Acta Oncologica Vol. 30 No. 2 1991

FROM THE DEPARTMENTS OF UROLOGY AND ULTRASOUND, HERLEV HOSPITAL, UNIVERSITY OF COPENHAGEN, HERLEV, DENMARK.

SCREENING FOR PROSTATIC CANCER

Investigational models

P. IVERSEN and S. T. TORP-PEDERSEN

Abstract

Prostatic cancer has a long natural history and a significant preclinical period, during which the disease is detectable. Thus, this common malignancy in males fulfills some of the most important criteria for initiating screening programs. However, the still enigmatic epidemiology also includes features of the disease, which make the possible gain from screening programs questionable. Thus, before embarking on expensive community or national screening programs, the beneficial effect of such an effort on morbidity and mortality must be demonstrated in large-scale trials comparing a screened population with non-screened controls. In this paper, some of the problems associated with such studies are addressed.

Key words: Prostatic cancer, screening.

Screening has been defined as the use of simple tests or examinations to differentiate those who probably have disease from those who probably do not (1).

The difference between normal medical practice and the situation where healthy individuals are exposed to screening procedures is obvious, and while the idea of screening is simple, the implementation and validation of the concept is complex.

Before considering screening for a specific disease, the following criteria should be met:

- The disease must contribute significantly to morbidity/mortality in society, i.e., it must be a common disease.
- The preclinical period should be sufficiently long, and the disease should be detectable during that period.
- An effective treatment should be available, and the combination early detection/early treatment should significantly alter the prognosis of the screened population.

Epidemiology and general considerations

Prostatic cancer is one of the most common malignancies in males and, when diagnosed, it is advanced beyond curability in the majority of cases. The probability that a screening program for prostatic cancer will influence morbidity and mortality from the disease in society more specifically depends upon:

1. The natural history of prostatic cancer:
   a) The duration of the preclinical period, where the cancer can be detected by screening.
   b) The proportion of cancers detected by screening that would progress, if left untreated.
2. A reliable screening test to detect the preclinical cancer.
3. The extent to which available treatment alters the natural history of prostatic cancer, i.e. the efficacy of therapy.
4. Compliance in the population, i.e. the percentage of the population in question attending the program.

In a screening scenario, an effective therapy is a necessity. The optimal management of localized prostatic cancer is still an unsettled issue, and randomized controlled trials are few and small. However, radiotherapy and radical prostatectomy are both modalities generally accepted as effective in the treatment of localized cancer (2, 3). Therefore, in this context, the problems of therapy will not be addressed further.

The natural history of diagnosed clinical prostatic cancer is long (4, 5) and ample evidence points to an even

Presented at the Conference on Early Prostatic Cancer. The WHO Collaborating Centre for Urological Tumours, Karolinska Hospital, Stockholm, March 21–23, 1990.

Accepted for publication 2 October 1990.

longer preclinical period. This very long preclinical period explains in part what makes cancer of the prostate unique in man—the finding of prostatic malignancy at autopsy in approximately 30–40% of males above the age of 60 (6). Further, probably reflecting the high prevalence at autopsy, 'incidental' cancer is found in approximately 10% of transurethral resections for clinically benign enlargement of the prostate (7). When comparing these percentages to the actual incidences of clinical cancer, it is clear that many men will outlive their prostatic cancer without ever enduring any morbidity from it.

The histology of these early cancers found incidentally at transurethral resections or at autopsy are indistinguishable from clinical cancers, and even though some rough correlations are established between histology and prognosis, we are not presently able to predict in the individual male if an early cancer of the prostate will cause disease and even death, or it will remain silent for the rest of his life.

We are faced with a paradox. The long natural history allows us to detect and treat the disease in a preclinical phase, but in many patients this preclinical phase exceeds the remaining life time of the individual. Thus, in a screening situation, we would not wish our screening test to detect these 'insignificant' cancers.

The ideal screening test is simple, safe, inexpensive and may be performed by a para-medical personnel. Specificity must be high to spare the many without disease unnecessary evaluation and anxiety.

A sensitivity of the test less than 100% theoretically results in preclinical cancers not being detected. However, the true sensitivity cannot be calculated without knowing the prevalence of the disease in the population. If the prevalence at autopsy is used, a sensitivity of 100% would lead to detection of cancers, that would never have required treatment. An estimate of the prevalence with relevance to screening should ideally include all preclinical cancers that would 'surface' as clinical cancers, i.e. the cumulated incidence rate. Still, a high sensitivity calculated from this prevalence would not ensure that the cancers detected at screening would be identical to those emerging clinically years later.

Clinical experience

Several non-controlled screening studies have been performed. Most investigators have used digital rectal examination (DRE) as screening test, some have employed transrectal ultrasonography (TRUS). Typically, DRE has led to a detection rate of up to 1.5% cancers (8), while TRUS has resulted in a detection rate as high as 2.6% (9). Interpretation of these studies is difficult. They are all subjects to one or more of the following biases: Selection bias, diagnostic bias, length time bias, and lead time bias (10–12). Detection of more localized tumors does not in itself prove screening to be beneficial, and an effect on mortality and morbidity from prostatic cancer has not been demonstrated in these studies.

To clarify the issue we need controlled randomized trials with a screening group compared to a control (non-screened) group.

Hypothetical screening study

In a recent study (Iversen and Torp-Pedersen: unpublished data), we designed and discussed a hypothetical screening study including a non-screened control group in a population of 65-year-olds. The hypothetical set-up and calculations were based upon Danish cancer statistics (13), a hypothetical biological model for prostatic cancer, a hypothetical screening test, and a hypothetical treatment.

The biological model proposes a preclinical period of 15 years for all tumors. When the cancers are diagnosed clinically (without screening), 20% are localized. When detected at screening one year before it would have emerged clinically (lead time = 1 year), 30% are localized; when detected 2 years before (lead time 2 years), 40% are localized . . . and 8 years before, 100% are localized. Thus stage distribution depends upon lead time, i.e. how much the detection is advanced in time by screening.

The screening test is thought to have a specificity of 100%. The sensitivity depends upon lead time. Ninety percent of cancers destined to 'surface' one year later will be detected, 80% of cancers which will surface 2 years later will be detected . . . ending with a sensitivity of 10% for cancers destined to surface 9 years later.

The treatment cures cancers confined to the prostate. Patients with advanced cancers are considered 'lost to disease' and will eventually die from the malignancy.

With these prerequisites we conducted a trial with 100 000 males screened once at age 65, and 100 000 65-year-olds serving as controls. Each year, incidence of prostatic cancer in the two groups was recorded. Stage for stage, individuals in the two groups received identical treatment.

After 10 years, 1 066 cancers were detected by our screening test in the screening group. Further 999 cancers in the group were 'overlooked' by our test and were diagnosed clinically.

In the control group, 1 944 cancers were diagnosed. The difference of 121 cancers between the two groups represents cancers 'overdiagnosed' in the screening group. This refers to cancers with the biological potential to emerge as clinical cancer had their hosts not died from other causes.

The total number 'lost to disease' in the screening group was 1 223, and in the control group 1 555. Thus, in this scenario, 332 was saved from prostatic cancer by screening 100 000 65-year-olds.

If we accept the assumptions in our hypothetical study as realistic, how many males do we need in a real study to reach statistically valid conclusions? Accepting a type 1 error of 5% and a type 2 of 10%, 23 700 males in each group are needed.

In this calculated example, the screening test only detected cancers destined to become clinical cancers. In real life, our test may detect cancers that would never surface clinically. This would lead to a much more significant 'overdiagnosis'.

In our trial model, the hypothetical presumptions may be changed, screening may be repeated and similar calculations may be performed. Even though the above example is a simplification of real life, it still serves the purpose to demonstrate the magnitude of the problem.

Problems and biases in randomized screening trials

A randomized trial is not without shortcomings. A control group ensure an average measure for how many cancers detected at screening that would have surfaced later if left alone. However, because of the treatment, we are not able to establish which cancers detected in screening that would have remained silent, and which that would have become clinical realities.

Along the same lines, we are in fact determining the effect of screening plus treatment. In the situation where treatment does not improve survival at all, we would not be able to detect any difference in mortality between the two groups, even though the screening served its purpose.

Further, how to form a screening group and a comparable control group? If we select the study population and randomize via a civil registry, it must be expected, that only a fraction of the individuals offered screening would accept the invitation. Thus, the screened population would be a (self) selected group. How to create a similarly selected control group?

Instead we could invite a number of males to participate in a randomized study. Following acceptance, half would be allocated to a control group. In this situation, the controls would be alerted to the risk of prostatic cancer, and might not behave like the background population.

Thus, in both these designs, we would be subjects to a selection bias.

As mentioned, treatment would have to be identical in the two groups in a randomized study. However, it is likely, that new and improved treatment modalities would appear. Even though the treatment was changed concurrently in the two groups, we would again be subjects to a bias, treatment bias, based on the fact that while a detected cancer in the screening group was managed with the 'old' therapy, the 'corresponding' cancer in the control group emerging x number of years later would be treated with the improved treatment. If this new treatment really improved prognosis, it is obvious that the effect of screening would decrease.

As discussed, mortality and morbidity are the only valid endpoints in a screening study. Evaluating these parameters may be very difficult in large populations, and it is possible that morbidity and mortality would be more easily recognized in the screened population which is more alerted to the risk of prostatic cancer. This possible bias may be identified as surveillance bias.

Conclusive remarks

The required size of studies to demonstrate an effect of screening on mortality from prostatic cancer must affect all with experience in clinical trials. Also, the problems in design and interpretation are of significant dimensions.

The huge cost of a screening program is beyond dispute, and whether or not to initiate large scale screening programs for prostatic cancer is in essence a political question. Even though the task seems overwhelming, further research into the topic and probably large scale multicenter/national trials are warranted to supply the best possible basis for making the right decision.

Corresponding author: Dr Peter Iversen, Department of Urology, Herlev Hospital, DK-2730 Herlev, Denmark.

REFERENCES

1. Commission on Chronic Illness in the United States. Cambridge, Massachusetts: Harvard University Press, 1968.
2. Bagshaw MA, Ray GR, Cox RS. Selecting initial therapy for prostate cancer. Radiation therapy perspective. Cancer 1987; 60: 521–5.
3. Walsh PC, Lepor H. The role of radical prostatectomy in the management of prostatic cancer. Cancer 1987; 60: 526–37.
4. Adami H-O, Norlén BJ, Malker B, Meirik O. Long-term survival in prostatic carcinoma, with special reference to age as a prognostic factor. A nation-wide study. Scand J Urol Nephrol 1986; 20: 107–12.
5. Johansson J-E, Adami H-O, Andersson S-O, Bergström R, Krusemo UB, Kraaz W. Natural history of localised prostatic cancer. A population-based study in 223 untreated patients. Lancet 1989; I: 799–803.
6. Franks LM. Latent carcinoma of the prostate. J Pathol Bacteriol 1954; 68: 603–16.
7. Sheldon CA, Williams RD, Fraley EE. Incidental carcinoma of the prostate: A review of the literature and critical reappraisal of classification. J Urol 1980; 124: 626–31.
8. Chodak GW, Keller P, Schoenberg H. Routine screening for prostate cancer using the digital rectal examination. EORTC Genitourinary Group Monograph 5: Progress and Controversies in Oncological Urology II. New York: Alan R. Liss, 1988: 87–95.
9. Lee F, Littrup PJ, Torp-Pedersen ST, et al. Prostate cancer: Comparison of transrectal US and digital rectal examination for screening. Radiology 1988; 168: 389–94.

10. Love RR, Camilli AE. The value of screening. Cancer 1981; 48: 489–94.
11. Adami H-O. Strategies for control of prostatic cancer mortality with some methodologic comments. Scand J Urol Nephrol 1988; (Suppl. 110): 15–22.
12. Iversen P, Torp-Pedersen ST. Screening for carcinoma of the prostate: Sound prophylaxis or harmful intervention? Curr Opinion Radiol 1989; 1: 255–63.
13. Danish Cancer Society, Danish Cancer Registry: Cancer Incidence in Denmark 1985. Denmark, 1985.