Mathematical Models for Desiging Optimal Medical Screening Programs

by

S.O. Duffuaa
Department of Industrial and Operations Engineering
University of Michigan
Ann Arbor MI 48109

Mathematical Models for Designing Optimal Medical Screening Programs

by

S.O. Duffuaa
Department of Industrial and Operations Engineering
University of Michigan
Ann Arbor, MI 48109
Tel: (313) 764-9410
Fax: (313) 764-3451
E-mail: SALIH.ENGIN.UMICH.EDU

Mathematical Models for Designing Optimal Medical Screening Programs

by

S.O. Duffuaa*
Department of Industrial and Operations Engineering
University of Michigan
Ann Arbor, MI 48109
E-mail: SALIH.ENGIN.UMICH.EDU

In this paper the concept of repeat inspections is borrowed from industrial quality control and used for designing medical screening programs. First a simple model is developed for determining the required number of repeated tests to minimize the effect of false negative results. Then a cost minimizing model is formulated to find the optimal number of repeated testing to minimize the expected total cost resulting from repeated testing, false positive and false negative results. A procedure is outlined to compute the optimal value of n. Then the models are demonstrated using the case of cervical cancer. Also extensions of the cost minimizing model is proposed for the design of an optimal screening medical program over the lifetime of a patient.

Keywords: Medical Screening Programs, Repeat Testing, Optimal, False Negative Result.

* On Sabbatical leave from King Fahd University of Petroleum and Minerals, Dhahran, 31261, Saudi Arabia

1

The main technique for secondary prevention is screening [1]. Screening is the application of a test to detect a potential disease or condition in a person who has no sign or symptoms of the disease or condition. The main purpose of screening is to detect a disease early in its natural history when treatment might be effective and relatively inexpensive. Many screening programs have been reported in the literature. For example, screening for hypertension, breast cancer, etc. A host of screening programs and their effectiveness and impact on society are given in [2]. The problem of screening is complicated by many factors. These factors include: age and sex of patients, incidence rate of disease, behavior of patients with respect to diagnosis, sensitivity of tests, cost of diagnosing, cost of early therapy and tests false positive and false negative results.

The most difficult part of the screening problem is its dynamic nature. Over time the probability of having the disease changes and that affect the ability of the test in detecting the disease. A disease is usually more difficult to detect at early stages and that may require the use of more accurate tests at early stages of the disease or the need to repeat the test several times prior to concluding that a disorder does not exist. Also the severity of diseases changes with time and that has an effect on the financial and nonfinancial outcomes, tending to be worse as the disorder (disease) progresses. Therefore its desirable to be able to detect the disease as early as possible.

Despite the complexity of the screening problem several screening programs are available with specific recommendations from such organizations as the American Cancer Society, the National Cancer Institute, the American College of Obstetricians, etc. [2]. Also models for evaluating screening programs have been developed by Eddy [4], Rosener and Polk[5] and Moskowitz et al. [6]. All these models evaluate an existing program and do not address directly the design of such programs. Also the model in [4] does not allow for the possibility of random false negative results.

In this paper the concept of repeat inspection is borrowed from industrial quality control and used in designing medical screening programs. First a simple model is developed for determining the required number of repeated tests in order to minimize the effect of false negative results. Then a cost minimization model is formulated to find the optimal number of repeated tests that minimizes the expected total cost resulting from repeated testing, false positive and false negative results. A procedure is outlined for computing the optimal n. Then the models are demonstrated using the case of cervical cancer. Extensions of the cost minimization model to design an optimal screening program over the lifetime of the patient is given.

The rest of the paper is organized as follows: First the simple model is presented, followed by the cost minimization model. Then both models are illustrated using the case of cervical cancer. Finally, extension of the cost minimization model for designing a comprehensive medical program over the lifetime of the patient is given. Then the paper is ended by summary and conclusions.

A Probability Minimizing Model

The model is developed for determining the number of repeated tests each time a patient visits for screening. The repetitions of the test is expected to minimize the probability of false negative results. When repeating the test, they are performed in a randomized manner in order to reduce dependency between repetitions.

The screening problem for which this model is developed can be stated as follows: A person arrives at the test site at time t with a probability $P(t)$ of having the disease or condition. Usually $P(t)$ is a nondecreasing function of time. A test is usually performed to detect the disease or condition. The tests are imperfect and they have certain probabilities of committing false negative and false positive results. Once the disease is detected it is interrupted and treated to try to prevent patients health. It

3

has been reported in [7] that the probability of committing false negative results is usually higher than false positive. The probability of false negative results is more than thirty times the probability of false positive in tests for cervical cancer [7]. A need exists to minimize this high level of false negative results. The main objective of the model is to minimize the effect of false negative results and that is achieved by minimizing the probability that the patient has the disease after tested.

An analogy exists between medical screening and industrial inspection. In industry items are inspected prior to shipping them to customers in order to make sure they meet specification or are nondefective. Inspectors are imperfect ant they commit type I error (classifying a nondefective item as defective) and type II error (classifying a defective item as nondefective). An item arrives at the inspection station with certain probability of being defective. Several inspection plans have been designed in order to minimize the effect of type I and/or type II errors [8,9]. Also plans exist that achieve certain target of average out going quality level (minimizing the effect of false negative) [10]. The analogy can be drawn if we resemble a screening program by an inspection plan, and the inspection process by the medical testing process.

Next a model is developed that minimizes the effect of false negative results in medical screening. Prior to stating the model the following notation is adopted:

n	=	number of repeated tests per visit,
α	=	probability of classifying a patient without the disease as having the disease. Probability of false positive,
β	=	probability of false negative,
$P(t)$	=	probability the patient is having the disease at time t,
$P(t,n)$	=	probability the patient has the disease after n random repetitions of the test are performed on the patient,

4

C_1 = cost of performing the test once,

C_2 = cost of false positive result,

C_3 = cost of false negative result,

$Œ$ = target value for the probability of having the disease after testing.

Each time the test is performed and the result is negative, the likelihood the patient has the disease reduces.

$$P(t,0) = P(t) \tag{1}$$

Using Bay's theorem, we can obtain the probability the patient has the disease after the test is performed one time,

$$P(t,1) = \frac{P(t)\beta}{P(t)\beta + (1 - P(t))(1 - \alpha)} \tag{2}$$

Applying Bay's theorem n times, the probability the patient has the disease after the test is applied to him n times is given as:

$$P(t,n) = \frac{P(t)\beta^n}{P(t)\beta^n + (1 - P(t))(1 - \alpha)^n} \tag{3}$$

It is clear as n increases $P(t,n)$ decreases. In order to apply this model for determining the value of n that guarantees the target set for $P(t,n)$, the following procedure can be followed:

1. Set $j = 1$, and let the target for false negative be $Œ$.

2. Compute $P(t,j)$ using equation 3.

3. If $P(t,j) \leq Œ$, stop, the value of n which guarantees the required level of false negative results is $n = j$. Otherwise, set $j = j + 1$ and go to 2.

This model will minimize the probability that the tested person has the disease after performing the tests n times and will find the best n that achieves any desired target for $P(t,n)$. In other words it minimizes that a tested person has the disease and therefore minimizes the possibility of false negative results.

5

Cost Minimizing Model

The previous model focuses on minimizing the probability of tested patients have the disease or disorder. In this section a model that minimizes the expected total cost (ETC) is formulated for the medical screening problem stated in the previous section. The expected total cost is the sum of the cost of testing, the cost of false positive results and the cost of false negative results. The cost of testing is easy to quantify, however the other two costs are not as simple as the cost of testing. The cost of false positive usually consists of the cost of extra testing and anxiety caused to the patient and the cost of false negative results consists of the cost of treating the disorder when it is discovered at an advanced stage. The cost of false negative results is very high for diseases such as cancer. If we let C_1 denote the cost of performing the test once, C_2 the cost of false positive and C_3 the cost of false negative. Then a cost minimizing model can be developed to find the optimal number of repeated tests, n that minimizes the expected total cost (ETC). Using the same notation in the previous section, the expected total cost after n randomized repeated tests performed on the patient is given by:

$$ETC(n) = nC_1 + C_2(1 - P(t,n))(1 - (1 - \alpha)^n) + C_3\beta^n P(t,n) \qquad (4)$$

where α and β are the probabilities of false positive and false negative results respectively. The expected total cost under no screening program is given by:

$$ETC(0) = C_3 P(t,0) \qquad (5)$$

which is the expected cost from having the disease and being treated at an advanced stage, i.e. when it becomes apparent. This cost is expected to be high for certain types of diseases such as cancer.

The following procedure can be used to obtain the optimal n for a fixed value of t.

1. Set $j = 0$, compute $ETC(0)$ using equation (5),

6

2. let $j = j + 1$, compute $P(t,j)$, from equation (3),

3. compute ETC(j), using equation (4),

3. If $ETC(j) < ETC(j-1)$, go to step 2, otherwise the optimal $n = j - 1$.

The model developed can be used to determine how many repeated testing is needed per visit to minimize the cost resulting from false negative, false positive and the cost of testing. The procedure outlined above is a scheme for finding the optimal number of repeated tests.

The data needed to run the simple model given in section 2 is $P(t)$ probability the patient has the disease, reliability of the test, i.e., the probability of false negative and false positive results and a target level needed to be accomplished. For the cost minimizing model in addition to $P(t)$, α and β, three types of costs are needed. The cost of performing the test, the cost of false positive and the cost false negative results.

Illustration

The models presented can be used in determining the optimal number of repeated testing each time a patient comes for screening and can also be used for designing a comprehensive medical screening program. The latter use will be elaborated on in the next section. In this section the models are demonstrated for the first use. The case used for the models demonstration is the screening of cervical cancer. The data for this case is obtained from [7]. Annual incidence rate for cervical cancer rises slowly beginning at about the age of 30 to 35 and quickly plateaus at about 20 per 100,000. Incidence rates of carcinoma in situ are higher peaking at 130 per 100,000. The lifetime probability of developing carcinoma in situ was estimated to be about 2%. The main test used to screen cervical cancer is the papanicolaou smear. It has false negative rate of 3% and false positive rate of 0.5%. The charge for the test is $75. The charge for working up a person with a false positive papanicolaou smear is

7

$150. The cost of false negative results until the disease is discovered at a later stage is taken to be much higher than the value given in [7], because that value does not account for the agony and psychological effects. It is taken to be $25,000,000. The data is given in Table 1.

Table 1. Summary of Illustration Case

Incidence of Cervical Cancer	0.0013 over individual life
false positive rate	$\alpha = 0.005$
false negative rate	$\beta = 0.03$
cost of test	$C_1 = \$75$
cost of false positive	$C_2 = \$150$
cost of false negative	$C_3 = \$25,000,000$

If we use the simple model in section 2 and we require that the test must discover 99.975% of patients having the disease. If we apply the procedure for finding n. It turns out at each visit the test must be performed randomly 3 times. The results of the calculation are shown in Table 2.

If we apply the cost minimizing model it turns out the test should be performed once every time the patient is screened. Details of the calculation are shown in Table 3.

Table 2. Results of the Probability Minimizing Model

Number of times the test repeated n	Probability the patient having the disease	Probability correctly identifying patient with disease
0	.0013	—
1	.003	96.981
2	.0009	99.909
3*	.000027	99.997

Table 3. Results of Cost Minimizing Model

Number of times the test is performed n	Expected total cost	Cost of false positive	Cost of false negative
0	325.00	0.00	0.00
1*	105.14	0.75	29.40
2	150.78	0.75	0.03

* optimal n.

Model Extension and Additional Uses

The cost minimizing model can be extended to design a screening program over the lifetime of a patient. Let us assume the screening period over the lifetime of a patient is T/year. Let N be the number of screening visits. Then the patient has to be screened every T/N year. Let n_i be the number of repeated tests at visit i made at time t_i, see Figure 1.

Figure 1. Screening Program Over Patient Lifetime
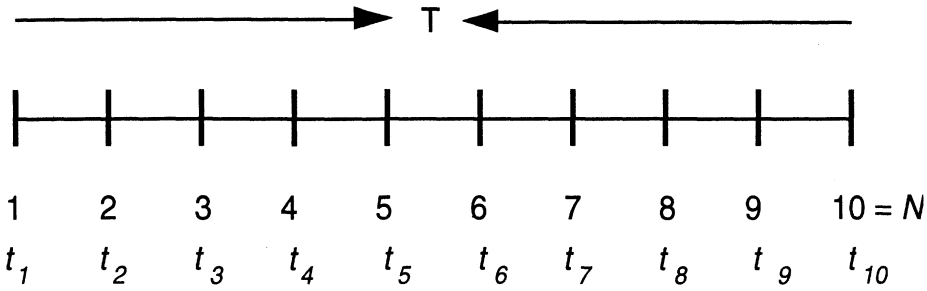
At each time t_i, the probability $P(t_i)$ a person has the disease can be estimated using the techniques in [4]. Then the following model can be used to determine N and n_i.

$$Min\sum_{i=1}^{N}[c_1 n_i + c_2(1 - P(t_i,n_i))(1 - (1 - \alpha)^{n_i}) + c_3 P(t_i,n_i)\beta^{n_i}]$$

Search techniques over N and possible n_i can be employed to determine the number of visits over the lifetime of a patient and the number of repeated tests at each visit.

Summary and Conclusion

In this paper two models are developed to aid in designing medical screening programs. The first model focuses on minimizing the effect of false negative results. It provides the required number of times a test need to be repeated to assure that the effect of false negative result is minimized and also maximize the probability that a person with disease will be discovered. A procedure is outlined for obtaining the value of n.

The second model minimizes the total cost resulting from repeated tests, false positive results and false negative results. A procedure is outlined to obtain the value of n that minimizes the expected total cost. Both models are illustrated using a case of cervical cancer.

In the last section the cost minimizing model is extended for designing medical screening programs over the lifetime of a patient. The model can be used for evaluation of existing screening program by comparing its cost and its false negative rates to the one the model provides. It differs from Eddy's Model for evaluation of medical

screening in the sense that the model in this paper can both be used for design and evaluation while Eddy's Model is an evaluation model. Also the model in the paper incorporates cost directly in the design of the screening program.