ is feasible with respect to the program and provides at least as small a value for the objective function. Once this is accomplished, among all "cyclic" schedules all future consideration may be restricted to testing schedules characterized by equal intervals between all test administrations (within a single subpopulation). The following lemma will be required:

Lemma 1: For $D(\cdot)$ an increasing function, $h(x) = \int_0^x D(t)dt$ is a (strictly) convex function in x.

Since $h(\cdot)$ is convex, $h(\bar{x}_i) = h\left(\sum_{\nu=1}^{m} x_\nu/m \right) < \sum_{\nu=1}^{m} h(x_\nu)/m$ for $i = 1,\ldots, m$. Adding these m equations together,

$$\sum_{i=1}^{m} h(\bar{x}_i) < \sum_{\nu=1}^{m} h(x_\nu)$$

$$\frac{N\lambda}{y_m} \sum_{i=1}^{m} h(\bar{x}_i) < \frac{N\lambda}{y_m} \sum_{i=1}^{m} h(x_i)$$

where $y_m = \sum_{i=1}^{m} \bar{x}_i = \sum_{i=1}^{m} x_i$. Using (4.3)

$$\bar{D}(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m) < \bar{D}(x_1, x_2, \ldots, x_m) .$$

Hence, the schedule $(\bar{x}_1, \ldots, \bar{x}_m)$ yields lower long-run expected disutility per unit time than does $(x_1, \ldots, x_m)$. Since both schedules require the same number of tests per unit time (m/y_m) and since $\bar{x}_1 = \bar{x}_2 = \ldots = \bar{x}_m$, then any m-cyclic testing schedule may be replaced with a uniform testing

schedule without increasing average testing frequency or expected disutility per unit time.

To return to the general case of n subpopulations, replace the schedule $\left(x_1^j, \ldots, x_{m(j)}^j\right)$ with a uniform schedule utilizing test intervals of $y_{m(j)}^j/m(j)$ for $j = 1, \ldots, n$. Since none of the subpopulations use more of the testing budget than before (the average testing frequency remaining the same for each subpopulation), the new testing schedule is feasible if the old one is.

The above result is given formally by

Theorem 2: If $p(t) = 1$ for $t \geq 0$, schedules with equally spaced test intervals are optimal within the class of m-periodic schedules for $m = 1, 2, \ldots$

Now let $x^j = \bar{x}_1^j = \bar{x}_2^j = \ldots = \bar{x}_m^j$, then by (4.3)

$$\sum_{j=1}^{Q} \bar{D}\left(\bar{x}_1^j, \ldots, \bar{x}_{m(j)}^j\right) = \sum_{j=1}^{Q} \frac{\lambda_i N_i}{y_{m(j)}} \sum_{i=1}^{m(j)} \int_0^{\bar{x}_i^j} D(t)dt$$

$$= \sum_{j=1}^{Q} \frac{\lambda_i N_i}{m(j)x^j} \sum_{i=1}^{m(j)} \int_0^{x^j} D(t)dt$$

$$= \sum_{j=1}^{Q} \frac{\lambda_i N_i}{x^j} \int_0^{x^j} D(t)dt . \qquad (4.4)$$

With uniform testing intervals, it becomes possible to define the testing frequency for subpopulation j as $r_j = 1/x^j$. Note that (4.4) is consistent with (3.5) of section 3 where the assumption of uniform

testing intervals was made and the problem of imperfect test reliability analyzed.

By (4.4), the mathematical program (4.1) can be rewritten

$$\left.\begin{array}{ll} \underset{(r_1,r_2,\ldots,r_Q)}{\text{Minimize}} & \sum_{j=1}^{Q} N_j \lambda_j r_j \int_0^{\frac{1}{r_j}} D(t)dt \\ \text{such that} & \sum_{j=1}^{Q} N_j c_j r_j \le b \\ \text{and} & r_j > 0 \quad \text{for } j = 1, 2, \ldots, Q \end{array}\right\} \tag{4.5}$$

where the constraint equation was transformed using

$$m(j)/y_{m(j)} = m(j)/\left(\bar{x}_1^j m(j)\right) = 1/\bar{x}_1^j = 1/x^j = r_j .$$

The solution of (4.5) will optimally allocate the screening budget among the Q subpopulations so as to minimize the long-run expected disutility per unit time over the entire population. (If r_j must be integer valued, then just add this restriction and solve (4.5) as a constrained knapsack problem.)

For the continuous case, let

$$f(r) = \begin{cases} r \int_0^{\frac{1}{r}} D(t)dt & \text{for } r > 0 \\ \lim_{r \to 0} r \int_0^{\frac{1}{r}} D(t)dt & \text{for } r = 0 . \end{cases}$$

At this point, assume that $D(\cdot)$ is a differentiable function on $[0, \infty)$. Therefore, $f(\cdot)$ is continuously differentiable on $(0, \infty)$ and

$$f'(r) = \frac{df}{dr} = \frac{d}{dr}\left[r \int_0^{\frac{1}{r}} D(s)ds \right] = \int_0^{\frac{1}{r}} D(s)ds - \frac{1}{r} D(\tfrac{1}{r})$$

$$= \frac{1}{r} D(\gamma) - \frac{1}{r} D(\tfrac{1}{r}) < 0 , \quad \gamma \in (0, \tfrac{1}{r}) , \quad r > 0 \qquad (4.6)$$

where the last equality follows from the mean value theorem for integrals. Also, since $D'(\cdot) > 0$

$$\frac{d^2 f}{dr^2} = \frac{d}{dr}\left[\int_0^{\frac{1}{r}} D(s)ds - \frac{1}{r} D(\tfrac{1}{r}) \right] = \frac{1}{r^3} D'(\tfrac{1}{r}) > 0 \quad \text{for } r > 0 . \qquad (4.7)$$

So, on the interval $(0, \infty)$, $f(\cdot)$ is a convex decreasing function in r. The same may therefore be said of the objective function in (4.5). The fact that the objective function is decreasing in each of its components implies that the constraint $\sum_{j=1}^{Q} N_j c_j r_j \leq b$ is binding. Consequently, this inequality may be replaced with equality in (4.5).

Lemma 2: If $\lim_{s \to \infty} D(s) = \infty$, then $\lim_{r \to 0} f(r) = \infty$.

Theorem 3: If $\lim_{t \to \infty} D(t) = \infty$, then there is a unique solution $(r_1^*, r_2^*, \ldots, r_Q^*)$ to (4.5) which satisfies $\lambda_i f'(r_i^*)/c_i = \lambda_j f'(r_j^*)/c_j$ for $i, j = 1, 2, \ldots, Q$.

In the proof of Theorem 3 in the Appendix, it can be seen that the hypothesis that $\lim_{t \to \infty} D(t) = \infty$ was needed only to establish (A.3). Therefore, an alternate theorem can be given.

Theorem 3': If $(r_1^*, r_2^*, \ldots, r_Q^*)$ is a unique solution to (4.5) and if $r_j^* > 0$ for $j = 1, \ldots, n$; then $\lambda_i f'(r_i^*)/c_i = \lambda_j f'(r_j^*)/c_j$ for $i, j = 1, 2, \ldots, Q$.

As a corollary, one would expect that subpopulations with greater susceptibility to the disorder should be tested more frequently. This turns out to be the case provided testing costs are the same.

Corollary 3.1: If $\lim_{t \to \infty} D(t) = \infty$ and $\lambda_i/c_i > \lambda_j/c_j$, then $r_i^* > r_j^*$.

The relation provided by this corollary between subpopulation specific defect incidence rates and cost per test on one hand, and the relative testing frequencies for the various subpopulations on the other hand, does not require any knowledge of the shape of the disutility function $D(\cdot)$, except that it is increasing to infinity. The next two results show how more specific knowledge of $D(\cdot)$ can be used to schedule a mass screening program optimally. The first provides an explicit solution to the program (4.5) when $D(t) = at^m$ for $m > 0$. The second result shows how the solution for the case $D(t) = at$ serves as a bound for the optimal solution when $D(\cdot)$ is any convex increasing function.

Corollary 3.2: If $D(t) = at^m$ for $m > 0$, then

$$r_1^* = b \Big/ \left(\sum_{j=1}^{Q} N_j \sqrt[(m+1)]{\frac{\lambda_j c_j^m c_1}{\lambda_1}} \right) = b \Big/ \left(\sum_{j=1}^{Q} N_j c_j \sqrt[(m+1)]{\frac{\lambda_j c_1}{\lambda_1 c_j}} \right)$$

and

$$r_j^* = r_1^* \sqrt[(m+1)]{\frac{\lambda_j c_1}{\lambda_1 c_j}} \qquad j = 1, 2, \ldots, Q .$$

The following result does not require knowledge of the exact form of the function $D(\cdot)$--provided that it is differentiable and convex increasing.

Theorem 4: If $D'(\cdot)$ exists and is increasing, then $r_2^*/r_1^* > \sqrt{\lambda_2 c_1/\lambda_1 c_2}$ for $\lambda_1/c_1 > \lambda_2/c_2$.

As an application of this theorem and Corollary 3.1, suppose subpopulation #1 has twice the incidence rate of the defect, per unit population, as does subpopulation #2. Then, if testing costs are the same, $r_2^* < r_1^* < \sqrt{2}\, r_2^*$, provided the disutility function satisfies the hypotheses of the theorem.

An analogous proof to that of Theorem 4 would show that if $D'(\cdot)$ exists and is decreasing (hence $D(\cdot)$ is concave increasing), then $r_2^*/r_1^* < \sqrt{\lambda_2 c_1/\lambda_1 c_2}$ for $\lambda_1/c_1 > \lambda_2/c_2$. One technical difficulty does arise, however. If $D'(\cdot)$ is decreasing, it is not necessarily true that $\lim_{t\to\infty} D(t) = \infty$, and hence Theorem 3 cannot be utilized. It is possible to get around this problem by adopting the hypotheses of Theorem 3'. This is summarized by writing

Theorem 4': If $D'(\cdot)$ exists and is decreasing and if the optimal solution $(r_1^*, \ldots, r_n^*)$ to (4.5) exists and has positive components, then

$$r_2^*/r_1^* < \sqrt{\lambda_2 c_1/\lambda_1 c_2} \quad \text{for } \lambda_1/c_1 > \lambda_2/c_2 .$$

Kirch and Klein obtain essentially this result for $c_1 = c_2$ and F a uniform distribution over an interval $[0, \beta]$. If $0 < r_j^* < \beta$ for $j = 1, \ldots, Q$, then their $F(\cdot)$ may be regarded as differentiable and increasing for purposes of our proofs.

5. Empirical Estimation of $D(\cdot)$

Although data may not be available currently to support the utilization of particular models, that should not deter the development of such models. It is reasonable to expect that in many application areas in the

future much additional data will become available; for example, more data will become available concerning the stochastic pattern of a particular disease's development. Furthermore, a good model will serve as a guide to the kinds of data which should be gathered.

It is also reasonable to expect the future development of improved testing technologies capable of detecting a disorder in a much earlier stage of development than is now the case. Such innovations will make screening programs more attractive and, consequently, make more important the analytical tools to design such programs intelligently.

However, it is possible to provide a simple procedure to estimate empirically the disutility function $D(\cdot)$ under the assumption of perfect detection. Such a procedure is necessary, since upon the detection of a disorder there may be no way of directly determining how long the disorder has been present; although that length of time could not exceed $x = \frac{1}{r}$. It is assumed, however, that at each detection of a defect, the degree of disutility incurred due to that defect can be observed. For example, in the case of medical screening, at the detection of a tumor its degree of development can be noted even if it is not possible to determine exactly when, since the last test, the tumor originated.

In order to allocate optimally a screening budget among differing subpopulations (susceptibility classes), as was discussed in section 3, it is necessary to know the shape of the disutility function. Under the reasonable assumption that $D(\cdot)$ is an increasing continuous function, the function may be derived in the following manner. For a particular population subject to a Poisson defect arrival $\{S^k\}$ with rate $N\lambda$, arbitrarily select a value for x and set up a prototype mass screening program for

the population in which tests are made (using a perfect test) at times 0, x, 2x,... The number of defects detected at the time jx, which are characterized by a disutility less than or equal to y, record and designate by $Q_j(y)$. Then,

$$Q_j(y) = \sum_{k=0}^{\infty} 1_{[0,y]}\left(D(jx - S^k)\right) 1_{[jx-x,jx]}(S^k), \quad j = 1, 2, 3.$$

$Q_j(y)$ is observable for all values of disutility $y > 0$. Further, $Q_i(y)$ and $Q_j(y)$ are independent and identically distributed random variables for $i \neq j$ and $y > 0$. Therefore, by the law of large numbers,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} Q_j(y) = E[Q_1(y)] \qquad \text{for all } y > 0.$$

Hence, the function $y \to E[Q_1(y)]$ may be estimated by collecting a sufficiently large sample of the functions $Q_1, Q_2, \ldots$

Let $H(y) = E[Q_1(y)]$. Once $H(\cdot)$ is estimated in this manner, the following theorem shows how the desired function $D(\cdot)$ may be obtained from $H(\cdot)$.

Theorem 5: If $D(0) = 0$ and $D(\cdot)$ is continuous and increasing on $[0, x]$, then

$$D(t) = H^{-1}(N\lambda t) \qquad \text{for } 0 \leq t \leq x.$$

6. Summary

In section 3, an explicit expression for the expected long-run disutility per unit time given a screening schedule of uniformly spaced test administrations was given. A general multi-population screening program was also formulated.

In the subsequent sections under the assumption of perfect test reliability, several features of optimal screening schedules for a heterogeneous population have been developed. Within the class of periodic schedules, uniform schedules were shown optimal. The Kuhn-Tucker conditions for assigning particular optimal uniform schedules (i.e., particular testing frequencies) to the respective subpopulations were computed. When $\lim_{t \to \infty} D(t) = \infty$, these conditions reduced to a simple form which made possible the following more specific results: (i) conditions (in terms of λ_i and c_i) under which one subpopulation should be tested more frequently than another, (ii) explicit testing schedules over all subpopulations when $D(t) = at^m$, and (iii) bounds on r_i^*/r_j^* when $D(\cdot)$ is an arbitrary convex increasing differentiable function. Finally, a procedure was suggested for the determination of $D(\cdot)$ under perfect test reliability.

Appendix

Lemma A. Let $\{\tau^k\}$ be a renewal process with associated renewal function $G(\cdot)$. And let (X_1^k), (X_2^k) be two sequences of stochastic processes such that

(a) The pair of processes (X_1^k, X_2^k) is stochastically identical to the pair (X_1^1, X_2^1) for $k = 1, 2, \ldots$

(b) (X_1^j, X_2^j) is independent of τ^k for any j, k.

Then

$$E\left[\sum_{k=1}^{\infty} H(X_1^k, X_2^k, \tau^k)\right] = \int_0^{\infty} E\left[H(X_1^1, X_2^1, s)\right] dG(s)$$

for any bounded function H.

Proof:

$$E\sum_{k=1}^{\infty} H(X_1^k, X_2^k, \tau^k) = \sum_{k=1}^{\infty} E\,E[H(X_1^k, X_2^k, \tau^k)|\tau^k]$$

$$= \sum_{k=1}^{\infty} E\,E[H(X_1^1, X_2^1, \tau^k)|\tau^k] .$$

The last equality uses hypotheses (a) and (b). Let

$$f(s) = E[H(X_1^1, X_2^1, \tau^k)|\tau^k = s]$$

$$= E[H(X_1^1, X_2^1, s)|\tau^k = s] = E[H(X_1^1, X_2^1, s)] ,$$

where the last equality utilizes (b). Now, since H is a bounded function, so is f. Hence, we may utilize a result in Çinlar (p. 286) to get

$$E \sum_{k=1}^{\infty} H(X_1^k, X_2^k, \tau^k) = \sum_{k=1}^{\infty} E\, f(\tau^k)$$

$$= \int_0^{\infty} f(s)dG(s) = \int_0^{\infty} E[H(X_1^1, X_2^1, s)]dG(s) .$$

Q.E.D.

Proof of Theorem 1:

$$E[B_{r,\ell,j}] = \sum_{k=1}^{\infty} E[D(\bar{S}^k_{r,\ell} - S^k)1_{j/r}(S^k)]$$

$$= \sum_{k=1}^{\infty} E\Big[E[D(\bar{S}^k_{r,\ell} - S^k)|S^k]1_{j/r}(S^k)\Big] . \qquad (A.1)$$

Now, since $\bar{S}^k_{r,\ell} \in \{1/r, 2/r, \ldots\}$,

$$E[D(\bar{S}^k_{r,\ell} - S^k)|S^k]$$

$$= \sum_{n=1}^{\infty} D(\tfrac{n}{r} - S^k)P(\bar{S}^k_{r,\ell} = n/r|S^k) .$$

And by (3.3)

$$P(\bar{S}^k_{r,\ell} = n/r|S^k)$$

$$= P\Big(\bigcap_{i=1}^{n-1} \{Y^k_\ell(\tfrac{i}{r} - S^k) = 0\} \cap \{Y^k_\ell(\tfrac{n}{r} - S^k) = 1\}\Big)$$

$$= p_\ell(\tfrac{n}{r} - S^k) \prod_{i=1}^{n-1} [1 - p_\ell(\tfrac{i}{r} - S^k)]$$

by (3.2) and the independence of $Y^k_\ell(s)$ and $Y^k_\ell(t)$ for $s \neq t$.

Combining the last two equations gives

$$E[D(\bar{S}^k_{r,\ell} - S^k) | S^k] = \sum_{n=1}^{\infty} D(\tfrac{n}{r} - S^k) p_\ell(\tfrac{n}{r} - S^k) \prod_{i=1}^{n-1} [1 - p_\ell(\tfrac{i}{r} - S^k)] .$$

This equation inserted into (A.1) yields

$$E[B_{r,\ell,j}] = E\left[\sum_{k=1}^{\infty} \sum_{n=1}^{\infty} D(\tfrac{n}{r} - S^k) p_\ell(\tfrac{n}{r} - S^k) \prod_{i=1}^{n-1} [1 - p_\ell(\tfrac{i}{r} - S^k)] 1_{j/r}(S^k) \right]$$

which by lemma A

$$= N\lambda \int_0^{\infty} \sum_{n=1}^{\infty} D(\tfrac{n}{r} - s) p_\ell(\tfrac{n}{r} - s) \prod_{i=1}^{n-1} [1 - p_\ell(\tfrac{i}{r} - s)] 1_{j/r}(s) ds$$

$$= N\lambda \int_{j/r}^{(j+1)/r} \sum_{n=1}^{\infty} D(\tfrac{n}{r} - s) p_\ell(\tfrac{n}{r} - s) \prod_{i=1}^{n-1} [1 - p_\ell(\tfrac{i}{r} - s)] ds$$

letting $u = \frac{n}{r} - s$

$$= N\lambda \sum_{n=1}^{\infty} \int_{\frac{n-j-1}{r}}^{\frac{n-j}{r}} D(u) p_\ell(u) \prod_{i=1}^{n-1} [1 - p_\ell(u - \tfrac{n-i}{r})] du$$

letting $m = n - i$

$$= N\lambda \sum_{n=1}^{\infty} \int_{\frac{n-j-1}{r}}^{\frac{n-j}{r}} D(u) p_\ell(u) \prod_{m=1}^{n-1} [1 - p_\ell(u - \tfrac{m}{r})] du$$

which, since $p(u) = 0$ for $u < 0$

$$= N\lambda \sum_{n=j+1}^{\infty} \int_{\frac{n-j-1}{r}}^{\frac{n-j}{r}} D(u) p_\ell(u) \prod_{m=1}^{n-1} [1 - p_\ell(u - \tfrac{m}{r})] du$$

letting $\nu = n - j$

$$= N\lambda \sum_{\nu=1}^{\infty} \int_{\frac{\nu-1}{r}}^{\frac{\nu}{r}} D(u)p_{\ell}(u) \prod_{m=1}^{\nu+j-1} [1 - p_{\ell}(u - \tfrac{m}{r})]du$$

$$= N\lambda \sum_{\nu=1}^{\infty} \int_{\frac{\nu-1}{r}}^{\frac{\nu}{r}} D(u)p_{\ell}(u) \prod_{m=1}^{\nu-1} [1 - p_{\ell}(u - \tfrac{m}{r})]du .$$

The last equality follows since for $m \geq \nu$ and for $(\nu - 1)/r \leq u < \nu/r$, $m/r \geq \nu/r > u$. But then $1 - p_{\ell}(u - m/r) = 1$ for $m \geq \nu$ and u within the region of integration.

Q.E.D.

Proof of Lemma 1: Since $D(\cdot)$ is increasing

$$\gamma \int_0^a D(s)ds = \int_0^{\gamma a} D(t/\gamma)dt > \int_0^{\gamma a} D(t)dt \quad \text{for } 0 < \gamma < 1 .$$

Now, for $0 < \alpha < 1$ and $\beta = 1 - \alpha$,

$$\alpha \int_0^x D(t)dt + \beta \int_0^y D(t)dt > \int_0^{\alpha x} D(t)dt + \int_0^{\beta y} D(t)dt .$$

Q.E.D.

Proof of Lemma 2: For an arbitrarily large number M, it is necessary to show that there exists a number $\delta > 0$ such that $f(r) \geq M$ for all $r \leq \delta$.

Since $\lim_{s \to \infty} D(s) = \infty$, there exists a number $1/2\delta$ such that $D(s) \geq 2M$ for $s \geq 1/2\delta$. Then

$$f(\delta) = \delta \int_0^{\frac{1}{\delta}} D(s)ds \geq \delta \int_{\frac{1}{2\delta}}^{\frac{1}{\delta}} D(s)ds \geq \delta \left(\frac{1}{2\delta}\right) 2M = M .$$

Since $f(\cdot)$ is a decreasing function, $f(r) \geq f(\delta) \geq M$ for $r \leq \delta$.

Q.E.D.

Proof of Theorem 3: Upon taking the closure of the feasible set in (4.5) and extending the objective function so as to preserve continuity, the following program is obtained:

$$\begin{aligned} &\underset{(r_1, r_2, \ldots, r_Q)}{\text{Minimize}} && \sum_{j=1}^{Q} N_j \lambda_j f(r_j) \\ &\text{s.t.} && \sum_{j=1}^{Q} N_j c_j r_j \leq b \\ & && r_j \geq 0 \quad \text{for } j = 1, \ldots, Q . \end{aligned}$$

Since the objective function is continuous and the feasible region is compact, the minimum $r^* = (r_1^*, \ldots, r_Q^*)$ exists. By the strict convexity of the objective function (Lemma 2), r^* is unique.

The linearity of the constraints and the differentiability and convexity of the objective function imply that the Kuhn-Tucker conditions are both necessary and sufficient conditions for r^*. The Kuhn-Tucker conditions are

$$\begin{aligned} &N_j \lambda_j f'(r_j^*) + u_1 N_j c_j - u_{j+1} = 0 \quad \text{for } j = 1, \ldots, Q \\ &\text{and} \qquad\qquad (A.2) \\ &u_{j+1} r_j^* = 0 \quad \text{for } j = 1, \ldots, Q . \end{aligned}$$

Suppose

$$r_j^* > 0 \quad \text{for } j = 1, \ldots, Q. \tag{A.3}$$

Then $u_{j+1} = 0$ for $j = 1, \ldots, Q$ and the theorem will follow from the first equation of (A.2). Of course, (A.3) is also necessary for feasibility with respect to (4.5). Hence, it only remains to verify (A.3).

If for some i, $r_i^* = 0$, then

$$\sum_{j=1}^{Q} N_j \lambda_j f(r_j^*) \geq N_i \lambda_i f(r_i^* = 0) = \infty$$

where the equality follows from Lemma 2. But, it is easy to exhibit a feasible point which maps to a finite value of the objective function; for example, $(r_1, \ldots, r_Q)$ where $r_j = b \Big/ \sum_{i=1}^{Q} N_i c_i$ for $j = 1, \ldots, Q$. This contradicts the optimality of r^*, hence establishing (A.3) and the theorem.

Q.E.D.

Proof of Corollary 3.1: From Theorem 3

$$f'(r_i^*)/f'(r_j^*) = (\lambda_j c_i)/(\lambda_j/c_j) = (\lambda_j/c_j)/(\lambda_i/c_i) < 1.$$

Multiplying each side by $f'(r_j^*)$, which is a negative number, $f'(r_i^*) > f'(r_j^*)$. The corollary now follows from (4.7) which shows $f'(\cdot)$ to be an increasing function.

Q.E.D.

Proof of Corollary 3.2: Since $\lim_{t \to \infty} D(t) = \infty$, Theorem 3 provides that for all $j = 1, \ldots, Q$

$$\lambda_j f'(r_j^*)/c_j = \text{constant}. \tag{A.4}$$

Now compute

$$f'(r) = \frac{d}{dr}\left[r \int_0^{\frac{1}{r}} at^m dt \right] = \frac{d}{dr}\left[\frac{a}{(m+1)r^m} \right] = \frac{-ma}{(m+1)r^{m+1}} .$$

Therefore, using (A.4)

$$1 = [\lambda_j f'(r_j^*)/c_j]/[\lambda_1 f'(r_1^*)/c_1] = \frac{\lambda_j c_1 r_1^{*m+1}}{\lambda_1 c_j r_j^{*m+1}} , \quad j = 1, \ldots, Q .$$

Hence,

$$r_j^* = \sqrt[m+1]{\frac{\lambda_j c_1}{\lambda_1 c_j}}\ r_1^* , \quad j = 1, \ldots, Q .$$

As mentioned above, the budget constraint in (4.5) is binding; therefore

$$b = \sum_{j=1}^{Q} N_j c_j r_j = r_1^* \sum_{j=1}^{Q} N_j c_j \sqrt[(m+1)]{\frac{\lambda_j c_1}{\lambda_1 c_j}} = r_1^* \sum_{j=1}^{Q} N_j \sqrt[(m+1)]{\frac{\lambda_j c_j^m c_1}{\lambda_1}}$$

which yields

$$r_1^* = b \Big/ \sum_{j=1}^{Q} N_j \sqrt[(m+1)]{\frac{\lambda_j c_j^m c_1}{\lambda_1}} .$$

Q.E.D.

<u>Proof of Theorem 4</u>: Since $D(\cdot)$ is convex increasing, $\lim_{t \to \infty} D(t) = \infty$.

Therefore, by Theorem 3

$$\lambda_1 f'(r_1^*)/c_1 = \lambda_2 f'(r_2^*)/c_2$$

and so

$$f'(r_1^*)/f'(r_2^*) = \lambda_2 c_1/\lambda_1 c_2 .$$

Hence, to prove the theorem, it suffices to show

$$(r_2^*/r_1^*)^2 > f'(r_1^*)/f'(r_2^*)$$

that is (recall $f' < 0$),

$$r_2^{*2} f'(r_2^*) < r_1^{*2} f'(r_1^*) . \tag{A.5}$$

Now, $\lambda_1/c_1 > \lambda_2/c_2$ implies by Corollary 3.1 that $r_1^* > r_2^*$. So (A.5), and thereby the theorem, will be established if $r^2 f'(r)$ is shown to be an increasing function; i.e., $\frac{d}{dr} r^2 f'(r) > 0$.

$$\frac{d}{dr} r^2 f'(r) = \frac{d}{dr}\left[r^2 \left(\int_0^{\frac{1}{r}} D(s)ds - \frac{1}{r} D(\tfrac{1}{r}) \right) \right] = \frac{d}{dr}\left[r^2 \int_0^{\frac{1}{r}} D(s)ds - rD(\tfrac{1}{r}) \right]$$

$$= 2r \int_0^{\frac{1}{r}} D(s)ds - D(\tfrac{1}{r}) - D(\tfrac{1}{r}) + \frac{1}{r} D'(\tfrac{1}{r})$$

letting $x = \frac{1}{r}$

$$= \frac{2}{x} \int_0^x D(s)ds - 2D(x) + xD'(x)$$

$$= \frac{2}{x}\left[xD(x) - \int_0^x sD'(s)ds \right] - 2D(x) + xD'(x)$$

$$= -\frac{2}{x} \int_0^x sD'(s)ds + xD'(x)$$

$$> -\frac{2}{x} \int_0^x sD'(x)ds + xD'(x)$$

since $x \geq s$

$$= -(2D'(x)/x)(x^2/2) + xD'(x) = 0 .$$

Q.E.D.

Proof of Theorem 5: Let $D(x) = d$. Since $D(0) = 0$ and $D(\cdot)$ is continuous and increasing on $[0, x]$, then $D^{-1}(\cdot)$ exists and is continuous and increasing on $[0, d]$ with $D^{-1}(0) = 0$ and $D^{-1}(d) = x$.

$$H(y) = E[Q_1(y)] = E\Big[\sum_{k=1}^{\infty} 1_{[0,y]}\Big(D(x - S^k)\Big)1_{[0,x]}(S^k)\Big]$$

by Lemma A

$$= N\lambda\int_0^x 1_{[0,y]}\Big(D(x - t)\Big)dt = N\lambda\int_0^x 1_{[0,y]}\Big(D(s)\Big)ds$$

$$= N\lambda\int_0^x 1_{[0,D^{-1}(y)]}(s)ds = N\lambda\Big(x, D^{-1}(y)\Big)^-$$

$$= N\lambda D^{-1}(y) \qquad \text{for } 0 \le y \le d .$$

Therefore,

$$D^{-1}(y) = H(y)/N\lambda \qquad \text{for } y \in [0, d] .$$

Applying $D(\cdot)$ to each side,

$$D\Big(H(y)/N\lambda\Big) = y \qquad \text{for } y \in [0, d] . \tag{A.6}$$

Let

$$t = H(y)/N\lambda \qquad \text{for } y \in [0, d] . \tag{A.7}$$

Then $t \in [0, x]$ and

$$y = H^{-1}(N\lambda t) \qquad \text{for } t \in [0, x] . \tag{A.8}$$

Inserting (A.7) and (A.8) into (A.6),

$$D(t) = y = H^{-1}(N\lambda t) \qquad \text{for } t \in [0, x] .$$

Q.E.D.

REFERENCES

Barlow, R., L. Hunter, and F. Proschan, "Optimum Checking Procedures," J. Soc. Ind. Appl. Math., 11, 4 (1963), pp. 1078-1095.

Blumberg, M. S., "Evaluating Health Screening Procedures," Operations Research, 5 (1957), p. 351.

Çinlar, E., Introduction to Stochastic Processes, Englewood Cliffs, New Jersey: Prentice-Hall, 1975.

Derman, C., "On Minimax Surveillance Schedules," Naval Research Logistics Quarterly, 8 (1961), pp. 415-419.

Flagle, C. D., "A Decision Theoretical Comparison of Three Procedures of Screening for a Single Disease," Proceedings of the Fifth Berkeley Symposium (1967), p. 887.

Hutchison, G. and S. Shapiro, "Lead Time Gained by Diagnostic Screening for Breast Cancer," Journal of the National Cancer Institute, 41 (1968), p. 665.

Keller, J., "Optimum Checking Schedules for Systems Subject to Random Failure," Management Science, 21 (1974), pp. 256, 260.

Kirch, R. and M. Klein, "Surveillance Schedules for Medical Examinations," Management Science, 20 (1974), pp. 1403, 1409.

_____, "Examination Schedules for Breast Cancer," Cancer, 33 (1974), pp. 1444, 1450.

Lincoln, T. and G. H. Weiss, "A Statistical Evaluation of Recurrent Medical Examinations," Operations Research, 12 (1964), pp. 187, 205.

McCall, J., "Maintenance Policies for Stochastically Failing Equipment: A Survey," Management Science, 11 (1965), pp. 493-524.

_____, "Preventive Medicine Policies," Rand Corporation, P-3368-1, 1969.

Neveu, J., Mathematical Foundations of the Calculus of Probability, San Francisco: Holden-Day, 1965.

Pierskalla, W. P. and J. A. Voelker, "A Survey of Maintenance Models: The Control and Surveillance of Deteriorating Systems," Naval Research Logistics Quarterly, 23, 3 (1976), pp. 353-388.

Prorok, P., "On the Theory of Periodic Screening for the Early Detection of Disease," Unpublished Ph.D. dissertation, University of New York at Buffalo, 1973.

Prorok, P., "The Theory of Periodic Screening, I: Lead Time and Proportion Detected," Advances in Applied Probability, 8 (March, 1976), pp. 127-143.

_____, "The Theory of Periodic Screening, II: Doubly Bounded Recurrence Times and Mean Lead Time and Detection Probability Estimation," Advances in Applied Probability, 8 (September, 1976), pp. 460-476.

Roeloffs, R., "Minimax Surveillance Schedules with Partial Information," Naval Research Logistics Quarterly, 10 (1963), pp. 307-322.

_____, "Minimax Surveillance Schedules for Replacement Units," Naval Research Logistics Quarterly, 14 (1967), pp. 461-471.

Shwartz, M. and H. Galliher, "Analysis of Serial Screening in an Asymptomatic Individual to Detect Breast Cancer," Technical Report, Department of Industrial and Operations Engineering, College of Engineering, University of Michigan, 1975.

Voelker, J. A., "Contributions to the Theory of Mass Screening," Ph.D. dissertation, Northwestern University, 1976.

Weiss, G. and M. Zelen, "A Semi-Markov Model for Clinical Trials," Journal of Applied Probability, 2 (1965), pp. 269-285.

Zelen, M. and M. Feinleib, "On the Theory of Screening for Chronic Diseases," Biometrika, 56 (1969), p. 601.

Zelen, M., "Problems in the Early Detection of Disease and the Finding of Fault," presented at the meeting of the International Statistics Institute, Washington, D. C., 1971.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report Number 3	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Model for Optimal Mass Screening and the Case of Perfect Test Reliability		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William P. Pierskalla John A. Voelker		8. CONTRACT OR GRANT NUMBER(s) N00014-67-A-0356-0030 N00014-75-C-0797
9. PERFORMING ORGANIZATION NAME AND ADDRESS Dept. of Industrial Engineering and Management Sciences Northwestern University Evanston, Illinois 60201		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-322
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE December 31, 1976 13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Maintainability, Inspection, Mass Screening, Preventive Maintenance, Surveillance of Equipment

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper considers a population (composed, for example, of people, machines, or livestock) subject to a randomly occurring defect or disease. If there exist testing procedures capable of detecting the defect before it would otherwise become known, and if such early detection provides benefit, the periodic administration of such a test procedure to the members of the population, i.e., a mass screening program, may be advisable. An analytical model of a mass screening program is developed and analyzed. In the development of the next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

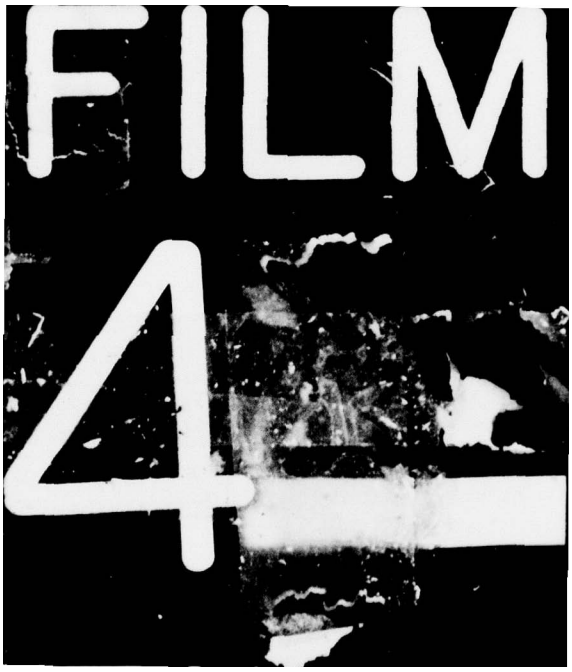
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20.

cont

model, the defect arrival process is Poisson and the ability of the test to recognize a defect is not, in general, assumed perfect.

The model focuses on the time since incidence of a defect as the factor which determines the reliability of a test administration and the value of a detection. The objective is to minimize an arbitrary increasing function of the detection delay which represents the disutility from the occurrence of the defect and its detection. This disutility is also represented as a function of the type of test applied, the probability of success of the test, the testing intervals and the arrival rates of the defects over Q subpopulations. A relatively simple expression is derived for the objective function and a comprehensive mathematical program is presented. The case of perfect test reliability is then considered and, from the class of "cyclic" schedules, the optimum schedules are shown to be equally spaced. For a polynomial disutility function, properties of the optimal schedules are presented. Finally, a method for determining the disutility function from experimental data is suggested.


FILM
4