Major clinical trials using prostate-specific antigen (PSA) as the screening test to detect localized early-stage prostate cancer and to attempt to change its natural history with early intervention have yielded conflicting interpretations. The US Prostate, Lung, Colorectal, and Ovarian (US PLCO) cancer screening trial concluded that PSA-based screening conferred no meaningful survival benefit, whereas the European Randomized Study of Screening for Prostate Cancer (ERSPC) and the GOTEBOG clinical trial (GOTEBORG) trials claimed statistically significant life-saving benefits. These divergent outcomes have not provided physicians with clarity on the best evidence-based treatment. To determine the extent to which these divergent outcomes are clinically meaningful, we evaluated these data and those of a long-term prospective cohort study in the context of the clinically documented harms of androgen deprivation therapy (ADT). We noted the unheralded fact that in both European trials far more patients received hormonal treatment in the control than the prostatectomy arm, whereas hormonal therapy in the US trial was balanced between arms. We examined this imbalance in ADT treatment and prostate cancer–related deaths in the contexts of contamination, stage migration, and attribution of cause of death, all of which impinge on data interpretation. The ERSPC and GOTEBOG data are compatible with the hypothesis that ADT treatment contributes differentially to an increase in prostate cancer deaths in control patients. If so, the claim of a reduction in prostate cancer deaths in the screened cohort requires reappraisal. The conventional interpretation that PSA screening and radical treatment intervention are the major contributors to the results of these two studies needs more rigorous scientific scrutiny, as does the role of ADT treatment of nonmetastatic disease.


Commentary

Prostate-Specific Antigen Screening Trials and Prostate Cancer Deaths: The Androgen Deprivation Connection

Ian E. Haines, George L. Gabor Miklos

Manuscript received May 9, 2013; revised July 25, 2013; accepted July 25, 2013.

Correspondence to: Ian Haines, MBBS, FRACP, FACHPM, Department of Medical Oncology, Monash University at Cabrini Hospital, Malvern, Victoria, Australia (e-mail: lan.haines@monash.edu) and George L. Gabor Miklos, PhD, Secure Genetics, Newport Beach, New South Wales, Australia (e-mail: gmiklos@securegenetics.com).

Prostate-specific antigen (PSA) screening of asymptomatic men from the general population facilitates the detection of more localized early-stage tumors of the prostate. After treatment with prostatectomy and/or radiation and/or androgen deprivation therapies (ADTs), it is widely believed that more of these patients live longer than those who were not screened and whose tumors were found at a later and possibly more advanced stage. The data underpinning this belief, however, are ambiguous, and the claimed benefits of radical treatment, particularly prostatectomy, for early-stage disease are still actively debated, despite its establishment in clinical practice.

We have reexamined the data largely from the three recent high-quality clinical trials (1–3) that have yielded contrasting results. The European Randomized Study of Screening for Prostate Cancer (ERSPC) and the GOTEBOG clinical trial (GOTEBORG) screening trials claimed that organised PSA screening and early treatment intervention saved lives because statistically significant differences occurred in the relative risk of prostate cancer mortality between the screening and control arms (1,2). In contrast, the Prostate, Lung, Colorectal, and Ovarian (PLCO) cancer screening trial revealed no statistically significant changes in risk (3). These diametrically opposite outcomes have generated controversy among clinicians (4) and confusion among patients and yielded contrasting recommendations from various professional bodies (5–8).

Examination of these data from the viewpoint of ADT allows for a more productive approach to the controversy. We propose the following hypothesis. Rather than PSA-based screening and improved treatment leading to reduced deaths in the screening cohorts, it is increased and possibly questionable use of ADT in the control cohorts that leads to increased deaths from prostate cancer. This proposal is evaluated in the context of the PLCO, ERSPC, and GOTEBOG data.

Radical Treatments and Prostate Cancer Deaths: The Prostate Cancer Intervention Versus Observation Trial (PIVOT)

The underlying premise of performing screening trials to detect cancers at an earlier stage is that the treatments employed to treat the increased number of patients diagnosed with early-stage disease will have previously been proven to cure or benefit more patients than other possible treatments or no treatment at all. Although this is true for surgery and bowel cancer, it has not been proven for radical treatment, particularly prostatectomy, for early-stage prostate cancer.
This divergence between prostatectomy treatment and relative risk is clearly demonstrated by the data of the Prostate cancer Intervention Versus Observation Trial (PIVOT) trial (9), in which men with histologically confirmed localized prostate cancer (T1–T2N × M0) were randomly assigned to prostatectomy or observation, with prostatectomy performed on 78.8% of men in the surgical arm and 10.1% in the observation cohort. After a follow-up of 12 years, 47.0% of men assigned to prostatectomy had died vs 49.9% assigned to observation. Of these deaths, 8.8% were definitely attributed to prostate cancer in the prostatectomy arm vs 9.8% in the observation group (9).

It is salient that despite nearly eightfold more prostatectomies in the surgical patients vs the observation patients, there was no statistically significant difference in either all-cause or prostate-cancer mortality compared with observation. Furthermore, although all patients were initially deemed negative for metastatic disease, bone metastases developed in 4.7% of men in the radical prostatectomy group vs 10.6% of the observation arm. Despite this twofold stage migration of control patients into the highest risk category, there was no statistically significant increase in deaths definitely due to prostate cancer.

### Critical Issues in the Interpretations of Prostate Cancer Data

There are several important issues that impinge on the survival benefits attributed to PSA-based screening and treatments for prostate cancer. The first is the correct attribution of cause of death. In the absence of the gold standard of autopsy, unambiguous assignment of death from prostate cancer, as inferred from death certificates, is subject to considerable uncertainty. Autopsy-proven immediate cause of death vs the clinical diagnosis is divergent 35% of the time in cancer patients (10). In addition, every country has different legal requirements concerning autopsies and death certificates, and these differ between different European countries, which themselves differ from the United States (11). This is particularly germane in the case of the ERSPC trial, where death certificates were from seven different countries.

Also, death is less likely to be attributed to prostate cancer for those patients who received aggressive treatment in screening cohorts (12). This bias artificially inflates the probability of lives saved because the difference between the screening and control arms favors the screening arm.

The second is the level of contamination of the control cohort by individuals who have undergone PSA-based screening. These contamination levels differed in the control arms of the US and European trials (13), leading many physicians to believe that only the European trials, with their lower levels of contamination relative to PLCO, are relevant to the life-saving benefits of screening and radical treatment. The third is stage migration, which means that those on the screening arm are more likely to have earlier stage and therefore more curable disease. The fourth is treatment and ongoing management. Patients assigned to a screening arm were more likely to receive radical prostatectomy and/or to be treated at a high-volume academic center.

Determining the extent to which radical surgery, radiation, and hormonal treatment each make their varying contribution to prostate cancer deaths in the screening and control arms of clinical trials is the outstanding challenge. This requires deconvolution of the data, the various steps of which are set out below.

#### Radical Treatments and Prostate Cancer Deaths: The PLCO, ERSPC, and GOTEBORG Trials

In these three screening trials, more early-stage cancers were diagnosed, as expected, in the screening arms of each study. The authors estimated the relative risks (RRs) of death from prostate cancer from cumulative mortality rates. After 13 years of follow-up, the relative risk of prostate cancer mortality in PLCO was a non-statistically significant 1.09 between the screening and control arms (95% confidence interval [CI] = 0.87 to 1.36), yielding no benefit from early diagnosis and treatment. After 11 years of follow-up in ERSPC, the relative risk was a statistically significant 0.79 between the two arms (95% CI = 0.68 to 0.91), interpreted as a 21% risk reduction owing to early diagnosis and treatment. After 14 years of follow-up in GOTEBORG, the relative risk was a statistically significant 0.56 (95% CI = 0.39 to 0.82) between the two arms, interpreted as a massive 44% risk reduction owing to the early diagnosis and treatment, which was usually radical prostatectomy.

If local treatment, particularly radical prostatectomy, is an important risk reducer in these studies, then the trials with the largest differences in surgical intervention between screening and control arms should yield the largest differences in relative risk, these being GOTEBORG (44% relative risk) and PLCO (with 0% relative risk). This expectation is not fulfilled (Table 1).

Prostatectomies in the screening arms of GOTEBORG and PLCO were nearly identical at 41.1% vs 40.0%, with rates of 33.6% vs 36.3%, respectively, in the control arms. The differences in prostatectomy rates between arms, of 7.5% in GOTEBORG and 3.7% in PLCO, are small and do not explain the large difference in relative risks. The biggest difference in prostatectomies between arms was actually 9.8% in ERSPC, with an intermediate

---

**Table 1. The numbers of radical prostatectomies (RPs) in the screening and control arms of the PLCO, ERSPC, and GOTEBORG trials**

<table>
<thead>
<tr>
<th>Trial</th>
<th>RPs in screening arms</th>
<th>RPs in control arms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>RP</td>
</tr>
<tr>
<td>PLCO</td>
<td>4250</td>
<td>1700</td>
</tr>
<tr>
<td>ERSPC</td>
<td>6963</td>
<td>2350</td>
</tr>
<tr>
<td>GOTEBORG</td>
<td>1138</td>
<td>468</td>
</tr>
</tbody>
</table>

* PLCO = the Prostate, Lung, Colorectal, and Ovarian cancer screening trial; ERSPC = the European Randomized Study of Screening for Prostate Cancer trial; GOTEBORG = the GOTEBORG clinical trial.
21% relative risk. Relative mortality risks do not track surgical intervention frequencies.

If the PLCO and PIVOT results are accurate, and early diagnosis and radical treatment does not cure more patients, what other likely explanation is there for the apparent benefits seen for early diagnosis and treatment in the ERSPC and GOTEBORG studies? Is it possible that the treatment of the control patients in these two studies increased the risk of earlier death from prostate cancer for some of these patients? Further analyses of the data from the three screening trials point to this possibility (Table 2).

In PLCO, ERSPC, and GOTEBORG, prostate cancer deaths per 10,000 men screened were 41, 41, and 44 (a tight clustering), whereas the respective control figures were divergent at 38, 52, and 78, (Table 2). Furthermore, deaths from prostate cancer (among only those men who had been diagnosed with prostate cancer) were 3.7%, 4.3%, and 3.9%, respectively (again a tight clustering), whereas the control figures were again divergent at 3.8%, 8.6%, and 10.9%, respectively, (Table 2).

Instead of the inferred reductions in relative risk and of lives saved among the screened patients, these data are compatible with the alternate proposal that the increased ancillary treatments used to control disease in patients of ERSPC and GOTEBORG produced an increase in prostate cancer deaths among control patients, relative to a near constant baseline of prostate cancer death in the screening patients.

**Differential Treatments and Deaths Between Screening and Control Arms**

Clinical trials are robust when the percentage of patients receiving ancillary treatments is well balanced between arms. Although such treatments were well balanced between the arms of the PLCO trial, they were quite divergent in the ERSPC and GOTEBORG trials (Table 3).

The percentages of different interventions in the screening and control arms of PLCO were as follows: ADT alone: 7.1 vs 8.6; ADT and radiation: 18.4 vs 21.3; and radiation alone: 21.2 vs 20.8. By contrast, there were large imbalances in the percentage of patients who received ADT in the screening vs control arms of ERSPC and GOTEBORG. Differential use of ADT was heavily skewed between arms of ERSPC (8.8% vs 19.6%) (1), and GOTEBORG (7.0% vs 22.6%) (2). A similar large imbalance occurred with ADT plus radiation treatment (8.7% vs 19.2% in ERSPC). Radiation alone was reasonably balanced in all three trials.

Are prostate cancer deaths and ADT treatment linked? The percentages of patients treated with ADT as initial monotherapy in the screening arms of PLCO, ERSPC, and GOTEBORG were close, at 7.1%, 8.8%, and 7.0%, respectively (Table 2), and deaths from prostate cancer in the screening arms were also close, at 3.7%, 4.3%, and 3.9%, respectively. In contrast, the percentage of patients who received ADT in the control arms was very different (8.6%, 19.6%, and 22.6%, respectively). Increasing ADT treatment tracked increasing prostate cancer deaths in the control arms (3.8%, 8.6%, and 10.9%, respectively) (Table 2).

Thus, increased ADT treatment and increased death from prostate cancer track each other, but to what extent is this a byproduct of stage migration, which occurred to a relatively similar extent in the control group of each study, or to a more fundamental effect?

In PLCO, hormone therapy treatment alone was given to 7.8% of the 8065 men diagnosed with prostate cancer. This treatment was well balanced between intervention and control groups for all prognostic subgroups. The stage breakdown was stage II T1 tumors (7.3% to intervention vs 7.9% to control), stage II T2 tumors (5.8% vs 8.5%), stage III (13.8% vs 10.8%), and stage IV (74% vs 69.4%). These figures are in stark contrast with the different prognostic subgroups of ERSPC (14) (excluding patients with metastatic disease where there was no difference between study arms) (Table 4).

A similar 8.9% of men overall received hormone therapy alone as primary therapy (5.7%, for screened patients vs 14.5% for control patients). The difference for risk categories was small for low-risk (2.0% vs 3.9%) and intermediate-risk (6.4% vs 8.0%) cancers but was statistically significantly different for high-risk disease, with 14.7% of screened patients and 29.5% of control patients

### Table 2. Deaths from prostate cancer in the screening and control arms of the three major trials and the percentage of deaths from prostate cancer among only those patients diagnosed with prostate cancer*

<table>
<thead>
<tr>
<th>Trial</th>
<th>Patients with PCa</th>
<th>PCa deaths</th>
<th>PCA deaths per 10,000 patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCO</td>
<td>38,240</td>
<td>4,250</td>
<td>158</td>
</tr>
<tr>
<td>ERSPC</td>
<td>72,891</td>
<td>6,963</td>
<td>299</td>
</tr>
<tr>
<td>GOT</td>
<td>99,952</td>
<td>11,38</td>
<td>44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trial</th>
<th>Patients with PCa</th>
<th>PCa deaths</th>
<th>% PCA deaths among PCa patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCO</td>
<td>38,240</td>
<td>4,250</td>
<td>158</td>
</tr>
<tr>
<td>ERSPC</td>
<td>72,891</td>
<td>6,963</td>
<td>299</td>
</tr>
<tr>
<td>GOT</td>
<td>99,952</td>
<td>11,38</td>
<td>44</td>
</tr>
</tbody>
</table>

* PLCO = the Prostate, Lung, Colorectal, and Ovarian cancer screening trial; ERSPC = the European Randomized Study of Screening for Prostate Cancer; GOTEBORG = the GOTEBORG clinical trial; PCa = prostate cancer.
treated with hormone therapy alone. This was a difference of 158 men in the high-risk group. Because it would only have taken a reduction of as few as 11 deaths attributed to prostate cancer in the control group overall to have negated the apparent improved prostate cancer mortality rate in the screened group, this large difference in treatments provided to patients of equivalent stage may explain the different outcomes for the two arms of this study (15).

Three other datasets are informative in this regard because they reveal the clinical consequences of ADT. The first is a recently concluded prospective cohort study on ADT that began before the PSA era and ran for a remarkable 32 years. The data are not only unambiguous but also quite remarkable in terms of their implications (16). Of 142 men who never underwent ADT, none died of prostate cancer; all died of other causes. However, of the 79 men who were prescribed ADT for symptomatic local progression of disease, 48% died of prostate cancer and 52% died of other causes. The reasonable conclusion from these data is that ADT as monotherapy is a contributor to prostate cancer death.

The second dataset contains the widespread use of ADT as primary treatment for localized tumors, where outcomes are worse compared with observation (17,18). Preclinical data reveal that reduced androgen levels promote the growth of high-grade aggressive hormone-refractory tumors (19,20) and reinforce the adverse clinical effects of finasteride (21), with the US Food and Drug Administration concluding that the risks involved in finasteride treatment of healthy men outweighed the benefits (22). Finally, although ADT in combination with radiation therapy for high-risk locally advanced disease produces statistically significant gains, such as less local progression, fewer distant metastases and less biochemical failure, overall survival remains unchanged (23,24).

The third dataset is an analysis of 19,271 men with localized (T1–T2) prostate cancer from the Surveillance, Epidemiology and End Results database of the US National Cancer Institute (25), in which 41% received primary ADT treatment and 59% were managed conservatively. There were more deaths from prostate cancer among patients receiving primary ADT than those treated conservatively. Thus contrary to the current belief that primary ADT treatment is beneficial for men with localized tumors of the prostate, this high-quality population-based dataset reveals that hormonal intervention increases prostate cancer deaths.

Finally, ADT as monotherapy is linked to multiple adverse events (26–31). Understandably, some physicians are reluctant to use ADT other than for proven metastatic disease because it has neither been prospectively tested as sole therapy for locally advanced disease in a randomized study nor approved by the US Food and Drug Administration for this usage. Nevertheless, its use has greatly expanded outside the group with metastatic disease as has the cost to the health-care system. Widespread use of ADT may cause death from cardiovascular disease and diabetes before men would ordinarily die of prostate cancer (32).

### Clinical and Statistical Caveats

The two European trials claimed that lives were saved by PSA screening and timely treatment intervention. However, the robustness of this claim depends absolutely upon having balanced ancillary treatments, which was not the case in ERSPC and GOTEBOG. Statistical modeling of earlier and incomplete ERSPC data claimed that after adjusting for the clinical stage of patient’s cancers, the differences in treatment between arms were unlikely to play a major role in prostate cancer death (14). However, that analysis was problematic because, among its modifications, 379 men with distant metastases (of whom 353 received ADT) were excluded “to improve statistical analyses,” and 3030 of 8010 missing data values were “filled in by a statistical method that accounted for correlations between variables” (14). Reanalysis of the latest published data (1) may clarify this situation.
It is unavoidable that owing to later diagnosis, the cohort of control patients in screening studies will exhibit stage migration and will contain a greater proportion with more advanced tumor grade and stage. Thus the claim that lives are saved by early diagnosis and treatment such as prostatectomy requires showing that this stage migration causes more deaths from prostate cancer in control patients. This can only be demonstrated if there is no variability in the initial treatment of patients with similar risk and similar stage disease. The screening and control patients in the PLCO study were all consented before randomization and were staged and managed similarly in the same centers, and this produced the confluence between initial treatments employed. In contrast, because randomization occurred first and only screened patients were consented, the screened and control patients diagnosed with early-stage prostate cancer in ERSPC and GOTEBORG were managed in different centers.

Furthermore, the inference of lives saved from the ERSPC trial has been portrayed by some medical and advocacy groups as a head-to-head comparison of PSA-screened patients vs men who were not screened. As emphasized by Brawley and Goldberg, this description diverges from the facts: “ERSPC should be considered a comparison of a group of men who were screened and got American-style prostate cancer treatment if diagnosed, versus a group who were not screened and got European-style treatment if diagnosed” (32).

The US Preventive Services Task Force evaluated the relative risk reduction of 21% inferred from the ERSPC trial and found it to be statistically fragile. It noted that this value was heavily influenced by the results from two countries, Sweden and the Netherlands, with no statistically significant risk reduction in the other five countries (Belgium, Italy, Finland, Spain, and Switzerland) (5). The magnitude of the claimed 21% relative risk reduction, and hence of lives saved by PSA screening and treatment, is first and foremost a result of the inappropriate pooling of heterogeneous results from different European countries. It is also compromised by the underreported fact that the Swedish component of ERSPC and the GOTEBORG data overlap. ERSPC contains a 60% subset of the GOTEBORG data that were transferred into it (5,32). It is important to acknowledge that the differential impact of ADT treatment between arms will further erode the 21% figure and move the data toward non-statistical significance. Similarly, any potential bias in attributing cause of death to prostate cancer for those patients who received aggressive treatment in screening cohorts (12) would further reduce this value. These caveats all point to a lower relative risk reduction than the claimed 21% value.

The clinical reality is that the mortality difference between arms of ERSPC and GOTEBORG is an amalgam of many factors, including differential ancillary treatments. Currently, it is impossible to deconvolute this difference into its contributing components from the published ERSPC and GOTEBORG data. Similarly, the trial data do not allow us to conclude that ADT monotherapy was the main driver of differential death between arms. Hence we are unable to demonstrate clear causality from the ERSPC and GOTEBORG data; however, the 32-year study, the morbidity data, the differential ancillary treatment data, and the US National Cancer Institute’s population-based Surveillance, Epidemiology and End Results data are all highly suggestive of ADT’s impact.

### Synopsis

The claims of lives saved from the ERSPC and GOTEBORG trials as a result of the PSA screening of asymptomatic men is both clinically and statistically unconvincing. The conventional explanation for the results is that the initial treatments used were balanced for all prognostic subgroups in ERSPC and GOTEBORG and that the difference in mortality between trial arms was due to reduced mortality in the screened group because of earlier diagnosis. This hypothesis rests on the foundation that there was much lower contamination of the control groups by patients having PSA-based screening tests in ERSPC than by patients in PLCO, where contamination approached 85% (13). This contamination may have contributed to a larger proportion of patients with lower-stage disease in the control arm of PLCO than in both ERSPC and GOTEBORG, meaning that the higher use of primary ADT in the control arm of ERSPC and GOTEBORG relative to PLCO is a direct result of this stage migration.

The second hypothesis, raised herein, is that widespread and inappropriate use of ADT in the control arm increases prostate cancer deaths in the control arm over and above those due to stage migration. Thus rather than PSA-based screening and improved treatment leading to reduced deaths in the screening cohorts, it is increased ADT treatment in the control cohorts that leads to increased deaths from prostate cancer.

The extent to which one can determine the contribution of each of these hypotheses to the clinical outcome depends ultimately on the quantitative and predictive capacity of each of them. As it stands, the published clinical data are insufficient to distinguish between the two hypotheses. More detail is required about the stage and grade of prostate cancers diagnosed in the ERSPC and GOTEBORG studies to allow further assessment of whether the difference in mortality results was due to an imbalance of initial treatments used and whether the greater use of ADT hormone therapy in ERSPC and GOTEBORG contributed to more prostate cancer deaths in their control groups. This detail would include specific information about the treatments used for all patients who definitely died of prostate cancer.

Given both the current uncertainty surrounding the benefits of ADT monotherapy and the risks of its well-documented harms to patients (32), it is imperative to rapidly resolve its contribution to all-cause and prostate-cancer deaths. To achieve this, the immediate priority is to obtain access to the detailed, but as yet unpublished, specifics of stage, treatment, and patient deaths from the assembled data of each study.

### References


Funding

All of this work has been carried out with personal funds. There have been no sponsors or external financial inputs.

Note

GLG Miklos states for the record that he does not accept funds from federal or state governments; does not accept grants, gifts, gratuities or remuneration for speaking commitments from drug companies or their proxies; and does not hold shares in companies that relate to any of the technologies mentioned in this work. Furthermore, he does not accept funds from federal or state governments; does not accept grants, gifts, gratuities or remuneration for speaking commitments from drug companies or their proxies; and does not hold shares in companies that relate to any of the technologies mentioned in this work.

Affiliations of authors: AMREP Department of Medicine, Monash University, Victoria, and Melbourne Oncology Group, Cabrini Hospital, Malvern, Victoria, Australia (IEH); Secure Genetics Pty Ltd, Newport Beach, NSW, Australia (GLGM)