Trials of prostate-cancer screening are not worthwhile

Hans-Hermann Dubben

Lancet Oncol 2009; 10: 294–98

University Medical Centre, Department of Primary Medical Care, Hamburg, Germany (H-H Dubben PhD)

Correspondence to:
Dr Hans-Hermann Dubben, University Medical Centre, Department of Primary Medical Care, Martinistrasse 52, D-20246 Hamburg, Germany
dubben@uke.uni-hamburg.de

About 3% of men in developed countries die from prostate cancer. No conclusive evidence, however, either supports or refutes the benefit of prostate-cancer screening. More than 200 000 participants are needed for a screening study with prostate-cancer-specific death as the endpoint. A relative reduction in prostate-cancer mortality of 25% leads to a decrease in absolute risk of less than 1%—a difference of 75 individuals between the control and screening group. Participant non-compliance and small inaccuracies in attributing cause of death need to be compensated for in study size, requiring several million participants. Screening trials with insufficient sample sizes might show a lowering of cancer-specific mortality but not detect increases in all-cause mortality related to screening. Studies of a manageable size have too little discriminatory power and last a long time. Furthermore, results become available decades after trial initiation, by which time they are probably antiquated. Whether screening for prostate cancer is beneficial cannot be assessed in trials, a statement that might also be true for other diseases with low specific mortality.

Introduction

By definition, screening is done for individuals without symptoms. Most people who are screened do not have the target disease, and, therefore do not benefit from screening but are at risk of possible side-effects. Screening studies require high scientific rigour to rule out a detrimental effect on the population. Cancer-specific mortality is a common endpoint in studies of cancer screening. Even for cancers that are leading causes of death the life-time risk is just a few percent in the general population. Large relative reductions in mortality are small absolute reductions. Therefore, very large numbers of participants are needed for screening studies to have adequate statistical power. Because of the necessary size and duration of such studies, their feasibility and usefulness are questionable.

Prostate cancer is the most common cancer in men worldwide.[1] Prevalence changes with age: 40% of 40-year-old men and 80% of 80-year-old men have the disease.[2] However, only a small proportion of men with prostate cancer have clinical symptoms and even fewer die. Screening programmes typically target men aged 50–75 years. For men older than 50 years, the lifetime risk of death from prostate cancer is about 3%, and the median age at death is almost 80 years.[3,4] A screening trial designed to search for a 25% relative reduction in disease-specific mortality has to include several hundred-thousand men (figure 1) who are followed up for many years. Trials of this size might discover about 75 men who did not die from prostate cancer because of screening.

Mirjam Colditz

Figure 1: **A football stadium that holds about 40 000 fans**
A prostate-cancer screening trial requires more than five times the number of men shown. The trial lasts for more than 10 years and the result is a survival difference of 75 men.

A trial might include 100 000 men in a control and 100 000 men in a screening group and continue until 10% of men in the control group have died. About 3% of these 10 000 men (300 individuals) will die from prostate cancer and 9700 from other causes. A successful screening programme might prevent 25% of deaths from prostate cancer. Therefore, in the screening group, there will be 75 fewer prostate-cancer deaths than in the control group, and 225 men will nevertheless die from prostate cancer. If screening does not affect other-cause mortality then 9700 men in this group will die from other causes.

Early detection of prostate cancer might not be advantageous. Screening is beneficial for men with cancer that would be incurable if detected clinically but in whom early treatment can prolong life; however, this group cannot be reliably defined. Screening is not beneficial for men with cancer that is curable if detected after clinical symptoms, cancer that is not curable even when detected by screening, and cancer that would not cause noticeable clinical symptoms during the patients life-time.[5] Screening might therefore lead to overdiagnosis and overtreatment of people who cannot gain from the procedure and are at risk of screening-related side-effects, including psychological stress. Is there an overall advantage or disadvantage for the population?

A systematic review concluded that evidence neither supports or refutes the use of routine screening to reduce prostate-cancer mortality.[6] Preliminary results from early detection programmes show a benefit for the screened population. However, these studies have methodological

flaws and lead-time, length, and volunteer biases.[6] Only controlled randomised trials with endpoints of all-cause and prostate-cancer-specific mortality avoid these biases and have the potential to provide valid evidence. Results from two ongoing randomised controlled trials, ERSPC (European Randomized Screening for Prostate Cancer trial) and PLCO (Prostate, Lung, Colorectal and Ovary cancer trial), will probably determine whether screening for prostate cancer is justified or not. But should evidence from one or two trials be enough to make decisions for the entire population? In this paper, I discuss methodological issues associated with large screening trials, such as accuracy in the attribution of cause of death, compliance of trial participants, and the effect of screening on all-cause mortality.

Size of screening studies

Sample sizes for screening studies are typically calculated for a statistical power of 90% and significance of 5%: if screening lowers mortality, results will show it with 90% probability; if there is no difference between screening and not screening there is a 5% probability that the results will show a difference by chance. A study designed to detect a 25% relative reduction in prostate-cancer mortality requires 300 instances of death (cases) in the control group and 225 in the screening group. In both Germany[7] and the USA,[8] the average mortality rate for prostate cancer is fewer than 60 deaths per 100 000 men per year. To accumulate 300 cases, 500 000 men in both the control and screening group have to be followed up for 1 year (table). Such a short trial is not reasonable because the screening period alone is longer than 1 year. Long follow-up periods are needed to the research question and require fewer participants to observe the same number of deaths.

A more realistic sample size than 1 million participants is 200 000–250 000 men (as in the ERSPC and PLCO trials[9]). Although 300 deaths might be expected in 4–5 years of follow-up (table), overall duration of an actual trial is much longer. Because of the natural time course of the disease, screening does not immediately affect mortality, and recruitment time and the duration of the screening period also contribute to trial length. Even in trials with a realistic size and follow-up, detection of a small difference in the number of deaths caused by prostate cancer can be like looking for a needle in a haystack (figure 2).

Diagnosis of cause of death

In studies of screening for prostate cancer,[9–11] cause-specific death is the primary endpoint. Therefore, accurate determination of cause of death is essential. In a trial with 10 000 deaths in the control arm and a risk of death from prostate cancer of 3%, 300 men will die from prostate cancer. Assuming the sensitivity of detecting death by prostate cancer is 100% and specificity is 99%, 397 deaths will be attributed to prostate-cancer but 97 will have been misdiagnosed. Thus, 24% (97/397) of supposed prostate-cancer deaths are misattributed. The proportion of misdiagnosed patients is almost the same as the proportion who benefit from screening (25%). In the screening arm of the same trial, if 75 deaths were prevented by screening, then there are 9925 deaths in total. Again, 97 were falsely attributed to prostate cancer, and 30% (97/322) of prostate-cancer deaths in this group would be misattributed. In both control and screening groups, the effects of misattribution would be much greater if sensitivity or specificity were lower.

	Participants			Prostate-cancer deaths		
	Total	Per group	Years of follow-up	Control group	Screening group	Difference (patients)
Trial 1	1 000 000	500 000	1	300 (0·06%)	225 (0·045%)	75
Trial 2	250 000	125 000	4	300 (0·24%)	225 (0·18%)	75
Trial 3	200 000	100 000	5	300 (0·3%)	225 (0·225%)	75
Trial 4	20 000	10 000	50	300 (3%)	225 (2·25%)	75

300 plus 225 deaths in the control and screening group have to be accumulated, respectively. Assumed prostate-cancer mortality rate is 60/100 000 per year.

Table: **Scenarios of prostate-cancer screening trials**

The relative reduction in prostate-cancer specific death rate between the control and screening arms is 19% rather than 25%, and the power of the trial is reduced from 90% to 80%. To compensate for this loss of power, the sample size has to be increased by about 35%. A small drop in specificity (99% instead of 100%) for the identification of cause of death means that 35% more patients are needed. With a specificity of 97% more than twice the number of participants is needed (figure 3). Relative reduction in prostate-cancer mortality is then only 20%, instead of 25%. In this scenario, the power of the trial is reduced, and to compensate the sample size has to be increased by 26%.

The hypothetical estimations above show that a screening trial with disease-specific death as an endpoint cannot rely on death certificates, the specificity and sensitivity of which are usually lower than 95%.[12,13] Labrie and colleagues,[10] whose study in Quebec, Canada, was included in a systematic review,[6] used information on prostate-cancer-specific death from the death registry of

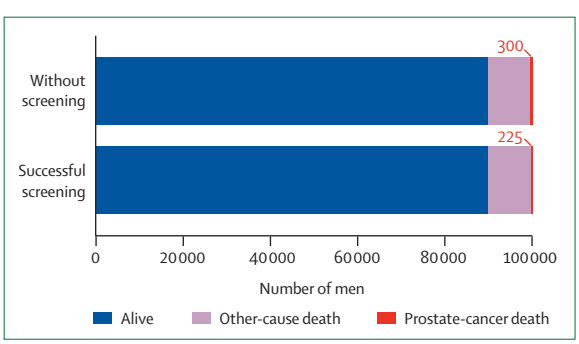

Figure 2: **Reduction of prostate-cancer deaths by 25% from 300 to 225 cases**
In a trial of 200 000 men followed up for 5 years, there are 75 fewer prostate-cancer deaths and 75 more men alive (at best) in the screening group.