BIOMETRIC DESIGN OF THE MAYO LUNG PROJECT FOR EARLY DETECTION AND LOCALIZATION OF BRONCHOGENIC CARCINOMA

WILLIAM F. TAYLOR, PHD, AND ROBERT S. FONTANA, MD

Several important aspects of the Mayo Lung Project demand evaluation. These are: 1. Acceptance. Will people accept such a screening program? 2. Case finding. Does the screen pick out the people most likely to have or develop bronchogenic carcinoma? 3. Effectiveness. If an early case of bronchogenic carcinoma is found, will prompt treatment extend life beyond the time at which death from this disease would have occurred if treatment had been delayed? Direct measurement of effectiveness is not possible, and indirect methods must be used. A group of patients, all of whom are considered suitable for the screening program, are being divided randomly into two subgroups, one to be screened and the other to be kept as an unscreened control. Mortality in the two groups is to be compared for 5 years, and hopefully for 10 years. We also consider here sample size requirements and reports on some of the characteristics of the first 500 patients.

THE MAYO LUNG PROJECT IS A SCREENING program designed to detect new cases of bronchogenic carcinoma. Its goals have to do with:

1. Reducing mortality from bronchogenic carcinoma (through early detection).
2. Identifying persons most likely to develop this disease (by serial sputum cytologic tests).
3. Finding the lesion in its earliest (or in situ) stage.
4. Predicting the clinical course and outcome (by prolonged intense surveillance).

The first goal above was the weightiest in our planning. Can we really reduce lung cancer mortality by use of screening to find early lung cancer cases?

Briefly, the screening comprises chest roentgenography, cytologic examination of sputum, and a report of the patient's symptoms. Screening is followed in suspicious cases by intensive effort to locate and treat the lesion. The program is a prospective one, since the initial screening, while of interest in regard to prevalence, is done mainly to eliminate preexisting cases (both actual and sufficiently suspicious). What we really want is to find and study new cases.

NECESSARY DECISIONS

In designing this study, we faced many problems. How long should the interval between screenings be? What sort of subjects should be used? How can we tell whether the screening is worthwhile? Many of our decisions were arbitrary but practical. We decided to use a 4-month screening interval because previous studies suggested that longer intervals were too long. We thought a 4-month interval would be acceptable to our patients and achievable by our technical personnel. (This is yet to be confirmed.) Ambulatory Mayo Clinic patients were chosen as subjects. This was not our first choice, but negotiations to use industrial workers failed on scientific and practical grounds.

ANTICIPATED EVALUATIONS

Of particular factors: Any screening done to find cancer for early treatment has factors that demand particular evaluation. (There are parallels with tuberculosis screening and with recent decisions about its value.)

Presented to a conference on "Carcinoma of the Lung—A Status Report of the National Cancer Institute's Research," Bethesda, Md., March 27, 28, 1972.

From the Mayo Clinic and Mayo Foundation, Rochester, Minn.

Received for publication July 15, 1972.

1. Acceptance. If the screening program is offered to appropriate persons, will they accept it in sufficient numbers to justify establishing the program permanently?

What proportion accept the screening offer? How many drop out after initial acceptance? What sort of persons drop out? Once evidence of disease is found, will the patient accept treatment?

2. Case-finding. Does the screen pick out the persons most likely to have or to develop bronchogenic carcinoma? How sensitive is the screen? How specific?

Frequent chest roentgenography, with close comparison for changes, is expected to reveal newly existent suspicious lesions. The cytologic examination is expected to identify subjects having cells of varying degrees of abnormality. How are these indicators related to subsequent evaluation and confirmation of bronchogenic carcinoma?

3. Effectiveness. If we find a case of bronchogenic carcinoma earlier than usual, will prompt treatment prolong life beyond the time at which death from this disease would have occurred if treatment had been delayed? This, of course, is the paramount issue. Obviously, we think we can do it, but proof is another matter. The effectiveness of early treatment could be tested directly by finding early cases and treating some promptly but leaving others to be treated later, when the usual, but presumably less sensitive, methods detect the disease. Because this is not possible with human patients, we have turned to indirect methods.

Of overall result: Suppose we offer our screening program (with prompt treatment) to one group of men and do not offer it to a comparable second group. Can we, after a number of years, detect meaningful differences in the lung cancer mortality of the two groups? This evaluation ignores specific reasons for differences and asks only, "Does screening work?" This is the basic question addressed by the Mayo Lung Project.

DESIGN OF PROJECT

Subjects and methods: In the course of usual procedures at the Mayo Clinic, we identify each male patient who is 45 years of age or older and who smokes at least one pack of cigarettes a day. As part of the routine health examination of such patients, a standard 14 by 17-inch posterior-anterior chest roentgenogram is made and studied and a pooled 3-day "deep cough" sputum specimen is examined cytologically. We have the patients answer a Lung-Health Questionnaire as part of this project. All men found free from clinical evidence of lung cancer and free from other serious diseases (to the degree that life expectancy is estimated as at least 5 years) are included in this study. These patients are assigned at random to one of two groups.

1. One group, designated *controls,* receives care and advice of the standard which is current practice at Mayo Clinic. This includes the recommendation of the Clinic's Division of Thoracic Diseases that a chest roentgenogram and a sputum cytology test be obtained at least once a year and that the patient stop smoking. However, these men will be told nothing of the screening program. Rather, they will be examined and will receive care at their own request as if no screening program existed. A routine follow-up communication will be made with each man at least once a year for at least 10 years to determine survival status. When a man dies, his death certificate will be obtained and the circumstances of his death will be determined from his local doctor.

2. The other group, called *participants,* will be treated initially just as the first group, but these men will also be urged to participate in the intensive bronchogenic carcinoma screening project. Men who refuse will not be dropped; they will be followed as closely as possible through correspondence and will be included when comparisons are made with the first group.

Analysis: If the work is carefully done and if adequate time is allotted for the project, a moderate difference in observed lung cancer mortality can be deemed significant statistically and can be attributed to some aspect and effect of the screening procedure. (We may not know which aspect, but at least we will have established that screening and early treatment have some effect, and we will have incentive to pursue the matter further. Such aspects as how intensive the screening should be or how costs can be reduced are perhaps better delayed until the question of gross effectiveness is answered.)

Notice that we will not merely compare survival time of early-discovered and late-discovered cases. There is an unknown bias in favor of early-discovered cases, even if no treatment is employed. Notice also that we do not rely

on volunteers for one group and let the comparison group consist of nonvolunteers. Instead, we divide the group of eligible people at random into two groups, offer the screening to one of them, and then compare the two groups in their entirety. Finally, it should be noted that we do not plan to make a comparison of the incidence of cancer or the survival of cancer patients among the unscreened controls with that of the participants, because to get such detailed information we would have to communicate with the control patients and thus lose part of the difference between control and screened patients. The screened group may have a higher *observed* incidence because we observe them more closely. We want the two groups to be observed with different intensity—within the bounds of currently acceptable medical practice—because this is what the study is all about.

A word about eligibility: An early benefit from this work results from the first screening. The cases of lung cancer found then will be interesting in themselves and will be worked up thoroughly. The initial screening should also eliminate from further study patients who for other reasons are considered to have an unusually short expectation of life. This, of course, will be somewhat subjective, but decisions will be made as consistently as possible, in accord with written guidelines.

Sample size and time required: We have considered sample size in relation to comparison of mortality from bronchogenic carcinoma in the two designated groups. Suppose we admit N men into each group. After 5 years there will have occurred T_1 and T_2 man-years of exposure in each group, and D_1 and D_2 deaths. If $T_1 \cong T_2$, as is likely, we merely must determine whether the control deaths D_1 exceed significantly the screened deaths D_2. It is reasonable to consider the D_1 and D_2 deaths as independent binomial trials. Let p denote the probability that, given a death occurs, it occurs in the controls. Let H_0 be the hypothesis $p = 1/2$, and let H_1 be the alternative of interest, $p = 2/3$. (This corresponds to reducing the lung cancer death rate in the screened group to half that in the controls.) We want the following two conditions to be met.

P (reject H_0 in favor of $H_1 | H_0$ true) = α = 0.05

P (reject H_0 in favor of $H_1 | H_1$ true) = β = 0.95

We reject H_0 in favor of H_1 whenever

$$\left[\left(\frac{D_1}{D_1 + D_2} - \frac{1}{2}\right) \Big/ \sqrt{\frac{1}{4}\frac{1}{D_1 + D_2}}\right] \geq 1.645.$$

The probability that this occurs under H_0 is about 0.05. The probability under H_1 is about 0.95 if $D_1 + D_2 = 90$.

Now the question is—how many men must we examine for how long to get about 90 deaths from bronchogenic carcinoma? (The following information is in the nature of a first attempt to estimate this quantity.) Suppose we wish to get an answer in 5 years, and assume from published data and some educated guessing that 5 deaths per 1,000 man-years will occur among the controls and 2.5 deaths per 1,000 man-years among the participants in the close surveillance. We expect to have 60 deaths among the controls and 30 among the participants if we observe 12,000 man-years in each. These estimates, based on averages, do not take into account chance variation. If we wish to be 95% sure of obtaining 60 and 30 deaths, respectively, we need to observe 16,000 mean-years in each group. We think we can obtain such numbers from our present case load but not without difficulty. Initial plans calling for a total of 6,000 men (3,000 in each group) may have to be modified and will be as soon as deemed essential. We anticipate some losses; there may well be men who refuse to continue under screening. These are not to be entirely lost; their cases will be followed anyway by mail. But it does dilute the difference between the groups and makes the true effects of screening more difficult to detect. The surveillance effort will have to be vigorous and encouraging.

Will 5 years be long enough, even with the numbers of subjects proposed? Perhaps not; but regardless of the early outcome and regardless of whether the actual screening goes on beyond 5 years, these men should continue to be traced for at least a total of 10 years. In our opinion, important information about survival following early treatment will require more than 5 years' study. This opinion is based on possible recurrence of the initial cancer, as well as concern over development of an entirely new primary cancer, particularly in individuals with squamous cell carcinoma.

First Observations

Of obvious interest is the nature of the pa-

tients to be studied in the Mayo Lung Project. Interviews with the first 500 men to appear at the project office as potential study subjects provided the data in Tables 1 and 2. These are a high-risk group as determined by age, amount smoked, duration of smoking, and relevant symptoms.

TABLE 1. Pertinent Characteristics of First 500 Candidates for Mayo Lung Project

Age (yrs)	Per cent	Cigarettes per day	Per cent	Years smoked	Per cent
45–49	26	20–29	43	<20	2
50–59	48	30–39	23	20–29	20
60–69	23	40–49	27	30–39	43
70–74	2	50+	7	40–49	27
				50+	7

COMMENT

Questions and criticisms on a number of points have been received from reviewers and other early readers of this report, and they deserve discussion here.

1. *The selection of 4-month intervals for the survey.* Radiation exposure will indeed be greater for our participants, who will undergo chest roentgenography every 4 months, than it has been in other surveys which used an interval of 6 months. But the amount still presents no particular danger to the participants, and we think it justifiable in this group of heavy smokers.

2. *The possibility of false positives.* Because of the intensive nature of this screening, some false positives are to be expected. All suspicious cases will be studied with exceeding care, however, and unnecessary therapeutic risk will be small indeed. Evaluation of therapeutic risk will be accomplished with the aid of exceptionally thorough medical records.

3. *The choice of Mayo Clinic patients as subjects.* Prople who come to the Mayo Clinic are clearly not representative of the "general population," nor are they a group of healthy industrial workers. They represent, however, a group of people for whom screening, if it works, would be desirable in other institutions involved in care of patients. Our patients may be unusual with respect to their willingness to accept the intensive screening program offered to them. Acceptance by a more general, nonpatient group might be so poor as to indicate that such screening should not be attempted in such a population. We cannot determine that from this project; all we will find here is how well our selected group of patients will accept the inconvenience and expense incidental to this screening procedure. (They will not pay for the actual examinations.)

4. *Analysis of data 4 or 5 years from now.* Since we will have good medical records for each of our participants, we can consider associations between various other health conditions and subsequent occurrence of bronchogenic carcinoma.

5. *The ability of find curable lung cancer by use of serial roentgenograms.* Not much success has been reported in the past. However, in this study, roentgenography is only part of the screening process. It is the combination of x-ray with sputum cytology and a variety of new approaches to the identification and location of cancer lesions which are to be studied very intensively here.

6. Why limit this study to men who smoked at least one pack of cigarettes a day? Why were light smokers excluded? In this study, considerable effort has been spent to find high-risk subjects. Light smokers are not considered, simply because they are not expected to have enough bronchogenic carcinoma to be worth studying in this way. The essence of the design is in the comparability of the participants frequently screened to patients who receive only standard medical care. All are thought to be persons at a fairly high risk of getting bronchogenic carcinoma because of their age, sex, and intensity of cigarette smoking.

TABLE 2. Pertinent Symptoms of First 500 Candidates for Mayo Lung Project

Symptoms	In candidates (per cent)	
	Present	Worsening
Cough		
Chronic	39	12
Producing: sputum	69	9
blood	2	0
Breath: short	11	20
wheezing	36	11
Pain in chest	23	8
Hoarseness	25	5
Swelling of joints	9	4