Periodic health examinations and multiphasic screening

DR. CLAIRE BOMBARDIER:* The belief that the periodic health examination guarantees longer and better life has led to the emergence of various screening programs. This trend was facilitated by the recognition of correlation of risk factors with most of the major diseases. Social pressures other than medical have also played a role in the establishment of these programs: from government to reduce long-range cost; from industry to provide fringe benefits to employees and to gain an outlet for equipment and tests; and from private entrepreneurs to establish computerized check-up clinics.

These programs vary considerably in content, but all are based on the same assumption — that earlier diagnosis will promote efficiency of present therapy. Is this assumption true? It was based mainly on studies done in the late 1950s showing that the periodic health examination did, in fact, "discover" many diseases. In a typical study[1] 50% of the patients were found to be abnormal. Some of the abnormalities were already known, but a large number were discovered with the examination. The abnormalities most commonly found were cardiovascular disease, obesity, tumours, gastrointestinal disorders and metabolic disorders. However, these studies never established that discovering these diseases could change the outcome. The failure to change the outcome may be attributed to the fact that not all these diseases are amenable to medical treatment after they are established, and because treatment policy in latent stages is even more nebulous.

Without conclusive evidence that we can alter the natural history of the diseases detected in a significant proportion of the people screened, the physician must consider whether screening is ethically justifiable.

Furthermore, screening may not be financially justifiable. At the present growth rate, health expenditures in Canada would absorb almost the gross national product by the year 2000,[2] and pressures are increasing to develop some rational method of allocating limited funds.[3] In the health sector there are various methods of evaluating the allocation of resources, economists and physicians tending to use different criteria.[4] The economist's view is reflected through studies such as the cost-benefit analysis illustrated in Table I.[5] The criteria used are the ratios of the monetary benefits to the costs of different programs. The monetary benefits are measured as "avoided earning time lost". This is easily measurable in terms of dollars, but introduces biases; for example, how does one compare the housewife's earning time lost with that of a man employed in industry. The other view is found in the medical literature, where the measure of benefits has been mainly in terms of changes in mortality rates.

Three medical groups have published data which evaluate the benefits from periodic health examinations in terms of change in mortality rates. The first, the Periodic Health Examination Co-operative Research Project,[6] involved eight executive health clinics (all for men) and a 15-year follow-up of patients. In their report comparing the ratio of the actual death rates of the screened people with the expected death rates, they concluded that the data were compatible with a favourable influence of the periodic health examination, but that they did not allow conclusion of such an effect. The second group, the Kaiser-Permanente Group, is a young study started in 1964 which has a good control group. At the present stage its results are still not conclusive. The third group is the Commission on Chronic Illness;[7] this is a retrospective study in which the mortality rates of a group of people who accepted screening in the 1950s have been compared with the rates of a group who refused screening. For most categories the confidence limits overlapped, and the only category in which the death rate was a little lower in the screened group was that of white women, aged 40 to 59. The conclusion of this study was that either screening might be effective, or might only reflect the selection bias for screening.

These studies are not conclusive as to the effect of periodic health examinations on mortality rates and are attended by many problems, such as the selection of the proper control group

Table I—Cost-benefit analysis of different government programs*

	Ratio: benefit/cost
Security belts	1351.4
Education of pedestrians	1117.1
Motorcycle helmets	144.3
Arthritis	42.5
Syphilis	16.7
Tuberculosis	4.4
Medical exam for driver's licence	3.8

*Prepared by the U.S. Department of Health

Held at the Royal Victoria Hospital, Montreal, November 29, 1972

* CLAIRE BOMBARDIER, M.D.: formerly Senior Assistant Resident and Clinical Scholar, McGill University; currently Chief Resident, Royal Victoria Hospital
** JACQUELINE MCCLARAN, M.D.: Assistant Resident and Clinical Scholar, McGill University
***DAVID L. SACKETT, M.D., M.S. (Epidemiology): Chairman, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario

Reprint requests to: Clinical Scholars Programme, c/o Dr. W. D. Dauphinee, Royal Victoria Hospital, 687 Pine Ave. W., Montreal 112, P.Q.

and the simultaneous evaluation of many variables. Another approach that has been used is to identify and examine the selected screened variables commonly applied by doctors in the periodic health examination.[8]

DR. JACQUELINE MCCLARAN:** Interviews were conducted with 18 doctors in private general practice in Montreal to find out what tests are most usually performed when an asymptomatic patient requests a check-up. Their answers revealed that the following are almost always done: measurement of pulse and blood pressure, examination of the optic fundi, eardrums, mouth, chest, heart, abdomen, extremities, reflexes and, for women, breast and pelvic examinations. Lymph nodes and search for vascular bruits were less uniformly included in the examination, and motor or sensory deficits were not sought unless they were indicated by other findings. Cervical cytology, urinalysis, chest x-ray and hemogram were usually considered routine, and for men over 40 an ECG and a serum cholesterol determination were also usually included. This was not necessarily the case with women in this age group. Many of the physicians interviewed were more likely to order an SMA-12 than a uric acid determination, since ordering estimation of a single parameter was not as economical as ordering a whole battery of tests. It is interesting that in the early days of medicare, tonometry was always performed. Now tonometry is rarely part of a routine check-up. A cervical culture never is routine.

Having identified the variables that are commonly looked at in a periodic health examination, we will now direct the following fundamental questions to Dr. Sackett: Is the periodic health examination able to detect disease likely to have an important impact on health? Will the treatment of risk factors have a major impact on the subsequent development of disease? What are the prospects that the behaviour of patients as it affects their health can be altered? Does the periodic health examination really alter the outcome of disease? Are we misled by traditional methods used in evaluating the clinical effect of early detection programs? Have we considered the entire range of the possible effects of early diagnosis and long-term therapy?

DR. DAVID SACKETT:*** I would like to make three introductory comments before addressing the questions raised today. First, general practitioners in an area near Hamilton have recently successfully petitioned the school board to remove the periodic health examination as a prerequisite for high school admission, indicating that there is some regional disparity in the views of general practitioners concerning the usefulness of these procedures among some age groups.

Second, in my opinion there are at least four reasons for doing screening or periodic health examinations, only one of which will be discussed today. The first could be described as a gamble in which you "bet" a life insurance company that you are going to die and they "bet" you that you are going to live. They screen you, not in an effort to explain anything about your health, nor necessarily to improve it, but in order to win their "bet". The second reason is in order to protect people other than the patient. An example is the screening of potential crane operators in steel mills for seizure disorders and cardiovascular disease. Here again, we are not necessarily interested in the individual patient; we are interested primarily in protecting those around him. The third reason relates to the use of screening for clinical baselines. And the fourth, the reason for screening that I want to discuss, relates to its use for the purpose of reaching an early diagnosis under the assumption that a disorder, if diagnosed in an early (or preferably asymptomatic) state, will result in an improved clinical outcome.

The only other comment I want to make is that I do perform a certain amount of screening of a highly "prescriptive" nature, so that although I may be an iconoclast, I don't believe that I am a nihilist!

If we consider the first question: "Is the periodic health examination able to detect disease likely to have an important impact on health?", we find surprisingly little information on this topic. However, one group has looked at the experience of approximately 10 major industrial periodic health examination programs in North America and obtained the results that are shown in Table II.[9] This analysis determined the proportion of individuals dying from specific disorders who had these disorders diagnosed as a result of participation in a periodic health examination program. Less than half the individuals subsequently dying of cancer had this disorder so diagnosed and slightly less than two thirds of individuals dying of coronary heart disease were identified prior to the point at which they developed lethal coronary disease. If we could assume that early detection of these diseases could lead to improvement in the outcome, such findings might be encouraging; however, we have to note at the outset that the periodic health examination has a relatively low sensitivity for the detection of major disorders which have lethal outcomes.[10]

Table II—Does the periodic health examination detect diseases likely to have an important effect upon health?

Cause of death	% of those dying from this cause in whom the diagnosis was made at a periodic health examination
Cancer	43
Coronary heart disease	58
All diseases causing death	51

If we turn to the second question: "Will the treatment of risk factors have a major impact on the subsequent development of disease?", I would point out something which is, I suspect, obvious to clinicians but which is usually ignored when we begin thinking about prevention. If we examine the Framingham Heart Epidemiology Study, a 25-year follow-up of several thousand individuals who were initially free of coronary heart disease, we discover that men destined to develop manifest coronary heart disease in this project exhibited serum cholesterol levels averaging only 245 mg./100 ml., very little above the average of 222 mg. among control men and below the level at which clinicians would institute therapy.[11] Similarly, most victims of coronary attacks do not have clinically abnormal levels of blood pressure, triglycerides, uric acid, or other risk factors; the numbers of victims with abnormal values for these coronary risk factors, despite their higher attack rates, are relatively few in number. When one subjects the gamut of coronary risk factors or predictors to this type of analysis and then recalls that the treatment of abnormal levels for the most prominent of these, blood pressure, does not appear to lower coronary risk, it must be acknowledged that the treatment of risk factors is not likely to have a profound impact upon the underlying burden of disability and untimely death.[12-14]

If we look briefly at the third question: "What are the prospects that the health behaviour of patients can be altered?", we might begin by asking ourselves how successful we have been with our own waistlines and our own consumption of cigarettes. When one systematically determines the extent to which patients do follow clinical instructions, the results are equally sobering. One landmark study found that less than one fifth of children complete the full 10-day course of oral penicillin prescribed for group A beta-hemolytic streptococcal sore throat,[15] while studies of several North American

cities have shown that well under 20% of individuals who could benefit from antihypertensive medication are taking these drugs.[16] Furthermore, these studies indicate that fewer than 50% of middle-aged men identified by their physicians as being on therapy are actually under adequate therapeutic control. Such findings are common in compliance studies. In summary, we must have far greater assurance that efficacious therapy will be followed, particularly among patients who are asymptomatic, before we can anticipate that the programs of early detection of the related disorders are going to have beneficial effects.

When we consider the fourth question: "Does the periodic health examination really alter disease outcome?", I would refer to a report of the Kaiser-Permanente Clinic. Two years ago this group reported a trial of the periodic health examination which they had performed.[17] Several thousand Kaiser Plan participants were randomly allocated to two groups, one of which received intensive encouragement on a regular, recurring basis; members of the control group received no such encouragement but were permitted to use Kaiser Plan services in the routine fashion. After several years of the program these investigations were unable to determine any favourable health effect of the periodic health examination on women, and only one group of men, between the ages of 45 and 54, showed differences in disability and absenteeism. Furthermore, these differences, while statistically significant, are clinically unimpressive — only 3.9% less disability and 1.3% better attendance at work. The results of this study are quite sobering.

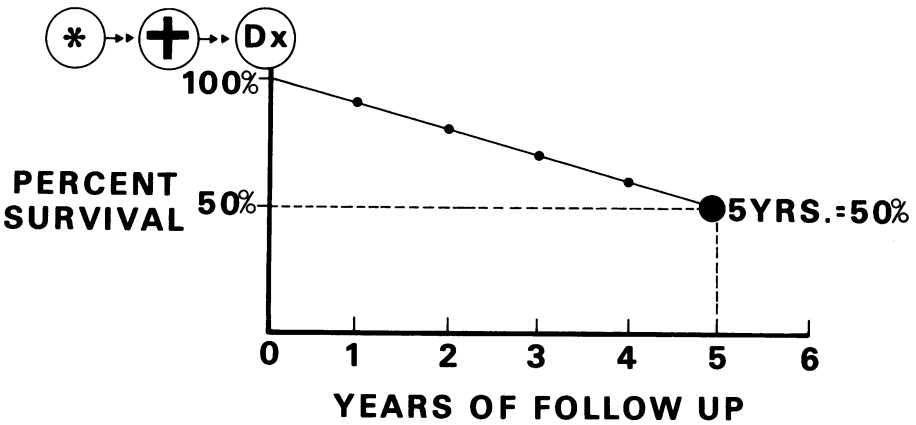

FIG. 1—Typical survival pattern for cancer patients. Dx = time of diagnosis; + = time early diagnosis could be achieved by screening; * = no disease.

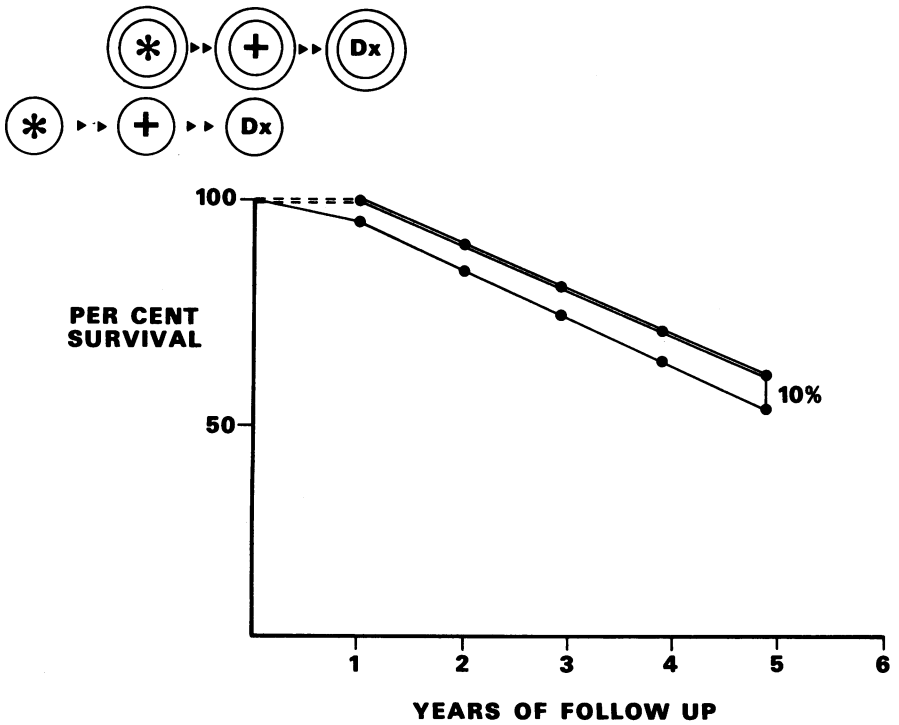

FIG. 2—Survival patterns when patients are diagnosed by screening (double line) and when symptomatic (single line).

In Canada we have a natural experiment which also addresses this issue. A vigorous program of cervical cytologic screening was introduced in British Columbia in the late 1940s, an approach followed much later in the other provinces. Since the "Pap smear" was never validated in a properly randomized clinical trial, it has only been possible to observe the death rate from cancer of the cervix in British Columbia and compare it with that reported from the other parts of Canada in order to determine whether this program has had any impact. The results of the comparison have been discouraging. As of 1965, such comparisons failed to demonstrate an effect of the Pap smear program on mortality from carcinoma of the cervix.[18] The updating of these comparisons by a group in Toronto has also given negative results, and the increasing application of cytology elsewhere in Canada renders future comparisons of this nature invalid.[19]

One may ask a fifth question: "Are we misled by traditional methods used in evaluating the clinical effect of early detection programs?" I suspect that much of the foregoing is in contrast with many of your beliefs about the value of the periodic health examination. Indeed, I am sure that each of you can recall at least one specific patient in whom you achieved an early diagnosis that was followed by what appeared to be a prolonged survival. Although I would not question your clinical judgement in these cases, I suspect that your interpretation of these patients' subsequent survival was affected by one of the following types of pitfalls in the evaluation of clinical outcomes. In Fig. 1, which summarizes a typical survival pattern for cancer, there is a steady decline in survivors which amounts to about 50% at five years, if we select the usual time of diagnosis, Dx, as the starting point for this five-year survival measurement. Thus, of a cohort of 45-year-old patients whose cancer was detected by the usual clinical means, we would expect half of them to be alive at age 50. Let us now assume that early detection techniques can identify this hypothetical carcinoma an average of one year prior to the usual time of clinical diagnosis, that is, the screening of asymptomatic populations could detect this carcinoma one year prior to that point in time at which the appearance of symptoms causes the patient to seek medical care. If we perform the usual types of survival analyses that appear in clinical journals, we will make the kind of mistake shown in Fig. 2. Although this figure assumes that the therapy for this cancer is no more effective when

applied early (the double line) than when it is applied at the time of usual diagnosis (the single line), we note that the five-year survival among the early diagnosed group is substantially better than that of the former group who were not diagnosed until they developed symptoms. This again, however, is entirely misleading, for all we have done is to shift the starting point for the five-year survival measurement one year backward, from the usual time of diagnosis, Dx, to the point at which early diagnosis could be achieved,+. Our group of 45-year-old patients referred to earlier would simply have the diagnosis made one year earlier, at age 44. Only 50% would be alive at age 50, as before. We would not have given them an extra year forward of life; we would have given them an extra year backward of disease! Selection of an inappropriate starting point for measuring survival, then, is one kind of mistake often made in looking at the survival rates of individuals whose disease is diagnosed at an earlier stage. Their increased survival rate is guaranteed, even if the therapy instituted does nothing at all to control or reverse the natural history of their disease.

The second common error in analysing the effectiveness of periodic health examination programs arises out of the relationship between the duration of the preclinical (early or asymptomatic) and clinical (late or symptomatic) stages of disease. Studies of cancer of the breast, stomach and colon have indicated that patients with these cancers who have long preclinical stages tend also to have long clinical stages of disease; conversely, individuals with short preclinical stages for these disorders tend to have relatively short clinical stages.[20,21] This is illustrated in Fig. 3.

Although this relationship probably characterizes most disease, its effect has usually been ignored in analysing programs of early diagnosis. Fig. 4 shows once again that early diagnosis (depicted as the vertical line) will always improve survival, whether or not therapy is effective.[22] This is because the periodic health examination will be more likely to pick out those patients whose disease has a long preclinical stage than those where it is short. As a result, when the disease is diagnosed through screening or periodic health examinations patients are guaranteed longer clinical stages of disease and better short-term survival rates than are patients whose diagnosis is made in the usual fashion, even if the therapy instituted as a result of this early diagnosis has no effect whatsoever.

The final question is: "Have we considered the entire range of the possible effects of early diagnosis and long-term therapy?" It has been suggested that even in the absence of sound evidence these programs are effective, and we simply cannot permit high-risk patients to wait for the results of proper randomized trials. Obviously this suggestion calls for the type of individual clinical decision we have always made about individual patients, whether we are talking about screening or about the use of unproved therapy; the patient simply cannot wait for the treatment to be validated, and we must make our individual clinical decision on the basis of incomplete evidence. When we begin advocating the periodic health examination as public policy, however, such a decision takes on additional dimensions. The individual clinical decision, even if futile, carries with it a relatively low financial cost; if advocated as public policy, however, the cost of periodic examinations is so large that their wider institution must force the reduction, delay or cessation of other programs of clinical care. Furthermore, in both the individual and general case, we must consider the possibility that the intervention, rather than simply being beneficial or useless, may be harmful to health. The magnification of harm through the widespread use of deleterious diagnostic or therapeutic strategies has occasionally had tragic consequences. Many must be aware of the epidemic of asthma deaths that occurred in Great Britain following the introduction of "over-the-counter" bronchodilator aerosols.[23] Furthermore, there is continuing controversy over the usefulness of drugs like tolbutamide and phenformin in the treatment of adult-onset nonketotic diabetes mellitus, as revealed in the randomized trial which demonstrated a death rate in those groups receiving oral hypoglycemic agents twice that of those receiving no therapy other than diet.[24] The prognostic stratification performed prior to randomization in that trial is open to question; the current debate, however, is mainly concerned with whether oral agents have nothing to offer when these patients are on a dietary regimen, or whether they are indeed harmful in terms of cardiovascular mortality.

I think we also have to consider the possibility that the "labelling" of patients as "diseased" may substantially decrease their social, emotional and occupational function. There is, for example, some limited evidence which suggests that both the labelling of an individual as hypertensive and the initiation of antihypertensive treatment may place him at somewhat of a disadvantage in terms of work attendance when compared with other individuals with similar levels of blood pressure who are neither labelled nor treated.

In conclusion, I think that with the exception of prescriptive screening among highly selected groups of patients,[25,26] existing screening and periodic health examination programs are being conducted either in the absence of, or in direct contradiction to, evidence for their clinical effectiveness, and have very little promise of improving, or even maintaining, the health of the general population. Furthermore, I believe that it is essential for groups such as the Clinical Scholars to engage in

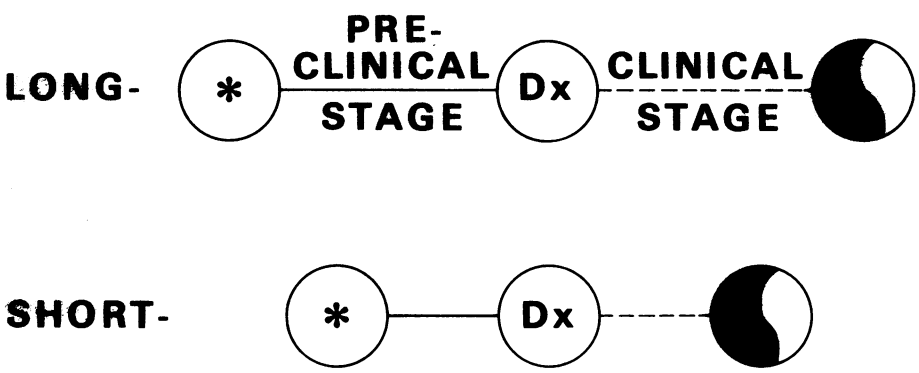

FIG. 3—Correspondence of duration of preclinical and clinical stages.

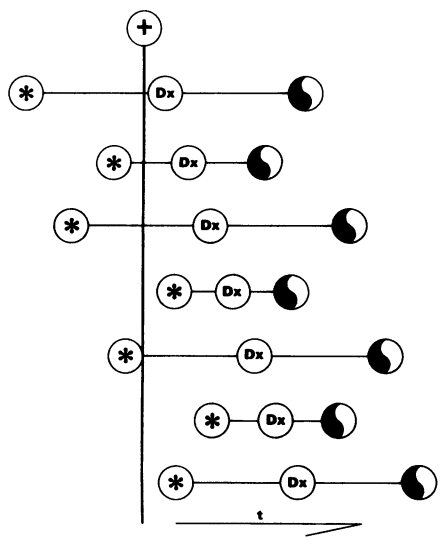

FIG. 4—Improved survival with early diagnosis.

the type of experimental, patient-based research which forms the absolute prerequisite for the demonstration of useful, efficacious and effective programs of early detection and clinical management. Ontario is facing a dilemma which may be somewhat similar to that which will soon have to be faced in Quebec; it has been determined that the proportion of the gross provincial product which will be allocated to health will remain fixed at its present level. On the other hand, Shapiro, Strax and Venet in New York are demonstrating that the use of mammography and clinical examination of the breast can substantially lower the death rate from breast cancer.[27,28] We will then witness an irresistible force, a cancer detection program of demonstrated efficacy, meeting an immovable object, the provincial budget. Unless we, as academic clinicians, rapidly expand our randomized clinical trials of screening, diagnostic and treatment maneuvers so that we can free resources spent on worthless clinical procedures and reinvest them in valid clinical innovations, we will have only ourselves to blame when we are faced with government edicts which restrict hospital beds and physicians' incomes.

References

1. FRANCO SC: The early detection of disease by periodic examination. *Ind Med Surg* 1956
2. Economic Council of Canada, 7th Annual Review: *Patterns of Growth,* chap 4
3. COCHRANE AL: *Effectiveness and Efficiency: Random reflexions on Health Services.* Published for the Nuffield Hospitals Trust by Oxford University Press, London, 1972
4. KLARMAN HE: Present status of cost-benefit analysis in the health field. *Am J Public Health* 57: 1948, 1967
5. GORHAM W: Deux années d'expérience dans l'application du P.P.B.S. ou comment améliorer le partage du gâteau public? *Anal Prev* 5: 403, 1968
6. ROBERTS NJ, IPSEN J, ELMSOM KO, et al: Mortality among males in periodic-health-examination programs. *N Engl J Med* 281: 20, 1969
7. KULLER L: Commission on chronic illness follow-up study. *Arch Environ Health* 21: 656, 1970
8. MCKEOWN T: *Screening in medical care: Reviewing the evidence.* Published for the Nuffield Provincial Hospital Trust by the Oxford University Press, London, 1968
9. SCHOR SS, CLARK TW, PARKURST LW, et al: An evaluation of the periodic health examination. The findings in 350 examinees who died. *Ann Intern Med* 61: 999, 1964
10. BRETT GZ: The value of lung cancer detection by six-monthly chest radiographs. *Thorax* 23: 414, 1968
11. KANNEL WB, DAWBER TR, FRIEDMAN GD, et al: Risk factors in coronary heart disease. An evaluation of several serum lipids as predictors of coronary heart disease: the Framingham study. *Ann Intern Med* 61: 888, 1964
12. Veterans Administration: Effects of treatment on morbidity in hypertension. I. Results in patients with diastolic blood pressures averaging 115 through 129 mm Hg. *JAMA* 202: 1028, 1967
13. Veterans Administration: Effects of treatment on morbidity in hypertension. II. Results in patients with diastolic blood pressures averaging 90 through 114 mm Hg. *JAMA* 213: 1143, 1970
14. CORNFIELD J, MITCHELL S: Selected risk factors in coronary disease. *Arch Environ Health* 19: 382, 1969
15. BERGMAN AB, WERNER RJ: Failure of children to receive penicillin by mouth. *N Engl J Med* 268: 1334, 1963
16. WILBER JA, BARROW JG: Hypertension — a community problem. *Am J Med* 52: 653, 1972
17. RAMCHARAN S, CUTLER JL, FELDMAN R, et al: *Kaiser-Permanente Multiphasic Health Check-up Evaluation Project. II. Disability and chronic disease after seven years of multiphasic health check-ups.* Am Public Health Assoc 99th Ann Meet, Minneapolis, 1971
18. AHLUWALIA HS, DOLL R: Mortality from cancer of the cervix uteri in British Columbia and other parts of Canada. *Br J Prev Soc Med* 22: 161, 1968
19. MORGAN RW: Personal communication, 1972
20. SUTHERLAND R: *Cancer: The Significance of Delay.* London, Butterworth, 1960
21. FEINSTEIN AR: Symptoms as index of biologic behaviour and prognosis in human cancer. *Nature* 190: 241, 1966
22. FEINLEIB M, ZELEN M: Some pitfalls in the evaluation of screening programs. *Arch Environ Health* 19: 412, 1969
23. INMAN WHW, ADELSTEIN AM: Rise and fall of asthma mortality in England and Wales in relation to use of pressurized aerosols. *Lancet* II: 279, 1969
24. University Group Diabetes Program: A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes. *J Am Diabetes Assoc* 19: suppl 2, 1970
25. SACKETT DL: Can screening programs for serious disease really improve health? *Sci Forum* 15: 9, 1970
26. *Idem*: The family physician and the periodic health examination. *Can Fam Phys* 18: 61, 1972
27. SHAPIRO S, STRAX P, VENET L: Periodic breast cancer screening in reducing mortality from breast cancer. *JAMA* 215: 1777, 1971
28. SHAPIRO S: *The effect of screening in breast cancer.* Work-in-progress seminar, McMaster University, Hamilton, Ontario, December 1972