Does Cardiovascular Disease Screening Save Lives in Asymptomatic Adults?

¿El screening de enfermedad cardiovascular salva vidas en adultos asintomáticos?

It is difficult to make predictions, especially about the future.

NIELS BOHR

INTRODUCTION

Screening in the search of asymptomatic disease in adults is a method increasingly used by physicians and most requested by patients, especially for the dreaded cardiovascular disease.

However, we should acknowledge that it is not a method of “primary prevention”, because in reality it consists of actions set in motion by the physician to detect the early stage of a disease either in a person or population. Therefore, it is a “secondary prevention” methodology of health problems, in order to initiate treatment in the asymptomatic period and, in theory, facilitate the cure of the disease.

Genuine “primary prevention”, also known as “health promotion”, involves the actions promoted by a physician or the community to avoid the causes of a health problem in a person or in the population before it appears. (1)

The rationale for screening seems simple, attractive and understandable in itself; it is the early detection of a disease in asymptomatic subjects to enable their treatment, reducing morbidity and mortality and in addition the associated costs.

However, the role of screening is questioned in current heated controversies and by the strong criticism of different groups of interest, including the patients themselves, who overestimate the benefits ruling out the damages. Moreover, on many occasions it cannot be shown that the benefits outdo the damages.

This happened with the United States Preventive Services Task Force (USPSTF) recommendation against the screening of prostate cancer in healthy men, because the screening damage with the prostate specific antigen (PSA) exceeded the benefits, as randomized clinical trials showed no improvement in long-term survival, and screening produced high risk of overdiagnosis with adverse consequences.

Even the popular colonoscopy to detect colon cancer is in discussion. (2)

We will briefly enumerate, according to our understanding, the criteria to decide whether screening should be performed in a medical condition, and we will extensively discuss them in the following sections:

1. To assess the severity of the medical condition in terms of disability or mortality load amplitude caused in the population.
2. Understand the quality of the screening test in terms of sensitivity, specificity and predictive value.
3. Evaluate whether the early medical condition has an effective treatment and advantages over the treatment performed at the moment of clinical presentation.
4. Screening leads the time of detection compared with clinical diagnosis, and deceivingly prolongs the time to death or the emergence in the incidence of morbidity. For example, coronary artery events would decrease (myocardial infarction, acute coronary syndromes, angina, coronary revascularization, etc.).
5. Therefore, it is necessary to develop clinical trials that randomly assign a group to screening and another to standard care. This is essential to avoid the confounding biases of observational trials.

Some screening tests were already firmly enshrined in clinical or public health practices long before randomized controlled clinical trials were widely used, which we currently acknowledge as necessary evidence to admit that the benefits significantly surpass the damages, enabling their massive use in the population.

MAGNITUDE OF MORTALITY FROM CARDIOVASCULAR DISEASE

There is currently no doubt on the severity of the cardiovascular disease (CVD) condition, in terms of disability or mortality load magnitude, since it is the first cause in adult subjects and its prevalence is very high in the population. In the Framingham Heart Study, lifetime risk was estimated in 50-year-old persons free from CVD, up to 95 years of age. The development of CVD in men was 51.7% (95% CI 49.3%-54.2%) and in women 39.2% (95% CI 37.0%-41.4%).

Compared with participants with ≥2 major risk factors, those who have optimal levels at 50 years, which is only 4% of the population, have substantially
less risk throughout their lifetime: 5% vs. 69% in men and 8% vs. 50% in women. (3)

Moreover, the probability that someone who is alive at the age of 35 years dies between 35 and 69 years, mostly from non-transmissible diseases (mainly cardiovascular), is 1/6 in developed countries and doubles in developing countries (1/3). Not only 80% of the population dies in underdeveloped countries, but they die at a much younger age.

This means that if screening worked, it would be relevant for coronary artery disease in our countries.

**SAFETY AND PRECISION OF DIAGNOSTIC TESTS**

Use of a population diagnostic screening test should allow discriminating in persons who are apparently well, those who will probably have an asymptomatic disease from those who will probably not have it.

The quality of a screening test is measured in terms of sensitivity, specificity and predictive value. It must be sufficiently sensitive to detect a great proportion of ill people with few “false negatives”, and in turn it must be sufficiently specific to have few “false positives”, thus increasing the positive predictive value of the test.

But in truth, the search of latent asymptomatic disease usually has very low prevalence of cardiovascular disease, even among those classified as intermediate or high risk, so the positive predictive value for screening tests will be low regardless of their specificity.

Sensitivity and specificity are combined in Bayes theorem to constitute the risk multiplier called the likelihood ratio (LR) (LR=Sensitivity /1-Specificity). (4)

In addition, screening tests should have other characteristics: they should be simple, fast, not cause discomfort, cheap, very safe and acceptable both for the patient and the physician.

Table 1 summarizes the safety and accuracy characteristics of screening tests (sensitivity, specificity and LR), comparing three ischemia evocative tests: exercise treadmill testing (ETT), exercise nuclear imaging (ENI) and exercise stress echocardiography (ESE). (5)

With relatively low prevalence for coronary artery disease, ETT gives false positive tests in 96% of cases with prevalence of 2%, 91% with prevalence of 5%, and 85% with prevalence of 8%. In turn, the improved ENI test has prevalence of 2%, 5% and 8% for false positive tests in 93%, 83% and 75% of cases, respectively.

If the probability before the test is 1%, even with LR as high as 5.7, the probability of having coronary artery disease after the test is 4%, that is 96% of positive cases will be false.

Therefore, in case screening of asymptomatic ischemia is positive, either with ischemia evocative tests or even with conventional imaging tests, it is like looking for a needle in a haystack, since most will be false positive results. Then, why are these tests useful for clinicians? Because when we suspect coronary artery disease due to the symptoms referred by the patient, we already have 40% to 60% probability before the test.

In the face of this prevalence, true positive ETT span from 57% to 75% and false positive tests are reduced between 43% and 25%, and with ENI they are even more reduced from 27% to 13%.

It is useful for us clinicians, because with a probability of 45% before the test and a LR as high as 5.7, if it is positive, 82% will have coronary artery disease and if it is negative (LR 0.19) only 13% will have the disease. In this case a test is reliable, because if positive it indicates high probability of coronary artery disease and if negative high probability of rejecting it. Let us see the recommendations of USPSTF to perform resting and exercise electrocardiogram to detect coronary artery disease in asymptomatic adults.

If the risk of events is low (<10% at 10 years in the Framingham score), the recommendation is to discourage the use of resting and exercise ECG, because the risk balance exceeds the potential benefit.

If the risk of events is intermediate (10% to 20% at 10 years) or high (>20% at 10 years) it makes no recommendation for screening because it indicates there is insufficient information and the balance between risk and benefit provided by screening cannot be established in this population.

**IS THERE AN EFFECTIVE AND EARLY TREATMENT THAT CONFER AN ADVANTAGE OVER TREATMENT AT SYMPTOM ONSET?**

The latest USPSTF review found no randomized controlled trial or prospective cohort study on the effects of screening in asymptomatic adults with ECG during ETT versus non-screening, with clinical endpoints. Nor are there studies of how the identification of high-risk individuals through ETT affects the use of treatment to reduce cardiovascular risk (e.g. statins or aspirin, etc.). (6)

No study estimated ETT correction in the classification of participants into groups of high, intermediate or low risk compared with only assessment of conventional risk factors. A study in men and women

---

**Table 1. Characteristics of test safety and accuracy (5)**

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Likelihood ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETT</td>
<td>46% (33-58)</td>
<td>77% (67-86)</td>
<td>2.00 (1.4-4.1)</td>
</tr>
<tr>
<td>ENI</td>
<td>87% (74-100)</td>
<td>78% (74-81)</td>
<td>3.96 (3.8-5.2)</td>
</tr>
<tr>
<td>ESE</td>
<td>63% (15-100)</td>
<td>87% (33-58)</td>
<td>4.85 (0.5-∞)</td>
</tr>
</tbody>
</table>

published a C statistics of 0.73 for traditional risk factors only assessed by the European score (EuroSCORE) versus 0.76 for the EuroSCORE plus ETT in all-cause death, but it is not known whether this small difference is significant, because the confidence intervals (CI) were not published. (7)

Together, 38 ETT prospective cohort studies evaluated abnormalities with subsequent risk of cardiovascular events. Various abnormalities were associated, such as ST-segment depression with exercise that presented a HR of 2.1 (CI 1.6-2.9), chronotropic incompetence, a HR of 1.4 (CI 1.3-1.9), abnormal heart rate recovery, a HR of 1.5 (CI 1.3-1.9); HR decreased exercise capacity in a range of 1.7 to 3.1 (in which a meta-analysis could not be performed due to the different measurement methods). (6)

In conclusion, certain ETT abnormalities are associated with a slight increase of later cardiovascular events, but the clinical implications of these findings are unclear due to the absence of information from prospective cohort studies or clinical trials comparing screening versus non-screening with development of clinical events.

SCREENING LEADS TIME COMPARED WITH MEDICAL DIAGNOSIS

The decreasing mortality trend of a disease can occur for different reasons, and cannot be mistakenly attributed to current screening programs.

Therefore, it is useful to briefly review the different stages of disease development. (8)

From the beginning (T1) to the time it begins to be detected by screening (T2), there is an unidentified period of the disease. The latency period before diagnosis (T3) is the entire period during which the disease is asymptomatic but detectable by screening, which will vary for each individual, with an average lead time value (t2) which could be of interest, because it would be a measure of the time gained for a potentially effective treatment. (Figure 1)

However, it would be misleading to simply compare the observation between groups with and without screening in the duration of survival, because the probability of events for the group without screening would begin at the time of clinical diagnosis (T3), but in the screening group it would begin much earlier (t2). Therefore, to prolong actual survival and not the fictitious one due to the anticipated screening time, survival prolongation must be greater than the lead time of diagnosis in screening.

But the estimated lead time is rather complex and uncertain, so the only reliable solution is to perform a randomized controlled clinical trial, assigning one group to screening and another to standard control, comparing the development of relevant clinical events. An important event is specific death, but even more definitive is all-cause death, because the addition of competitive or rival causes of death removes the benefit of the reduction in specific mortality.

Many newspaper publications tell us that research on early cancer and cardiovascular disease detection has been fruitful, as demonstrated by the sustained improvement in survival at 5 or 10 years, the most commonly used assessment to communicate progress in the war against cancer and cardiovascular disease.

As already mentioned, it can be shown, with a crude example, that this assessment is misleading. If in the past cancer diagnosis was always done with a palpable tumor, whereas in current patients the diagnosis includes those with microscopic abnormalities in a biopsy, then it would be expected that survival at 5 to 10 years increases by supplementary lead time, even if the new screening strategy were ineffective.

This phenomenon is clearly displayed in the statistical data from the Surveillance, Epidemiology and End Results of the National Cancer Institute in the United States, based on the entire US population from 1950 to 1995. In that period of 45 years, 5-year survival was estimated for 20 types of common solid tumors. Using the tumor as a unit of analysis, it was correlated with two other measurements of cancer burden based on the population; its denominator includes all the population at risk of the disease: mortality (Ncancer deaths/Npopulation), and incidence (Nnew cancer cases/ Npopulation).

From 1950 to 1995, an increase in 5-year survival was observed for each of the 20 types of tumors. The absolute range of increased survival was 3% (pancreatic cancer) to 50% (prostate cancer). Nevertheless, during the same period, the actual mortality in the population declined in 12 cancers, but it incredibly increased in the remaining 8 types. There was little or no correlation between the 5-year survival change for a specific tumor and the change in mortality in the population related to the tumor (Pearson r=0.00, Spearman R=-0.07). Conversely, the change in 5-year survival was positively correlated with the change in the population tumor incidence (Pearson r=+0.49, Spearman R=+0.37).

This clearly shows that increased survival of 20 solid tumors seems primarily to be due to the marked increase in the incidence of each type of cancer, associated with the changing diagnostic patterns imposed by screening. Instead, it has very little association with actual changes in the population mortality decline.
Why are temporal changes in 5-year survival not related with mortality?

As shown in Table 2, (9) there are 3 ways to increase 5-year survival. The first would be that treatment for cancer is actually more effective and with unchanged incidence, mortality will decrease and patients will live longer; this effect would be shown in a typical interventional controlled clinical trial (new drug or surgery).

In both the other forms of increased survival, this would be due to the detection of more patients in the early stage of the disease. Any progress in the time of the disease will increase the 5-year survival due to the spurious effect of lead time, with the consequent increase in disease incidence, with no change in population mortality and without demonstration of effect in a controlled trial.

If early treatment is effective, then there will be an additional increase in 5-year survival, mortality will decline, although less than indicated by the 5-year survival and a randomized trial could prove it.

**Table 2. Three ways to increase 5-year survival**

<table>
<thead>
<tr>
<th>Incid.</th>
<th>Expected change in 5-y survival</th>
<th>Popul.mortal.</th>
<th>RCT mort.</th>
</tr>
</thead>
<tbody>
<tr>
<td>More effective treatment</td>
<td>NC</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Increased survival due to screening</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Is not effective</td>
<td>↑</td>
<td>↑</td>
<td>NC</td>
</tr>
<tr>
<td>Is effective</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
</tr>
</tbody>
</table>

NC: No change. ↑ Increase. ↓ Decrease. ↑ Great increase. * Modified from ref. 9.

It was assumed that the incidence of breast cancer in the United States remained constant over those 30 years, because there was no change in the share of population that generally has not been exposed to screening, women younger than 40 years.

It is shown that the introduction of mammography screening in the United States was associated with more than twofold increase in the number of early stage cases of breast cancer detected each year, from 112 to 234 cases per 100,000 women (an absolute increase of 122 cases per 100,000 women). Concomitantly, the frequency with which women had cancer at the late stage decreased from 102 to 94 cases per 100,000 women (an absolute decrease of 8 cases per 100,000 women).

It should be observed that out of the 122 additional cancers diagnosed at an early stage there are only 8 late-stage cancers less.

Therefore, of all breast cancers diagnosed, an estimated 31% overdiagnosis is considered, more than 70,000 overdiagnoses in 2008 in the United States.

Unfortunately, the study suggests that mammography did not meet the prerequisite to reduce specific mortality (a reduction in the number of women presenting with late stage cancer), because the absolute reduction in deaths (20 deaths per 100,000 women) is greater than the absolute reduction of late-stage cancer cases (8 cases per 100,000 women), confined to cancers with regional invasion, which may now be treated successfully with 85% survival at 5 years. And they have no effect on those who present with distant disease, with a survival rate of only 25% at 5 years.

The good news of the downward tendency of breast cancer, should be largely attributed to improved treatment and not to screening, since the decline in breast cancer among women aged 40 or older was 28% and the concurrent decline among women under 40 was 42%, a higher relative reduction of mortality among women who were not exposed to screening mammography. (10)

This situation which occurs with the use of screening was confirmed for other diseases such as prostate cancer.

How many asymptomatic patients undergoing SPECT screening, who end with coronary angioplasty or sometimes coronary bypass surgery are actually
WHAT IS THE EVIDENCE THAT SCREENING SAVES LIVES IN ASYMPTOMATIC ADULTS?

Scientists at the Stanford Prevention Research Center systematically assessed the evidence from randomized controlled trials (RCT) on whether screening reduces mortality of diseases where death is the common result. Therefore, they focused on the categories of “cancer” and “heart and vascular disease,” as well as type 2 diabetes and chronic obstructive pulmonary disease.

The USPSTF recommended screening for 6 diseases with 12 screening tests.

Evidence from RCT meta-analyses were available for 6 diseases and 9 different tests: abdominal aortic aneurysm (ultrasound), breast cancer (mammography and self-examination), colon cancer (fecal occult blood, flexible sigmoidoscopy), lung cancer (chest X-ray + cytology, CT scan), ovarian cancer (CA-125) and prostate cancer (PSA).

The 95% CI excluded the null hypothesis in 4 of the 11 tests estimated for disease-specific mortality. It was reduced with ultrasound in abdominal aortic aneurysm, with mammography for breast cancer and with fecal occult blood test and flexible sigmoidoscopy in colon cancer, ranging from 16% to 45%.

But it did not diminish in any case all-cause mortality, since all relative risks were very close to 1.0.

Researchers conclude that with the screening tests currently available for diseases where death is a common result, disease-specific mortality reduction is rare and all-cause mortality reduction is unusual or nonexistent.

Why is the decrease in disease-specific mortality not reflected in total mortality? We should acknowledge that because of the many other competing causes of death, it is very difficult to document reductions in all-cause mortality; for example, screening may decrease the risk of death from a ruptured abdominal aorta, but instead the patient may die from myocardial infarction as a concurrent cause, common in the pathology. Total mortality might decrease when the disease of screening interest is the dominant or leading cause of death, or when extremely large RCT are performed.

WHAT HAVE WE LEARNED FOR THE FUTURE?

We will discuss a tremendously current situation. In 2012, the advent of new treatments for hepatitis C virus (HCV) led the US Center for Disease Control and Prevention (CDC) to recommend screening for all people born during the 1945-1965 period, as it is estimated that three quarters of all infected people are in that age cohort. (12)

Subsequently, in April 2014, WHO proposed an expanded screening in the Guidelines for the screening, care, and treatment of persons with hepatitis C, which proposes to perform screening to people at high risk of HCV (injection drug users, HIV infected or incarcerated subjects, children born to HCV infected mother, those who received transfusions before 1992, hemodialysis, sexual partners infected with HCV, etc.). (13)

The widespread screening has been strongly supported by many experts and greeted as an opportunity to save hundreds of thousands of lives around the world.

They base their recommendation on the substantial prevalence of HCV in the world. Around 170 million have HCV antibodies (it identifies those who have been infected by the virus), but since about 30% of people infected with HCV have a strong immune response that cures the infection, it is necessary to perform a second test with HCV RNA to confirm chronic infection; finally it is estimated that, there are at least 120 million people with active infection, and between 350,000 and 500,000 deaths per year. (12, 13)

Now there are highly effective treatments available (ledipasvir and sofosbuvir) which used in combination with another drug (e.g. ribavirin) makes the virus disappear from the serum in 90% of cases, for at least 24 weeks after stopping the 12-week treatment, which has started the discussion about the “cure” of hepatitis C.

Is this high response rate maintained and transferred to long-term clinical benefit? Most of the information comes from old observational studies in those who develop positive responses, but the ability of therapy to reduce the incidence of end-stage liver disease (cirrhosis, hepatocellular carcinoma and death) has not been tested in any RCT.

However, since most people infected with HCV never develop symptoms and will die from other causes, we expose this group to treatment damage without any possible benefit, which should be widely exceeded by the benefit achieved in the minority who will develop end-stage liver disease.

Ribavirin commonly causes anemia and may cause leukopenia, generalized rash, gastrointestinal disorders or insomnia in 10-20% of patients.

Protease inhibitors cause severe anemia and rash, including the potentially fatal Stevens Johnson syndrome. The safety data are scarce with new direct action antiviral drugs. In a sofosbuvir vs. peginterferon plus ribavirin trial, 3% of participants taking sofosbuvir experienced serious adverse effects compared with 1% in the other branch.

The World Health Organization has included sofosbuvir and other medicines for HCV in the List of Essential Medicines, even if the price for a 12-week treatment with sofosbuvir is 84,000 dollars in the United States, sold by Gilead Company, which holds the patent. Recently, on May 9th, 2016, the Indian Patent Office granted the license to the American
company, which prevented the sale of an Indian generic version of the drug available at a retail price of around 500 dollars. (14) The cost of manufacturing sofosbuvir; including a reasonable profit, results in an estimated price of 68 to 136 dollars. (15) Thousands of patients with HCV and public health advocates have already taken to the streets of Madrid to demand “Treatment for All”. (16)

The only solution to this global public health issue is the compulsory license, which is a government decision to allow someone who is not the owner of the patent to produce, sell or buy the registered product without the consent of the patent holder. (14, 17)

Ronald Koretz et al. (12) argue that although an extensive screening for HCV may be a cost-effective strategy to reduce the development of end-stage liver disease, it could result in damage. It is therefore necessary to determine in a well-designed RCT the real benefit of screening. It proposes to include 120,000 participants, in whom liver disease is expected to produce 250-500 deaths. The study would have an excellent power to demonstrate 30% relative risk reduction in the number of deaths from liver disease in the screening branch; the secondary end-point would be all-cause mortality and the composite end-point, liver transplant or death from liver disease. If this difference between groups after 4 years is not observed, due to a low frequency of liver disease, it may be continued for another 2 years.

If treatment for HCV will be scaled to cover the 120-150 million infected people worldwide, regulatory agencies should ensure that the drugs will be evaluated in the long term by clinical endpoints and not with surrogate markers, in several thousands of patients, in whom liver disease is expected to produce 250-500 deaths. The study would have an excellent power to demonstrate 30% relative risk reduction in the number of deaths from liver disease in the screening branch; the secondary end-point would be all-cause mortality and the composite end-point, liver transplant or death from liver disease. If this difference between groups after 4 years is not observed, due to a low frequency of liver disease, it may be continued for another 2 years.

If treatment for HCV will be scaled to cover the 120-150 million infected people worldwide, regulatory agencies should ensure that the drugs will be evaluated in the long term by clinical endpoints and not with surrogate markers, in several thousands of patients, in whom liver disease is expected to produce 250-500 deaths. The study would have an excellent power to demonstrate 30% relative risk reduction in the number of deaths from liver disease in the screening branch; the secondary end-point would be all-cause mortality and the composite end-point, liver transplant or death from liver disease. If this difference between groups after 4 years is not observed, due to a low frequency of liver disease, it may be continued for another 2 years.

If treatment for HCV will be scaled to cover the 120-150 million infected people worldwide, regulatory agencies should ensure that the drugs will be evaluated in the long term by clinical endpoints and not with surrogate markers, in several thousands of patients, in whom liver disease is expected to produce 250-500 deaths. The study would have an excellent power to demonstrate 30% relative risk reduction in the number of deaths from liver disease in the screening branch; the secondary end-point would be all-cause mortality and the composite end-point, liver transplant or death from liver disease. If this difference between groups after 4 years is not observed, due to a low frequency of liver disease, it may be continued for another 2 years.

CONCLUSION
To claim that patients with detected disease in screening will live longer is nothing more than a self-fulfilling prophecy, as they advance the lead time, but this is not an evidence of decreased mortality due to screening.

If the progress of diagnostic screening is not reflected in decreased end-stage disease or death, there is no benefit for the patient and one wonders on the question of the English writer John Milton “... why man has to predict the date of his misfortunes, while these remain unknown, and why go out to meet what he should most avoid?” The only justification for population screening is to have the evidence of its benefit demonstrated by impartial, well-designed RCT, with large numbers of patients and with hard events as mortality.

**REFERENCES**

2. Redberg RF. Fecal blood testing or colonoscopy. What is the best method for colorectal cancer screening? JAMA Intern Med 2016; published online June 15.