



Editorial

Prostate Cancer Screening: The Need For Problem-Solving that Puts Men's Interests First

Lars Holmberg^{a,b}

^a Division of Cancer Studies, Medical School, King's College London, London, UK

^b Regional Oncologic Centre, University Hospital, Uppsala, Sweden

1. Introduction

No one professionally involved in prostate cancer diagnosis and treatment would have expected the first analyses of the European Randomised Study of Screening for Prostate Cancer (ERSPC) [1] and the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) [2] to provide the *final answers* to all major questions in prostate cancer screening. We know that screening for prostate-specific antigen (PSA) level is not the ideal test, and we have a limited understanding of the effects of different alternatives for managing early disease. Our lack of understanding of the biology of the disease hampers rational decision making for the individual patient. We need to determine how these new data change this problematic situation and our discussion about the present priorities.

2. The two trials

The ERSPC results are based on a median follow-up of close to 9 yr in 72 952 men aged 55–69 yr who were randomised to screening and 89 435 men in the same age group who were randomised to control. In the men randomised to screening, 82.2% were screened at least once, and 214 men died of prostate cancer. In the control group, 326 men died of prostate cancer. With the exception of one centre, the screening interval was 4 yr, and for most of the subjects the trigger point for biopsy was a PSA level of 3.0 ng/ml. Inferring from a previous publication of Ciatto et al [3], the incidence of biopsy from screening in the control arm was <15% (with lower levels in the first years of recruitment) in the two largest centres (Finland and the Netherlands) and <10% in Spain but near 28% in Italy. The cumulative incidence of prostate cancer was 8.2% in the men allocated to screening, and it was 4.8% in the control group. At this

follow-up, there was a reduction of incidence of advanced cancer stage at diagnosis in men allocated to screening. These results corresponded to a relative risk reduction of 0.80 for dying of prostate cancer; in absolute terms, 1410 men needed to be screened, and, of these, 48 men needed to be treated for prostate cancer (or rather, managed, since some were allocated to active surveillance).

The PLCO trial analyses mainly relied on data from 7 yr of follow-up in 38 343 men aged 55–74 yr who were randomised to screening and 38 350 men aged 55 yr to 74 yr who were randomised to control. Forty-four percent of those enrolled had undergone one or more PSA tests previously. The screening procedure was annual PSA testing for 6 yr and annual digital rectal examination (DRE) for 4 yr. The trigger point for biopsy was a PSA level of 4.0 ng/ml. The compliance with screening was 85% for PSA testing and 86% for DRE in the group randomised to screening, and 50 men died of prostate cancer. In the PLCO control group the proportion PSA screened was 40% at the beginning of the trial, rising to 52% at year six, with similar figures for DRE (41–46%). Forty-four men died of prostate cancer in the control group. The incidence of prostate cancer was 116 per 100 000 persons per year in the group allocated to screening and 95 per 100 000 persons per year in the control group. At 7 yr, there was no reduction in the incidence of advanced cancer. The relative risk ratio for dying of prostate cancer was 1.13, and the incidence rate ratio was 1.22.

One short paper cannot discuss all the relevant details contained in these studies. There are also many details that we do not yet know. Neither the scientific community nor the clinical or lay communities can, in a few weeks, identify and grasp all the possible inferences. Without a doubt, these trials will be hotly debated over a long period of time. It is a fact, however, that both trials have the following char-

acteristics: (1) both are randomised, (2) both have a near-complete follow-up for the present end points, (3) both have blinded evaluation of end points, and (4) both analysed according to intention to treat. The evaluation of overall mortality indicates that the randomisation was well balanced in both studies. The analyses are straightforward and show the facts of primary interest. Both studies were undertaken by large networks of very competent researchers, and clinicians and were scrutinised by experienced data-monitoring and safety committees. These circumstances guarantee a high degree of internal control and scientific rigour. Thus, it would seem rational to take the basic results at face value and to try to see them in a larger context.

3. An attempt at interpretation

From an overall perspective, the results of the two studies seem logical and compatible: There was a modest reduction in rate of prostate-cancer screening following PSA screening at the cost of a very high risk of overdiagnosis and overtreatment. ERSPC cannot yet give a more exact estimate of the effect of, for example, a protocol with a 4-yr interval using a PSA cut-off of 3.0 ng/ml, due to different populations included and protocols involved. There is also some uncertainty with regard to how much the contamination interfered. The fact, however, that a homogeneous effect is apparent in different centres and populations is reassuring with regard to whether the qualitative mortality effect really is a reduction. It is also reassuring that ERSPC reports a reduction of the incidence of advanced cancer by screening. This modest reduction in prostate cancer mortality is very difficult to measure in a substantially smaller study population with fewer events (the PLCO had 94 deaths compared with 540 deaths in ERSPC) and with a high level of contamination in the control arm. The more modest difference in incidence between the PLCO arms compared with the substantial difference in the ERSPC was probably also a reflection of a contamination as well as of frequent prescreening. Furthermore, the substantial degree of prescreening and aggressive radical treatments in both PLCO arms improved prognosis overall, making further improvement difficult. The subgroup analyses in PLCO by prescreening status or screening status carried limited information because the prescreening already eliminated some at-risk patients and the other subgroups had very few deaths.

Even if the studies seem only to confirm what many saw as evident, the data are extremely important. The researchers should be much congratulated for steadfastly conducting the studies under many constraints and sometimes heavy pressure. They have now provided evidence from two randomised settings to improve informed decision making on several levels. Many of us who thought we saw it coming should remember that our hypotheses were nothing more than hypotheses until these results emerged. The results imply that there is, indeed, a window of opportunity where early treatment can affect the natural history of some prostate cancers. At the same

time, the results also refute unrealistic expectations about dramatic effects, which is very helpful for the debate. The ERSPC provides a glimpse of the realistic long-term expectations of the pros and cons of PSA screening in terms of lives saved and risk of overdiagnosis. Empirically, this is an very important backdrop to the discussion that will hopefully temper speculations and emotional arguments.

4. The issues

Even a modest 20% reduction in relative risk of suffering the atrocious consequences of metastatic disease and then dying from a common cancer such as prostate cancer is a worthwhile goal to consider. In PSA screening—even more so than in screening for cervical cancer, breast cancer, or colorectal cancer [4]—the price to pay is high, and the problems that remain to be solved are many and difficult. The studies illustrate this, and they touch upon a series of serious questions and obstacles.

Both research groups are acutely aware of the great problem with overdiagnosis and overtreatment. Two ways to handle this problem would be improving the test, so that only clinically relevant cancers are found, or finding biomarkers that effectively separate slow-growing and/or indolent cancers from aggressive ones. Yet, no breakthrough has come in either of these fields. Additionally, research on new diagnostics and new biomarkers will take a long time to validate; essentially, such developments would have to go through the same long-term clinical research as the ones we have today. One reservation should be made here: Biomarkers that can be measured with high validity in biobanked material could be tested, at least preliminarily, in well-defined cohorts for which follow-up information is already available.

Intimately related to the risk of overtreatment is another limitation: Not only do we have little knowledge from randomised studies about the effects of radical prostatectomy and radical radiotherapy compared with active surveillance, but so far we also have nothing to show for the modern PSA era. For men today, we have a hope of avoiding radical treatment but we have no long-term tested policies of active surveillance. We do not know whether they are safe or whether they have acceptable psychosocial consequences. The information we have about the comparative long-term cost utilities of these initial management alternatives is sparse.

The researchers themselves mention many different fields where more knowledge is of utmost importance: aspects of quality of life, cost, cost-effectiveness, optimal indication for biopsy, and optimal interval. These questions are far from trivial. To address them fully, we would need randomised data on screening procedures and patient management in terms of medical treatment as well as other support. The PLCO study, however, shows that comparing two screening scenarios will require very large studies to get an acceptable level of statistical power. This is also mirrored in the breast-cancer experience: Despite 30 yr of research and 20 yr of routine practice, the optimal interval

in breast-cancer screening still has to be gauged mostly by modelling.

Two more questions that emerge are directly linked to the performance of the PSA test that emerge, but so far they have not been discussed separately in the present publications. First, since the mortality reduction is modest, the data imply that the PSA screening as used in these studies only finds a limited fraction of those cancers that will metastasise and kill. Thus, some of these cancers may be in a curable local state only a short time after they have reached the threshold level for biopsy, and, consequently, most of them will not be picked up if the screening interval is too long. Some cancers may have already metastasised before reaching the threshold level and can only be detected with a low or very low threshold, thus signalling poorly if at all with a rising PSA level. Shortening the interval and lowering the threshold level for biopsy would, however, have drastic consequences for both the side-effects of the programme and for resource demands.

Second, the PSA test has a low specificity. In the ERSPC trial, for example, 75.9% of the men who underwent PSA testing had a false-positive result. The published research reports do not yet tell us in detail what happened to all those men. Were many of them subjected to repeat biopsies? What are the psychological side-effects of being followed with new investigations over the shorter or longer term for an elevated PSA? Hopefully, we will soon have results from studies into these questions because they are pressing.

In a subgroup analysis, the ERSPC results indicate that PSA screening has little effect in men >70 yr of age. This is in line with the lack of effect of radical prostatectomy in the age group >65 yr in the Scandinavian Prostate Cancer Group study 4 (SPCG4) [5]. Subgroup analyses should always be interpreted very cautiously, but we have two that together raise an important warning sign of more harm than benefit following PSA screening and radical treatment among older men.

5. What can we do?

It is certainly not the task of the authors of these studies, of one or another editorial, or of any small group of individuals to give specific advice about how recommendations related to PSA testing in asymptomatic men should now be changed. A rational discussion among many is needed: The question is utterly complex. The stakeholders are numerous, ranging from a broad health-policy perspective (where we want a fair and rational use of resources for all diseases), down to the individual at high risk of getting a prostate cancer and to the specialist in urology or urological oncology. Although it is probably safe to guess that the controversy will continue [6] and that there will be no major change in current practices before more data emerge, the results nevertheless point to some priorities:

- We need all the short-term and long-term information we can possibly get from these studies, and, since over-diagnosis is the overwhelming problem, data on quality of life are just as essential as the cost-effectiveness

estimates. So, investigators and funders: Do not give up and do not delay analyses that can be done today!

- The results, for primary as well as secondary end points, to come from studies such as the Prostate Cancer Intervention versus Observation Trial (PIVOT) and Prostate Testing for Cancer and Treatment (ProtecT) trials [7,8] will be extremely important to shed light on treatment effects in a screening scenario.
- Screening for cancer and communicating a cancer diagnosis without offering active intervention is a contradiction and a profound ethical dilemma. We can anticipate that a substantial proportion of men with early disease detected by screening will still receive the recommendation for active surveillance. We should have tested this concept sooner in large-scale randomised studies, but we definitely cannot go on putting asymptomatic men with screening-detected disease on an untested management regime. Most of these men should be offered a trial, and one is ongoing [9]. An important research question is whether active surveillance can be combined with an intervention with a low level of toxicity that can delay progression and make a conservative approach safer.
- As pointed out, the need for research to improve the test and to find biomarkers for clinically relevant disease lies at the core of the whole predicament.
- The rule that full and understandable information based on evidence should be given to all men before a PSA test still stands.

The appeal for more and better research is not a surprise to anyone. The current situation, however, also actualises two other pleas that can help serve present and future prostate cancer patients. First, none of the research results will come easily. The ERSPC and PLCO trials show that, even with large patient groups and long-term follow-up, there is still considerable uncertainty. This means that long-term, large, broadly multidisciplinary, and multiprofessional networks need to collaborate in scientifically well-designed, creative, and professionally driven studies. Smaller-scale initiatives and less rigorous studies detract energy from what really needs to be done. Second, prostate-cancer screening is already, and promises to stay, one of the big medical debates of our time; it illustrates so many of the dilemmas of cancer screening, diagnosis, and treatment. Regrettably however, we can infer from other screening debates that there is a high risk of meaningless controversies over which study is *right* and which is *wrong*, thus hiding the real issues behind a smokescreen. All studies will have problems and can be criticised, especially when seen in a historic light, but the destructive type of *armchair epidemiology* [10], that only aims at finding errors that will invalidate studies rather than at rationally trying to understand what knowledge these studies can contribute and where they fail to be informative, should be avoided at all costs. Validity occurs along a continuum; we should avoid thinking in black and white, and we definitely should not rate studies by their *p*-values [11]. Both the PLCO and ERSPC contain valuable information and can, together,

inform us! Too much is at stake for all men potentially exposed to the offer to be screened, so narrow medical or other territorial fights must not be allowed to dominate the scene. I would strongly urge the different stakeholders to adopt Karl Popper's motto: "I may be wrong, and you may be right and by an effort we may get nearer to the truth" [12]. To solve all the difficult issues involved in prostate cancer screening, a truly open, critical debate founded on rational arguments soundly based in empirical findings and scientific inference is needed.

Conflicts of interest: The author has nothing to disclose.

References

- [1] Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 2009;360:1320–8.
- [2] Andriole GL, Grubb III RL, Buys SS, et al. Mortality results from a randomized prostate cancer screening trial. *N Engl J Med* 2009;360:1310–9.
- [3] Ciatto S, Zappa M, Villers A, Paez A, Otto SJ, Auvinen A. Contamination by opportunistic screening in the European randomized study of prostate cancer screening. *BJU Int* 2003;92:97–100.
- [4] Hakama M, Auvinen A. Cancer screening. In: Heggenhougen K, Quah SR, editors. *International Encyclopedia of Public Health*. San Diego, CA: Academic Press; 2008. p. 464–80.
- [5] Bill-Axelsson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med* 2005;352:1977–84.
- [6] Barry MJ. Screening for prostate cancer—the controversy that refuses to die. *N Engl J Med* 2009;360:1351–4.
- [7] Wilt TJ, Brawer MK, Barry MJ, et al. The Prostate Cancer Intervention versus Observation Trial: VA/NCI/AHRQ cooperative studies program #407 (PIVOT): design and baseline results of a randomized controlled trial comparing radical prostatectomy to watchful waiting for men with clinically localized prostate cancer. *Contemp Clin Trials* 2009;30:81–7.
- [8] Donovan J, Hamdy F, Neal D, et al. Prostate Testing for Cancer and Treatment (ProtecT) feasibility study. *Health Technol Assess* 2003; 7:1–88.
- [9] Klotz L. Active surveillance for prostate cancer: trials and tribulations. *World J Urol* 2008;26:437–42.
- [10] Paul JR, Freeman AW, Perkins JE. What and who is an epidemiologist? Comments on an editorial published in the April 1942 issue of the journal. *Am J Publ Health* 1942;32:867–9.
- [11] Savitz DA. *Interpreting Epidemiologic Evidence: Strategies for Study Design and Analysis*. New York, NY: Oxford University Press; 2003
- [12] Popper K. *The Myth of the Framework: In Defence of Science and Rationality*. Abingdon, Oxon, UK: Routledge; 1995.

Embracing Excellence in Prostate, Bladder and Kidney Cancer

27-29 November 2009

Barcelona, Spain



www.emucbarcelona2009.org

2nd European Multidisciplinary Meeting on Urological Cancers organised by:

EAU European Association of Urology

ESMO European Society for Medical Oncology

ESTRO European Society for Radiotherapy and Oncology