

Screening, the Everlasting Quest for the Needle in the Haystack

Raz Dekel MD

Occupational Health Branch, Israel Defense Forces

Key words: screening, diagnostic tool, health policy

IMAJ 2005;7:650-652

The ability to foresee the future is considered an attribute of the divine. According to the Judaic interpretation, this was the idea behind the “fruit of knowledge” – the apple – that God forbade Adam and Eve from eating. But the serpent proved to be more tempting, promising them: “For God doth know that in the day ye eat thereof, then your eyes shall be opened, and ye shall be as gods, knowing good and evil” (Genesis 3:5). This temptation is universal and timeless; the desire of people to discover their future has existed throughout history, as evidenced by the diverse techniques they have used in its pursuit – oracles, divination, sorcery and medicine. Excavations in Mesopotamia uncovered cuneiform texts in which anatomic findings in a sacrificed animal liver were interpreted to forecast the future, including the future health condition of the person who provided the sacrificial sheep [Figure 1].

Western science is not without those temptations, but the need to justify their use by epidemiologic and biostatistic proof limits their use significantly. Screening can be defined as the application of a test or procedure to asymptomatic, apparently well individuals, in order to separate those with a relatively high probability of having a given disease from those with a relatively low probability of having the disease (performing tests on individuals at high risk would be called diagnosis rather than screening).

The study on electroskin resistance of dermal-visceral zones published in this issue of *IMAJ* [1] represents such an attempt, and a rather daring one, as it seeks scientific support for a technique that has never been used in western medical diagnostics. Yet, screening tests of all kinds present many methodological and policy problems that in many instances cast them in the “twilight zone” of medical science.

One should keep in mind the inherent advantages of screening. The tests are relatively inexpensive, relatively deployable, prevention-oriented, they usually use cheaper, lower level technology, are usually performed by a more generalist personnel, and can be applied rapidly to large groups. However, one must not forget the inherent disadvantages: false negatives (increase in risky behavior), false positives (unnecessary anxiety, workup and treatment), diminishing returns when performed systematically (fewer new cases are diagnosed), increased demands for treatment facilities and increased service expectations (a sort of omnipotent treatment – a panacea), as well as several biases that render analysis more complicated.

We must also not forget that in some cases even a true-pos-



Figure 1. Babylonian clay model of sheep's liver used for divination, 1900-1800 BC.

itive screening test does not end in a clinical illness. Those cases of inconsequential disease often receive excessive treatment and some true positives never even become symptomatic. What happens is that we start labelling people, i.e., calling those with risk factors “diseased.”

A classical study on this question demonstrated the long-term results of screening for hypertension in steelworkers in Canada [2]. Workers diagnosed as hypertensives adopted a “sick” role. This led to a twofold increase in absenteeism, decreased satisfaction with both work and marriage and, in many cases, reduced income and less promotion. This effect also occurred in labelled but untreated workers!

In the 1960s, Wilson and Jungner [3] presented their criteria for population-based screening programs for the World Health Organization:

- Screening for important health problems only
- Screening for problems in which early diagnosis results in better prognosis
- Sufficient availability of facilities for diagnosis and treatment
- Problems in which a latent or early symptomatic stage is detectable
- Presence of a valid screening test
- Screening test is acceptable by target population
- Considerable knowledge about natural course of disease

- Clear decision strategy about further examination or treatment
- Calculation of cost and benefit analysis
- Screening process should not be incidental but continuous.

In other words, diseases that are rare, benign, rapidly evolving, hard to diagnose, hard to treat, undetected throughout latency, obscure to medicine and without a scientifically accepted and clear decision strategy for further examination or treatment, should not be screened, since the chance that the population at large will profit from the screening is small. There seems to be no point in screening for diseases that are very common or very rare. When disease is unlikely to be present before the test (low pre-test probability), only a test with very high specificity can establish its presence (such screening tests are very rare). Otherwise, most positive results will be errors (false positive), which can harm patients either by useless treatment or by "labelling." Conversely, when disease is fairly likely to be present before a test (high pre-test probability), only a test with very high sensitivity can establish its absence (such screening tests are also rare). And, it follows, most negative results will be errors (false negative), which can harm patients not getting the treatment they need. Another conclusion from Wilson and Jungner's criteria is that screening tests that are not scientifically well established (sensitive, specific, etc.), not accepted by the target population, not proven effective in a cost-benefit analysis and not performed on a continuous basis, should not be used since their value is questionable.

But there's another basic problem, by definition, in screening. Even in the rare cases that the test is valid and appropriate, the disease may sometimes not behave as predicted. The famous example is screening for lung cancer. The U.S. Preventive Services Task Force concluded that there is insufficient evidence to recommend or oppose screening asymptomatic persons for lung cancer with low radiation dose computerized tomography, chest X-ray, sputum cytology or a combination of these tests. The Task Force found fair evidence that screening with low radiation dose CT, chest X-ray or sputum cytology can detect lung cancer at an earlier stage than lung cancer would be detected in an unscreened population; however, there was only poor evidence that any screening strategy for lung cancer decreases mortality. Because of the invasive nature of diagnostic testing and the possibility of a high number of false-positive tests in certain populations, there is potential for significant harm from screening [4].

Another example of problematic procedure is screening for colorectal cancer. A study of over 2,600 patients showed extremely low sensitivity (4.9%) of the in-office fecal occult blood test done at the time of rectal examination. In other words, this approach missed 95% of significant abnormalities. This is in marked contrast to the much higher detection rates seen for the recommended 3-card, take-home FOBT. Although the sensitivities (44% for cancers, 24% for all abnormalities) may still ap-

pear somewhat low, it is important to keep in mind that fecal occult blood testing on a one-time basis has never been purported to be of great value [5]. The test derives its benefit from a program of continuous screening over a number of years, which is why the guidelines explicitly state that if FOBT is the method chosen for screening it must be done every year [6]. Even minute changes in screening protocols can render them entirely useless; and if the test requires frequent repetitions, it will reduce population compliance significantly. Furthermore, if follow-up of positive FOBT is inadequate, the whole procedure becomes an exercise in futility. In one study, approximately 30% of patients who were told they had a positive FOBT reported that this test was either followed with a repeat FOBT or no diagnostic workup (positive screening test should be followed by a definitive diagnostic test, not by re-screening or disregard) [7]. Screening tests have meaning only in the sequence of screening→diagnosis→treatment→follow-up. If done in isolation it is usually useless.

As we know, even those rather basic and logical criteria are very hard to follow. For screening to be successful, it requires high quality methodology, a high rate of participation and state-of-the-art treatment for diseases detected. If screening fails to achieve any of these parameters, it will be less effective. If the screening methodology is not of the highest order, then the act of screening is an empty exercise for the population and eventually confidence in the program will decline, not to mention the huge waste of money (another important parameter, but not discussed here). Another key reason for questioning the usefulness of tests is the interpretation of results that might be problematic, as the cutoff point between normal and abnormal values might be elusive even in valid tests and might vary significantly between populations.

The electroskin resistance of dermal-visceral zones study by Zimlichman et al. [1] shows the highest sensitivity for cardiovascular illnesses and high sensitivity for respiratory, gastrointestinal and genitourinary disease. However, for a population in their sixties (as in their study), would the screening results significantly alter the physician's clinical decision-making? Would we rule in any cardiac disease with a test that has 85% sensitivity or rule it out for 52% specificity? Can we apply it to a population with such a large pre-test probability of having a cardiovascular disease (ischemic heart disease, or congestive heart failure, or infectious disease, or hypertension, or vascular pathologies including thrombosis and vasculitis according to the study definitions)?

I believe that until further information is gathered, and more specific and sensitive definitions for diagnosing specific diseases are tested, proven and validated – such tools will merely become gadgets for lavish executive screening programs. In my opinion, the major problem with medical screening is not a lack of tools but a problem of policy. When we screen, instead of implementing a public health system-based approach we rely on a market-based approach. Most screening is opportunistic, i.e., it depends on a coincidence of interest and encounters. The lack of population registers or reminder systems means

FOBT = fecal occult blood testing

that most adults are not screened on a regular basis. This increases the risk of diagnosis of an advanced cancer among men and women who attend screening irregularly. The solution lies in the concept of organized screening, which relies on the systemization of care (institutional policy, population registers, computerization, reminders, chart tools, audits, and centers of excellence) to meet standards of care.

Another realistic solution might be abandoning “pure screening” as an ineffective, wasteful method and deciding that a screening program should begin by classifying the individual patient’s level of risk based on personal, family and medical history to determine the appropriate approach to screening that person – in other words, using the test as a diagnostic rather than a screening tool.

References

1. Zimlichman E, Lahad A, Aron-Maor A, Kanevsky A, Shoenfeld Y. Measurement of electrical skin impedance of dermal-visceral zones as a diagnostic tool for inner organ pathologies: a blinded preliminary evaluation of a new technique. *IMAJ* 2005;7:631–4.
2. Haynes RB, Sackett DL, Taylor DW, Gibson ES, Johnson AL. Increased absenteeism from work after detection and labeling of hypertensive patients. *N Engl J Med* 1978;299(14):741–4.
3. Wilson JM, Jungner YG. Principles and practice of mass screening for disease. *Bol Oficina Sanit Panam* 1968;65(4):281–393.
4. Humphrey LL, Teutsch S, Johnson M. Lung cancer screening with sputum, cytologic examination, chest radiography, and computer tomography: an update for the US Preventive Services Task Force. *Ann Intern Med* 2004;140:740–53.
5. Winawer S, Fletcher R, Rex D, et al. Colorectal cancer screening and surveillance: clinical guidelines and rationale – update based on new evidence. *Gastroenterology* 2003;124:544–60.
6. Collins JF, Lieberman DA, Theodore ED, David JW. Accuracy of screening for fecal occult blood on a single stool sample obtained by digital rectal examination: a comparison with recommended screening practice. *Ann Intern Med* 2005;142:81–5.
7. Nadel MR, Shapiro JA, Clabunde CN, et al. A national survey of primary care physicians’ methods for screening for fecal occult blood. *Ann Intern Med* 2005;142:86–94.

Correspondence: Dr. R. Dekel, 95 Rothschild Boulevard, Tel Aviv 65235, Israel.

Phone: (972-52) 343-3872

Fax: (972-3) 560-8687

email: erasmus@gmail.com