Milbank Memorial Fund Quarterly/*Health and Society, Vol. 56, No. 4, 1978*

The Emergence of the Concept of Screening for Disease

STANLEY JOEL REISER

Program in History of Medicine,
Harvard Medical School

AT THE DAWN OF THE 20TH CENTURY, a movement emerged in the United States that encouraged people to visit doctors periodically to take stock of their health. The movement is with us today. An exploration of the events and arguments that have shaped its identity can help us to understand its promise, and its failings.

An early effort to inspect a seemingly healthy population for evidence of hidden disease was made in the late 19th century on school children, and centered around examination for contagious disease. By the first decade of the 20th century, the school examination was expanded to include a search for general physical impairments. Several surveys found that between one-fifth and one-third of children had vision defects important enough to interfere with their studies, and that about one-half had some disability that required medical attention (Physical Examination of School Children, 1905:587; Medical Inspection of School Children, 1906:878–879).

Some authorities at this time also recommended that all persons would benefit from regular checkups. In 1900, for example, George Gould, in an address before the American Medical Association, proposed a plan for life conservation that included a periodic health examination every 1 or 5 years (Gould, 1900:134).

Another early supporter of preventive checkups was the National Association for the Study and Prevention of Tuberculosis,

0026-3745-78-5604-0403-23/$01.00/0

formed in 1904. It alerted people to the need for detecting early signs of tuberculosis and was joined by a growing number of health authorities in urging families with a tubercular person in their midst to have a precautionary chest X-ray (Precautionary X-Ray Examinations, 1912:560–561). In 1915, the Association set aside a week during which it intensively publicized the virtue of an annual general physical examination (Croft, 1916:220–221). Advocates of the health checkup argued that it made the same good sense as the business practice of making a yearly inventory of stock, or as having one's automobile inspected and tuned up periodically. The widespread use of the car had brought home to many the need for regular inspection of home appliances and machines. Were the heart, lungs, and kidneys any less in need of checkups than the automobile engine?

Industrial leaders joined the chorus of those who applauded the benefits of the checkup. In 1910, an anti-tuberculosis society in Chicago interested a group of employers in sponsoring examinations to detect tuberculosis in workers. Soon the concept was enlarged to general examinations designed to locate any impairment to health. Factory owners saw this idea as a means to improve work efficiency. By 1917, more than 10% of the 300 largest American corporations sponsored regular examinations of employees. Results helped employers to select workers best suited for particular positions and to prevent the spread of contagious diseases within factories. Workers profited not only from learning of hidden physical defects, but also from feeling confident that colleagues would not inflict contagious illness on them, and that employers took a personal interest in their welfare.

Some unions objected that sharing health data with employers infringed on the worker's liberty, might produce blacklists of workers with particular defects for circulation among employers, and result in the personal nature of health data being ignored. Counter arguments were made. No person was forced to submit to an examination, workers ultimately benefited by being assigned to appropriate jobs, and health records were kept confidential (Clark, 1917:239–240). Supporters noted that the average worker generally called on a doctor only when seriously ill. In one estimate, 60% of the workers had no family doctor. Thus, periodic examinations would improve the health and efficiency of the working community. As Clark (1917:244) wrote: "If universally adopted, it would mean a

physically and mentally better country. The sinews of production ever strengthened and guarded, the factory would cease to be considered a consumer of human lives, but would be considered rather an educator and supervisor of health."

The potential for health improvement through periodic examinations led to the founding of the Life Extension Institute in 1914. It furnished checkups through a panel of physicians to private individuals and to institutions such as insurance companies. The Institute made over 250,000 examinations in its first decade. An important by-product of these efforts was providing several thousand general practitioners with an opportunity to perfect themselves in the technique of making preventive evaluations.

The annual checkup was devoted to more than the discovery of undetected disease. A major aim was finding and correcting harmful habits and poor hygiene. The movement's leaders considered the checkup a logical next step in public health, which they thought was at a stage of development where hygiene was more important than environmental control. Sanitary engineering had triumphed over many health-threatening environmental conditions and, broadly speaking, over many communicable diseases. As the toll from communicable diseases declined, disorders such as cancer, heart, and kidney troubles seemed to increase; the weapons needed to combat them were early diagnosis and education on the rules of personal hygiene (Emerson, 1922:402; Tobey, 1923:610–611; Edie, 1925:602–603).

The popularity of the periodic evaluation was significantly furthered by concern over the large numbers of defects being found in apparently healthy people. Examination of army draftees during World War I was crucial, for it became the only real national physical examination the United States had had up to that time. Of the some 3,764,000 males between 18 and 42 examined, about 550,000 were rejected as entirely unfit for service. Of even the approximately 2,700,000 eventually called into service, 47% had physical impairments. The large rejection rate among men supposedly in their prime was considered a dark blot on American health care: medical journals referred to it as "the horrible example" (Tobey, 1923:611). Many causes of rejection were preventable or remediable—foot problems, venereal diseases, tuberculosis, defective vision—and the prospective recruits were often unaware they had such problems.

Doctors and the public expressed willingness to support measures, such as the periodic health examination, to deal with the defects revealed by the army draft. "A physical preparedness [program] is urgently needed in these critical days of reconstruction," wrote one physician (Yount, 1920:18). Their resolve was strengthened by study after study in the post-war period which revealed infirmities in all segments of the population: the rate among 10,000 industrial workers, 100%; 872 immigrants, 97%; a general population of 5000, 77% (Tobey, 1923:611, 648; Emerson, 1922:400–401). Some of these defects were not disabling. Others depended upon defining what "normal" was—a problem that many viewed with concern (Lee, 1923:930–931; Tobey, 1924:875–878).

During the 1920s, industries, the armed services, city and state health departments, and individual doctors encouraged the annual examination. It won the formal endorsement of the American Medical Association in 1922 (Emerson, 1922:399).[1] During that year, the National Health Council voted to designate 3 days for a campaign to induce everyone in the country to visit a doctor for a checkup. The following year, with the slogan, "Have a Health Examination on Your Birthday," the Council launched a public drive with the goal of fostering 10 million examinations (Tobey, 1923:648–649; Edie, 1925:603–605). In New York subway cars an advertisement appeared about this time reading (Emerson, 1922:399): "Your body is a wonderful machine. You own and operate it. You can't buy new lungs and heart when your own are worn out. Let a doctor overhaul you once a year." Evaluating health and teaching ways of keeping it were declared as much the warrant of the practicing doctor as treatment of disease.

Aspects of the periodic examination were like the ordinary general examination. It differed in testing such things as basic sight

[1]The AMA's endorsement said: "Whereas, the value of periodic medical examination of persons supposedly in health is increasingly appreciated by the public, it is recommended by the Council on Health and Public Instruction that the House of Delegates authorize the Council to prepare suitable forms for such examinations and to publish them in the *Journal of the American Medical Association*, and that the county medical societies be encouraged to make public announcement that their members are prepared and ready to conduct such examinations, it being understood that the indigent only shall be examined free of charge and that all others are expected to pay for such examination." (Emerson, 1922:399)

and hearing, and in the attitude of the doctor making the evaluation; the doctor's gaze was directed at appraising health rather than only seeking the causes of physical complaints. Discussion with patients provided not only an interpretation of their physical condition but also an assessment of the relationship of work and habits to health. Patients were helped to understand their condition in the context of how best to live life, rather than given directions limited to the relief of some special symptom. This meant reversing the attitude of doctors toward disease (ibid, 408) in which "its detection and treatment have crowded out of their minds the capacity and habit of studying individuals to learn what are the variations from health from which they suffer and how health may be more complete and better safeguarded."

Change of attitude was not the only problem in getting physicians, mainly generalists, to do such examinations. There were questions of time and technique. An adequate history and physical examination required about 1 hour. The United States in the 1920s had about one doctor per 800 people. To examine the entire population, it was estimated each doctor would have to spend about 2½ hours per day. Given the exigencies of practice, this was too much time. Another question arose: Since the examination involved some techniques usually the province of specialists, could the general practitioner be adequately trained to employ them? Further, experience showed considerable resistance by most people to the concept. For example, in 1914 Metropolitan Life Insurance Company policyholders were offered free periodic examinations. By 1921 only 62,478 had been examined, 2.5% of those invited to take part (Edie, 1925:605).

Still, advocates were optimistic. They believed it possible to convince the public the human machine was the most intricate, delicate, best-constructed, valuable machine they owned, and thus worthy of periodic inspection (Tobey, 1923:649).

By the 1930s, four constituencies interested in general disease-preventive examinations could be discerned: 1) insurance companies, to detect disease in applicants for policies; 2) public services and the military, to maintain a certain standard of physical and mental efficiency; 3) industrial organizations, to evaluate people for employment and promotion; and 4) individuals in the general population, to forestall serious illness (Buck, 1939:883–887).

In addition to the general examination there were special tests to detect highly prevalent diseases. By the 1930s, the two most commonly used tests were the serologic blood test for syphilis, and the chest X-ray for tuberculosis. Both were specific, relatively simple to administer, and thus facilitated a rapid disease detection. Further, the disorders they uncovered could be treated with increasing effectiveness, and the community recognized the importance of controlling them. However, many people were not reached by either physical examination or specific tests.

The 1940s heralded new possibilities. In an address to a distinguished group of British radiologists in 1941, S. C. Shanks observed (Brailsford, 1944:135): "The war has brought into prominence a problem which many of us have been dreaming about, and have regarded as Utopian—that of mass radiology of the chest." Americans, like the British, favored having all of their recruits examined by X-ray, for economic and military reasons. In World War I, routine X-ray examination was not used. By 1940 one investigator projected a loss to the United States government of nearly $1 billion from disability payments to veterans who had entered the army with unrecognized tuberculosis. And the efficiency of the fighting force could be protected by preventing exposure to a debilitating and contagious disease like tuberculosis (Long and Stearns, 1943:149; Marcy, 1944:241–243; Spillman, 1940:1371–1373). An official directive ordering all American recruits to receive a chest X-ray was issued in January, 1941. Studies revealed about 1% of inductees examined by X-ray were disqualified because of active tuberculosis and other diseases detected by the films (Alpern and Benjamin, 1944:548–549; Kinzer, 1945:501–502).

The end of World War II brought new concerns about the state of the nation's health that went far beyond chest disease. Between 1940–46, about 20 million men had been examined, of whom approximately one-third were found mentally, physically, or educationally unfit for service. Of some 3.5 million registrants between 18 and 37 years of age who had been rejected as of May, 1944, the highest number, 16.2%, had mental disease; the next highest, 13.9%, mental deficiency including illiteracy; 7.5%, musculo-skeletal disease; 7.1%, syphilis; and 6.5%, cardiovascular disease (Bauer, 1942:204; Rowntree, 1944:825; Eanes, 1946:573–574). These figures were believed to provide a fair picture of the prevalence of major

defects in the population at large, with allowances made for variations in the examining techniques and the judgments of physicians.

To students of national health problems, these findings were not new. Although data had never been obtained before on so large a number of young men, recent studies had revealed the same basic facts. A 1941 survey (Perrott, 1947:424) of men and women between 16 and 24 years of age, covering 47 states, found that about one-third had a health defect that at least partially restricted choice of work. On the average, at least one defect, such as dental lesions, or eye, heart, and kidney ailments, was found in each person examined. A 1935–1936 National Health Survey (Perrott, 1947:424–425), which covered all age groups, had found 25 million people suffering from some chronic disorder. Both of these studies reiterated the Selective Service conclusions: that the American population was riddled with physical and mental infirmities. Comparison of World War II Selective Service figures with those of World War I, although differences in terminology and classification made this difficult, did not reveal evidence of substantial improvement in the population's health. The death rate had decreased a significant 3.1% in the interval between the wars, with that for men between 20 and 34 years of age, going down nearly 30%. But statistics about the prevalence of defects in the population pointed to weakness in the methods and organization of diagnosis and treatment (Perrott, 1947:426).

The Selective Service data alarmed doctors and laymen. "It comes as a rude shock to learn that the flower of our young manhood is a somewhat substandard bloom," wrote one physician (Johnson, 1941:111). A National Committee on Physical Fitness was created by the President in 1943 to help indoctrinate the public and social institutions on the need for physical conditioning. The periodic examination was pressed generally as a means of uncovering unknown disorders, and the public was urged to seek treatment for defects immediately upon discovery. Many stressed the need to convince industry that fitness increased worker efficiency, and encourage it to make physical examinations available for all workers. Schools also seemed critical places to develop better disease detection programs; their tightly organized structure permitted systematic checkups with minimal effort and people. School health examinations also lay the groundwork for preventing many of the disabilities found in recruits (Eanes, 1946:578–580; Perrott, 1947:427–428).

Many believed that a more comprehensive approach was required, however, reaching not only selected populations but all of society as well. One solution proposed reorganizing medical services around community health centers. Doctors in this setting would combine the functions of health officer and family physician. Diagnostic resources available to all physicians in the area could be pooled, and the health information on each person examined—physical, social, and mental—more readily integrated. The scheme envisioned a regional organization of the hospital system, facilitating the transfer of patients between hospital and health center. Group practices could complement the community centers. An essential element was a method of prepaying the costs of medical care so financial problems would not prevent people from seeking help in early stages of disease (Perrott, 1947:432–435).

Others saw the nation's health problems as, fundamentally, the result of the rapidly changing environment of an increasingly urbanized society. To deal with health problems required elimination of inadequate housing and diets, reduction of illiteracy, abolition of child labor, and the use of government authority to enforce health standards—through compulsory reporting of venereal disease, for example, and physical examinations and Wasserman tests before marriage (Rowntree, 1944:827; Herpel, 1944:350).

Still others believed the crisis called for re-examination of past ideas concerning physical fitness. Measures of native ability had been developed during the 20th century and included tests of concentration, observation, imagination, strength, endurance, dexterity, quick grasp, kinesthetic discrimination, judgment, and intelligence. Capacity for work was seen by some as a mechanical concept, and the human body as a machine for its execution. It thus could be tested in different ways.

Police departments, for example, tested muscular strength by measuring the work done in a given time rotating a wheel by hand, or the force of muscular contraction exerted on a spring. A popular measure of physical fitness was cardiovascular efficiency, which was measured by pulse rate and blood pressure before and after exercise. Psychological aptitude tests also received great attention. But all had flaws in design, such as the inability to evaluate total body capacity by selecting and abstracting an evaluation from only one function, or the difficulty in separating the effect of native ability from training (Kessler, 1940:1591–1595). Advocates had placed too much signif-

icance on psychological and aptitude tests, while human energy and capacity had been underestimated.

Defining normal as "freedom from physical defects" seemed inappropriate: few fit the description. "Freedom from symptoms" was an equally inappropriate definition, since autopsies established that many individuals never experiencing subjective sensations of physical signs of disease had evidence of serious illness. The presence of physical defects was frequently presumed to indicate limitation of work capacity. But studies of the physically handicapped in industry revealed they were as fully productive, and showed no greater loss of time for sickness, accidents, or personal reasons as their nonhandicapped colleagues. Disabled people could not only be ordinarily productive but could excel in a chosen field, as had so many in the arts, sciences, and politics. The organic defect often acted as a stimulus to overcompensation of the whole personality. Physical fitness, in the widest sense, seemed therefore to mean (Kessler, 1940:1594) "the ability to perform productive and continuous work."

As the 1950s began, annual comprehensive health examination for the general population, and more frequently for infants and pregnant women, was the approach to health conservation advocated by many physicians, public and voluntary health associations, life insurance companies, and industries. The examination generally included a medical history, complete physical survey, and selected laboratory and X-ray procedures. Veterans of military service had been indoctrinated with the idea of frequent examinations and were particularly responsive to it as civilians. Industries also expressed growing interest in having corporate executives receive comprehensive checkups (Shillito, 1953:12; Fremont-Smith, 1953:170–173; Breslow, 1959:1148–1149).

The periodic health examination was praised for making participants more conscious of the harmful effects of overweight, bad posture, emotional tension, inadequate diet, lack of exercise, smoking, and drinking. Even if the examination did not uncover dramatic pathology, supporters claimed the small deviations from normal that were found in an estimated 80% of those seen, and the possibility of preventing small defects from leading to major disability, justified the effort (Shillito, 1953:7–9).

It was also seen as a hedge against government control of medical care. As federal and state governments expressed greater in-

terest in early diagnosis and disease prevention, some doctors believed that only a concerted attempt to examine more apparently well people would prevent a takeover (Fremont-Smith, 1953:173). But one of the most valuable results of the examination was its apparent help in establishing greater rapport between doctor and patient (Shillito, 1953:10–11).

Perhaps the closest approximation to these ideals of early detection of pathology, preventive action, patient education about health and illness, and establishing a solid foundation for understanding and care between doctor and patient occurred in pediatrics. This specialty was changing from one largely directed to the diagnosis and treatment of very sick children to one increasingly concerned with health supervision. Greater provision was being made for medical students to follow children over long periods so they might appreciate the many factors influencing their health. Some suggested that pediatricians might be the prototype of future family doctors in their emphasis on preventive medicine, and attention to the family's health (Breslow, 1959:1149).

Despite these accomplishments, a certain restiveness, even disillusion, prevailed in the early 1950s concerning the actual and potential accomplishments of the periodic checkup. Compulsory annual examination of school children seemed grinding to a halt. The Life Extension Institute's activities, initiated early in the century, were discontinued. Some criticized the program of comprehensive study of executives as a luxury that highlighted the difference between preventive medicine for the rich and poor. And evidence grew that relatively few people in the whole population received preventive care. For example, of an average of 5.2 physician visits by Californians in 1954–5, 10% were for specific preventive services and 4% for general checkups (Breslow, 1959:1149–1150). Examination of the doctor's work-week again revealed how difficult providing checkups to most Americans would be.

In the early 1950s, there were about 200,000 physicians serving 175 million people. A 50-hour week and 50-week year provided about 500 million medical man-hours. At least two fifths of this time was devoted to specialty practice that could not be used for periodic examinations. To provide annual checkups for the entire population, at 1 hour an examination, half of the 300 million physician man-hours left would have been required. Unless the number of physicians appreciably increased, this goal was not feasible.

In addition to the limitation of the doctors' hours, many physicians had neither an orientation toward nor interest in health maintenance. Patients seeking such care were often brushed off. "The doctor is apt to say he has no time for this type of health survey," noted the Boston practitioner, Maurice Fremont-Smith (1953:173), who recalled his own attitude during his training in the 1920s at the Massachusetts General Hospital. "What I wanted then were sick people who would get well and be grateful, or cases of advanced disease with good abnormal physical signs for differential diagnosis and teaching purposes. Many physicians still feel the same way about patients with little or nothing apparently the matter with them" (ibid., 170).

Public apathy also contributed to the failure of the annual examination. Most people were not motivated to seek such care considering, among other things, the expense involved. But even for those who were motivated, the examination was frequently of poor quality. Wrote one distraught patient (The Complete Physical Examination, 1957:814–815): "In the last three years, I have gone to two different doctors for *complete* physicals. Neither one of them made any motion to go beyond taking my blood pressure and giving an ear to my heart and lungs . . . Maybe if clinics were set up for checkups for people who think they are in good health, doctors wouldn't be so annoyed at having their precious time taken up by unbroken bones and healthy tissues."

An entirely different approach from that of the physician to the detection of early disease had been used by public health agencies—the mass screening test. They defined as screening the identification of unrecognized disease through the use of rapidly applied tests. Screening sorted out those likely to have the disease from those who did not, but was not intended to be diagnostic. The patient with positive tests would be referred to a physician for further study (Breslow, 1959:1151). The mass X-ray for tuberculosis and serological testing programs for syphilis, the main weapons in early screening programs, were joined in the 1940s by the development of mass techniques to detect diabetes, heart disease, and certain cancers (Cauley, 1950, 631–636; Boucot and Cooper, 1950:1255–1258; Blotner and Marble, 1951:567–575). In contrast to periodic health examinations, screening procedures were capable of wide application, were relatively inexpensive, and required little physician time. The doctor was not needed to administer the procedure, only to interpret it.

Traditionally, screening tests had been used in separate case finding programs; one population group was screened for tuberculosis, another or perhaps the same group at a subsequent time for syphilis or diabetes. As the number of new screening techniques grew, however, some doctors concluded that (Breslow, 1950:275): "Time was ripe for reviewing our categorical approach to case finding and considering the possibility of combining these separate screening methods." "Multiphasic screening" was the term applied to this new approach. It attempted to locate incipient chronic disease rather than, as with most previous single surveys, to focus mainly on communicable disease. Like their predecessors, advocates of the new approach hoped to develop consciousness about the value of preventive medicine in patients and doctors, and forestall premature death and the accompanying economic losses. They sought to fit the screening program into the existing medical care pattern of the community, rather than establish a new superstructure (Chapman, 1949: 1311–1314; Levin and Brightman, 1952:2602). They used tests that were simple to administer, easy to interpret, relatively inexpensive, and needed little professional time (Chapman, 1950:40). The late 1940s and early 1950s thus marked the origin of the notion of multiphasic screening.

For the individual, this approach provided the advantage of a comprehensive examination; for administrative agencies, it saved the expense of several campaigns and multiple record-keeping. The follow-up of positive reactions could be organized on a general rather than categorical basis. The rationale for multiphasic screening was even more apparent when the positive returns from any individual test were considered: in syphilis and tuberculosis, 3 to 5 positives each per 1000 tests; diabetes, 5 to 10 per 1000. Combining such tests thus yielded many more positive cases for each 1000 persons screened. Some workers urged that tests of vision, hearing, weight, blood pressure, hemoglobin, and albumin be added to the screening, as well as a short history form, to provide clues of other disabilities (Smillie, 1951:1254).

The physician's time was also saved. Technicians made the screening examination, physicians reviewed the findings and combined them with other data to formulate a diagnosis. Efforts were made to delimit multiphasic screening from diagnosis: screening produced suggestive, not definitive, findings, and was not a substitute for a visit to a doctor (Breslow, 1950:277–278).

Initial reaction to the concept was mixed. Multiphasic examination was well received by the public and also administrators, who were pleased at its low cost and its rapidity (usually conducted in under 1 hour). The tests were comprehensive, with no special emphasis on certain conditions, and no appeal to "horror" psychology. This helped forestall development of phobias concerning dread disorders such as cancer, tuberculosis, or heart disease. It provided new opportunities for health departments to create public interest in preventive services. It gave new opportunities to the public health nurse, who could follow up people with positive tests by going to their homes, explaining the tests' meaning, and getting them to go to the appropriate clinic (Smillie, 1951:1255).

The multiphasic screening effort uncovered many cases of important unsuspected degenerative disease: an early pilot study conducted in Boston found pathology in more than one-half of those tested, including obesity, impaired vision, hearing problems, hypertension, cancer, heart disease, diabetes, tuberculosis, and syphilis (Getting, 1951:726; Getting and Lombard, 1952:460). Such results stimulated an apathetic and medically unsupervised public to seek medical care (Petrie, 1952:1024). They also demonstrated that the majority of physicians had not met their responsibilities in preventing disease, trained as they were to deal with sick patients rather than presumably healthy people.

But multiphasic screening soon came under attack. In a stinging editorial, the *North Carolina Medical Journal* (1950:356–357) portrayed the effort as introducing into medicine the tendencies of American industry to substitute machines for people in an effort to achieve mass production. The screening process was characterized as a "glorified medical production line." Patients were envisioned entering a line of trailer trucks staffed by technicians who performed various procedures: a serological syphilis test, chest X-ray, electrocardiogram, blood sugar and urine examination. The patients, as the scenario went, then saw a nurse who took their blood pressure, and next a physician who examined them for signs of cancer, and listened to the heart.

The technicians who put the data together had no substantive contact with the "testee" (he could not be called a patient). It was mechanistic medicine, which viewed patients as objects and which lacked an integrator—a skilled observer who did not see the patient as a disconnected series of organs but as a person requiring sym-

pathy and guidance, who must be evaluated in terms of personality and living environment.

Multiphasic screening was also attacked on the ground of omitting significant techniques of evaluation—the history and physical examination. The history, taken by a physician, not one checked on a sheet by the patient, was declared to be the most important factor in diagnosis—providing correct answers 75% of the time. Critics asserted that no method could be devised to take an adequate history of 400–500 patients a day. The public was warned against being deluded that examination by machine could exclude or detect all incipient disease. Regardless of how the mechanical tests were combined, the physician remained the vital diagnostic and therapeutic ingredient in medical care. Stated the *North Carolina Medical Journal* (1950:357): "It seems impossible to convince people that the individual doctor-patient relationship is not merely a sacred fetish handed down from antiquity to which all physicians bow in reverence, but that it is the best method developed in 3000 years for giving good personal medical care."

Planners believed initially that multiphasic screening would consume an insignificant amount of the doctor's time. Unfortunately, they focused on the time needed to interpret the specific test as positive or negative, and not the effort needed to follow up the test. As multiphasic screening expanded, private doctors found they could not thoroughly investigate and treat the many patients who came to them with positive test results (Levin and Brightman, 1952:2602). Furthermore, the tests were generally constructed to detect all pathology that might be diagnostically significant: the normal range was imprecisely delineated from the pathological. As *Lancet* (1967:83) mentioned: "The grey area of uncertainty will probably always be large, and clinical judgment and skill will be especially required in handling patients in this zone." This placed a great burden on the doctor, now required in addition to the normal problems of practice to deal with the anxieties of the false-positive or the marginally pathological finding, and the hollow security of the false negative (Smillie, 1951:1255; Reiser, 1978:191–192). Critics asserted these problems demonstrated that laboratory tests could not replace the diagnostic acumen of the examining doctor (Smillie and Hahn, 1952:2612) and that, if preventive medicine and health maintenance were to be realistic medical and social goals, more physicians and a changed point of view in their education were

needed (Breslow, 1950:1152). Doctors also expressed their concern that the multiphasic screening concept placed a government agency between them and patients. They argued that primary responsibility for disease prevention rested with them and not the health department. They insisted the individual not go first to the health center for tests and then to the doctor: the order must be reversed (Smillie, 1951:1255).

Despite early setbacks to multiphasic screening, the late 1950s saw it slowly accepted. As with many innovations, there was an initial flurry of excitement following an initial introduction, disappointment after flaws were noted, and gradually increased use as basic benefits were recognized. The most comprehensive study of illness in the decade—The President's Commission on the Health Needs of the Nation (1953)—recommended screening as the first step in disease detection. It seemed feasible, given the health resources at that time (12 tests for about $5); screening led to the discovery of much new disease, and thereby fostered initiation of a doctor-patient relationship. The District of Columbia used multiphasic screening as part of a comprehensive approach to the problem of chronic disease, which included history and physical examination. At the same time, federal and state governments gave money to health departments to start multiphasic screening programs, which began to spring up all across the country. Arizona, for example, sent a mobile screening unit to county fairs (Breslow, 1959:1153–1154). Unions were particularly eager to explore the possibilities of screening. They believed it made more sense to develop a system that gave its membership access to care that prevented or reached illness in the bud than to spend money on insurance against the high cost of neglected health (ibid, 1154).

In the 1960s and 1970s, however, a number of serious shortcomings were found in the multiphasic screening concept. It became recognized that screening was valuable only as part of a larger health program that made provisions for follow-up diagnosis and therapy. Physicians, already alarmed at the operation of screening programs by local government as an encroachment on their authority, were also especially critical of an approach that not only dumped large numbers of disease suspects on them, but provided no additional financing or facilities for diagnosis and therapy. The situation was described in 1967 by an official of the Georgia Department of Health (Thorner, 1969:1038): "When we went back into a community that

we had screened a couple of years before, we would find the same old cases of T.B., less the ones that had died, plus new ones. To me, screening is no objective in itself, and just finding cases and dropping them has no value. I am certain that multiphasic screening, as we did it in Georgia 20 years ago, is capable of literally swamping doctors with more cases than we [sic] can handle." Also troubling was the issue of whether persons participating in periodic screening lived longer or had less illness than non-participants. There was little good evidence on either side of the question, although a number of studies had been conducted.

Concern over expending manpower and social resources on screening before adequate research was done to determine its value was increased by the finding of fundamental flaws in the accuracy of the screening procedures (Barnett, Civin, and Schoen, 1970:490). A group of studies in Great Britain sponsored by the Nuffield Provincial Hospitals Trust evaluated 10 of the most widely used screening procedures. Only four survived critical, expert scrutiny. One of the most common, for example, the "Pap" smear screening test for cervical cancer, appeared to have been introduced widely without adequate examination of the complex problems of its application and interpretation (Knox, 1968:43–54).[2]

Another study of 11 screening procedures used routinely by a clinic as part of its periodic health examination found that only five contributed significantly to the diagnosis, and they in only 20% of the cases (Clark, Schor, Elsom et al., 1961:1213).

The requirements for a good screening test, that it be effective and make better use of resources than other alternatives, applied equally to all medical measures. The reason that answers were thought to be more urgently needed in screening was because the position of the doctor in relation to the patient was different than in conventional diagnosis. In the conventional examination, it was the patient who sought assistance, making the doctor's position ethically

[2]The other five screening tests found deficient in some way were the ones for detecting bacteria in the urine of pregnant women, breast cancer, deafness in children, glaucoma, and phenylketonuria (PKU). The four found acceptable were tests for iron deficiency, diabetes, tuberculosis, and Rh disease.

simple. Doctors did their best with the available knowledge and resources; they could not be criticized when the state of knowledge did not allow them to diagnose accurately or treat effectively. The case was quite different in screening. Here the doctor or a public authority took the initiative in investigating the possibility of illness in people who had not complained of symptoms. In this case, an implicit compact existed that the disorder would be not only accurately identified but that those affected would derive benefit from the subsequent therapy. In tuberculosis, for example, where the natural history of the disease was understood, there were good tests with which to recognize the disease in latent and early stages and effective treatment available; the biological criteria for screening tests were thus met. This was less so in detecting cancer of the cervix, however, where its natural history was unclear, the predictive reliability of abnormal cells uncertain, and the long-term effects of the diagnostic biopsy used still to be determined.

Once certain biological criteria were met, there were economic issues to consider. The introduction of an elaborate screening program was at the expense of other medical uses of the resources. Was it worth it? Further, the costs of screening involved not only the health personnel associated with the test itself, but also those involved in the diagnosis and treatment of the diseases discovered.

From the social viewpoint, the principal benefit of the screening program was the more effective treatment of disease caught in its early stages. For the sick individual, it often meant avoiding pain, suffering, and premature death; for the healthy, it gave reassurance but also precipitated anxiety when a false-positive result arose. The psychic costs of pain presented a difficult problem of quantification, and investigators concentrated their attention on loss of production as an index of the cost of suffering, although recognizing the arbitrariness and limitation of this measure as an index (McKeown, 1968:1–13; Pole, 1968:141–158).

Proposals to improve this situation were made in the late 1960s and 1970s. One was the "health hazard appraisal" (Sadusk and Robbins, 1968:1108–1112). While screening sought to identify unrecognized defects, health-hazard appraisal identified risk facts that indicated impending or likely disease, on a probability basis. Physicians would intervene to reduce the risks—for example, by

altering diet, habits, or work patterns. Thus, as physicians developed techniques to learn with increasing accuracy who would become ill, their responsibilities and the problems obtaining and using the information increased.

One of the most interesting changes that occurred in the use of screening in the 1960s and 1970s was from a technique of case finding into an integral part of the ordinary diagnostic and therapeutic visit of the patient to the doctor. This idea was significantly advanced by its elaboration and testing in the Kaiser-Permanente Medical Program. The Program began in 1933, and gradually developed about the notion of providing medical services under a prepayment rather than fee-for-service arrangement through a group of physicians who worked together in the same clinic and hospital facilities. By 1970, it had accumulated over 2 million subscribers, served by outpatient centers, 51 clinics, and 22 hospitals; it provided comprehensive care at an annual cost of $100 per person, employed 2000 doctors, and 13,000 non-physicians (Garfield, 1970:16–17).

Elimination of the fee as the price of entry created a new set of problems. It became clear that it was not only a payment mechanism but a potent regulator of flow into the delivery system. It limited the entry of well and early sick people, but not the late sick person who generally sought help regardless of cost. Accordingly, the fee reduced the total numbers seeking care and created a patient mix of predominantly very ill individuals. With the regulator of flow into the system removed, and nothing in its place, an uncontrolled flood of well, worried-well, early sick, and sick people invaded the facilities on a first-come first-serve basis. The problem called for a new regulator to replace the fee. Automated multiphasic screening was chosen for the task (Garfield, 1970:15–20).

Kaiser had begun using multiphasic screening in 1951 to meet the demand for periodic health examinations. In 1964, it turned to automated techniques and the computer to augment its screening program, and changed it from one devoted to the detection of a few diseases into one that evaluated a broad range of illnesses, separating persons who probably had a disorder from those who probably did not. Thus, as screening became more comprehensive, precise, and quantitative through the new technology, it approximated disease identification, or diagnosis—and automated multiphasic screening approached automated diagnosis (Collen, Rubin, Neyman et al., 1964:741–742). At Kaiser, automated techniques were evolving from

a way of improving the administration of numerous periodic health checkups to serving as the planned entry regulator into the system (Collen, 1966a, 1966b).[3]

Conclusions

The screening concept, filled with potential, has been burdened with problems—many of which remain unsolved. Scientific and technical puzzles abound. For example, the construction of accurate tests that do not produce large numbers of false positives and false negatives challenges our ingenuity and is a key obstacle to the widespread application of screening. Also important is the apathy of the public—on the whole unwilling to participate in health maintenance and early disease detection programs, valuing health mostly in reaction to symptoms that threaten it. Equally significant is the attitude of physicians to the concept of disease prevention. Twentieth-century medical education has not emphasized maintenance of health and prevention of illness as major duties of physicians. Physicians see their principal task as identifying and treating disease in a patient complaining of symptoms. This task, overwhelming as it is, can drain doctors' energy and time and leave little room in their schedules for apparently healthy persons. Part of the solution is developing forms of practice that spread out the burdens of treating disease. But reorganization of practice and delegation of tasks are not enough. The educational institutions of modern medicine must examine the view of illness they teach to medical students and residents, and recall the one described in the second century A.D. by Galen (*Hygiene* 1951:5): "But since, both in importance and in time, health precedes disease, so we ought to consider first how health may be preserved, and then how one may best cure disease."

---

[3]The automated examination combined a detailed computerized medical history with a comprehensive number of psychological tests administered by paramedical personnel, which included tests of the heart, thyroid and respiratory systems, vision, hearing, blood pressure, urine, and a series of 20 blood chemistry measurements plus hematological studies. The chest, and in women the breasts, were X-rayed. All the findings were submitted to a computer that analyzed the data and recommended further tests or an immediate or routine appointment with a physician, depending on the abnormalities uncovered. The entire record was stored in the computer as a health profile, to which reference could be made later.

References

Alpern, A., and Benjamin, J.E. 1944. The Induction Film as an Aid in Appraising Subsequent Pulmonary Lesions. *Radiology* 43 (6): 548–562.

Barnett, R. W., Civin, W. H., and Schoen, I. 1970. Multiphasic Screening by Laboratory Tests—An Overview of the Problem. *American Journal of Clinical Pathology* 54 (3): 483–492.

Bauer, W. W. 1942. An Interpretation of Draft Rejection Figures. *Pennsylvania Medical Journal* 46 (3): 204–207.

Blotner, H., and Marble, A. 1951. Diabetes Control: Detection, Public Education and Community Aspects. *The New England Journal of Medicine* 245 (15): 567–575.

Boucot, K. R., and Cooper, D. A. 1950. A Critical Evaluation of Mass Roentgen Surveys. *Journal of the American Medical Association* 142 (16): 1255–1258.

Brailsford, J. F. 1944. Mass Radiography. *Postgraduate Medical Journal* 20 (222): 135–150.

Breslow, L. 1950. Multiphasic Screening Examinations—An Extension of the Mass Screening Technique. *American Journal of Public Health* 40 (3): 274–278.

———. 1959. Periodic Health Examinations and Multiple Screening. *American Journal of Public Health* 49 (9): 1148–1156.

Buck, R. 1939. The Physical Examination of Groups. *The New England Journal of Medicine* 221 (23): 883–887.

Cauley, J. H. 1950. The Results of the Boston Chest X-Ray Survey. *The New England Journal of Medicine* 243 (17): 631–636.

Chapman, A. L. 1949. The Concept of Multiphasic Screening. *Public Health Reports* 64 (42): 1311–1314.

———. 1950. Multiple Screening for a Variety of Diseases. *Hospitals* 24 (5): 37–40.

Clark, T. W., Schor, S. S., Elsom, K. O. et al. 1961. The Periodic Health Examination: Evaluation of Routine Tests and Procedures. *Annals of Internal Medicine* 54 (6): 1209–1222.

Clark, W. I. 1917. Physical Examination and Medical Supervision of Factory Employees. *Boston Medical and Surgical Journal* 176 (7): 239–244.

Collen, M. F., Rubin, L., Neyman, J. et al. 1964. Automated Multiphasic Screening and Diagnosing. *American Journal of Public Health* 54 (5): 741–750.

———. 1966a. Computer Medicine—Its Application Today and Tomorrow. *Minnesota Medicine* 49 (11): 1705–1707.

———. 1966b. Periodic Health Examinations Using an Automated Multitest Laboratory. *Journal of the American Medical Association* 195 (10): 830–833.

The Complete Physical Examination. 1957. *The New England Journal of Medicine* 256 (17): 814–815.

Croft, B. P., 1916. Benefit of Physical Examinations. *Boston Medical and Surgical Journal* 174 (6): 220–221.

Eanes, R. H. 1946. Need for a Campaign for Physical Fitness. *Occupational Medicine* 2 (6): 572–580.

Edie, E. B. 1925. Health Examinations Past and Present and Their Promotion in Pennsylvania. *American Journal of Public Health* 15 (7): 602–606.

Emerson, H. 1922. The Protection of Health Through Periodic Medical Examinations. *Journal of Michigan Medical Society* 21 (9): 399–403.

Fremont-Smith, M. 1953. Periodic Examination of Supposedly Well Persons. *The New England Journal of Medicine* 248 (5): 170–173.

Galen. *Hygiene.* 1951 (trans. R. M. Green). Springfield, Ill.: Charles C Thomas.

Garfield, S. R. 1970. The Delivery of Medical Care. *Scientific American* 222 (4): 15–23.

Getting, V. A. 1951. Mass Screening or the Multiphasic Clinic. *Journal of the Michigan Medical Society* 50 (7): 726–729.

———, and Lombard, H. 1952. The Mass Screening or Health-Protection Program. *The New England Journal of Medicine* 247 (13): 460–463.

Gould, G. M., 1900. A System of Personal Biologic Examinations: The Condition of Adequate Medical and Scientific Conduct of Life. *Journal of the American Medical Association* 35 (3): 134–137.

Health Inventory—By Machine? 1950. *North Carolina Medical Journal* 11 (7): 356–357.

Herpel, F. K. 1944. Medical Examination at an Armed Forces Induction Station: Observations of a Medical Examining Physician. *The Military Surgeon* 94 (6): 345–350.

Johnson, H. J. 1941. America's Neglected Health. *Proceedings of the Life Extension Examiners* 3 (5): 111–114.

Kessler, H. H. 1940. The Determination of Physical Fitness. *Journal of the American Medical Association* 115 (19): 1591–1595.

Kinzer, R. E. 1945. The Chest X-Ray Examination. *Journal of the American Medical Association* 128 (7): 499–502.

Knox, E. G. 1968. Cervical Cancer. In Nuffield Provincial Hospitals Trust, ed., *Screening in Medical Care: Reviewing the Evidence, A Collection of Essays.* pp. 43–54. London: Oxford University Press.

Lee, R. I. 1923. The Physical Examination of Apparently Healthy Individuals: Its Importance, Limitations and Opportunities. *Boston Medical and Surgical Journal* 188 (24): 929–937.

Levin, M. L., and Brightman, I. J. 1952. The Place of Multiphasic Screening in the Chronic Disease Program. *New York State Journal of Medicine* 52 (21): 2600–2604.

Long, E. R., and Stearns, W. H. 1943. Physical Examination at Induction. *Radiology* 41 (2): 144–150.

Marcy, C. H. 1944. The Problem of Tuberculosis in Apparently Healthy Man as Shown in the Routine Examinations for Induction into the Armed Forces. *Diseases of the Chest* 10 (3): 241–245.

McKeown, T. 1968. Validation of Screening Procedures. In Nuffield Provincial Hospitals Trust, ed., *Screening in Medical Care: Reviewing the Evidence, A Collection of Essays.* pp. 1–13. London: Oxford University Press.

Medical Inspection of School Children. 1906. *British Medical Journal* 1 (April 14): 878–879.

Perrott, G. St.J. 1947. Some Public Health Implications of Selective Service Rejections. *Bulletin of the New York Academy of Medicine* 23 (7): 420–435.

Petrie, L., Bowdoin, C. D., and McLoughlin, J. 1952. Voluntary Multiple Health Tests. *Journal of the American Medical Association* 148 (12): 1022–1024.

Physical Examination of School Children. 1905. *Boston Medical and Surgical Journal* 152 (20): 587.

Pole, J.D. 1968. Economic Aspects of Screening for Disease. In Nuffield Provincial Hospitals Trust, ed., *Screening in Medical Care: Reviewing the Evidence, A Collection of Essays.* pp. 141–158. London: Oxford University Press.

Precautionary X-Ray Examinations. 1912. *Boston Medical and Surgical Journal* 167 (16): 560–561.

President's Commission on the Health Needs of the Nation. 1953. *Building America's Health, A Report to the President.* Vol. 1. Washington, D.C.: U.S. Government Printing Office.

Reiser, S. J. 1978. *Medicine and the Reign of Technology*. New York: Cambridge University Press.

Rowntree, L. G. 1944. National Program for Physical Fitness. *Journal of the American Medical Association* 125 (12): 821–827.

Sadusk, J. F., and Robbins, L. C. 1968. Proposal for Health-Hazard Appraisal in Comprehensive Health Care. *Journal of the American Medical Association* 203 (13): 1108–1112.

Screening and the Practice of Medicine. 1967. *Lancet* 2(7506): 83–84.

Shillito, F. H. 1953. Periodic Health Examinations. *Annals of Internal Medicine* 39 (1): 7–14.

Smillie, W. G. 1951. "Multiphasic" Screening Tests. *Journal of the American Medical Association* 145 (16): 1254–1256.

———, and Hahn, R. G. 1952. Inherent Inadequacies of Multiphasic Screening. *New York State Journal of Medicine* 52 (21): 2610–2613.

Spillman, R. 1940. The Value of Radiography in Detecting Tuberculosis in Recruits. *Journal of the American Medical Association* 115 (16):1371–1379.

Thorner, R. M. 1969. Whither Multiphasic Screening? *The New England Journal of Medicine* 280 (19): 1037–1042.

Tobey, J. A. 1923. The Health Examination Movement. *The Nation's Health* 5 (9): 610–611; 648–649.

———. 1924. A Layman's View of Health Examinations. *Boston Medical and Surgical Journal* 191 (19): 875–878.

Yount, C. E. 1920. The American Youth as Mirrored by the World War: Some Timely Lessons. *Southwestern Medicine* 4 (7): 13–19.

Address correspondence to: Dr. Stanley Joel Reiser, Program in History of Medicine, Harvard Medical School, 10 Shattuck Street, Boston, Massachusetts 02115.