Absolute Effect of Prostate Cancer Screening: Balance of Benefits and Harms by Center within the European Randomized Study of Prostate Cancer Screening

Anssi Auvinen1, Sue M. Moss2, Teuvo L.J. Tammela3, Kimmo Taari4, Monique J. Roobol5, Fritz H. Schröder5, Chris H. Bangma5, Sigrid Carlsson6,7, Gunnar Aus6, Marco Zappa8, Donella Puliti8, Louis J. Denis9, Vera Nelen9, Maciej Kwiatkowski10,11, Marco Randazzo10,12, Alvaro Paez13, Marcos Lujan14, and Jonas Hugosson6

Abstract

Purpose: The balance of benefits and harms in prostate cancer screening has not been sufficiently characterized. We related indicators of mortality reduction and overdetection by center within the European Randomized Study of Prostate Cancer Screening (ERSPC).

Experimental Design: We analyzed the absolute mortality reduction expressed as number needed to invite (NNI = 1/absolute risk reduction; indicating how many men had to be randomized to screening arm to avert a prostate cancer death) for screening and the absolute excess of prostate cancer detection as number needed for overdetection (NNO = 1/absolute excess incidence; indicating the number of men invited per additional prostate cancer case), and compared their relationship across the seven ERSPC centers.

Results: Both absolute mortality reduction (NNI) and absolute overdetection (NNO) varied widely between the centers: NNI, 200–7,000 and NNO, 16–69. Extent of overdiagnosis and mortality reduction was closely associated [correlation coefficient, \( r = 0.76 \); weighted linear regression coefficient, \( \beta = 33; 95\% \) confidence interval (CI), 5–62; \( R^2 = 0.72 \)]. For an averted prostate cancer death at 13 years of follow-up, 12 to 36 excess cases had to be detected in various centers.

Conclusions: The differences between the ERSPC centers likely reflect variations in prostate cancer incidence and mortality, as well as in screening protocol and performance. The strong interrelation between the benefits and harms suggests that efforts to maximize the mortality effect are bound to increase overdiagnosis and might be improved by focusing on high-risk populations. The optimal balance between screening intensity and risk of overdiagnosis remains unclear.

Clin Cancer Res; 22(1); 243–9. © 2015 AACR.

Introduction

Prostate cancer screening remains controversial despite evidence from the European Randomized Study of Screening for Prostate Cancer (ERSPC) showing a 20% reduction in prostate cancer mortality (1).

The major adverse effect of prostate cancer screening is overdiagnosis, that is, detection of cases that would not have been diagnosed during a man’s lifetime in the absence of screening. The extent of overdiagnosis has been estimated as 25% to 50% of the screen-detected cases (2, 3). Overdiagnosis is due to both substantial lead time gained by screening and increased detection of slow-growing or even nonprogressive disease. Overdiagnosis can rarely be recognized at individual level at the time of diagnosis depending on a man's remaining lifespan. Therefore, the full extent of overdiagnosis can generally be shown only through comparison of screened and unscreened populations. However, modeling studies have suggested that if the long-term benefits of screening are predicted on the basis of ERSPC, they would eventually outweigh the adverse effects in terms of quality-adjusted life-years (4).

Absolute measures of effect (such as risk or rate difference) provide more concrete indication of intervention effect, as they reflect the probability of the benefit or adverse effect, that is, the increase or decrease in frequency of occurrence (and depend usually also on the level of background risk). In contrast, relative measures (for instance rate ratio or risk ratio) do not capture this aspect and cannot be interpreted as probabilities of benefit and
Our findings indicate that in a population-based screening program, 700 to 2,000 men need to be invited for screening to avert one prostate cancer death during 13 years, depending on the screening algorithm and background risk in the population. The excess numbers of cases detected by screening are considerably larger, with one additional diagnosis per each 16 to 69 invited men. If the ratio of averted prostate cancer deaths and overdiagnosed cases (the pros and cons) is largely constant, more aggressive screening is justified mainly for high-risk groups that are likely to obtain the largest benefits (very likely men with little comorbidity, those with family history) and those who do not emphasize the potential adverse effects of overdiagnosis (men who can cope with treatment-related side effects or active surveillance as treatment strategy).

Within the ERSPC, there are differences in both study populations (e.g., underlying prostate cancer incidence and mortality) as well as screening procedures (despite common core protocol; e.g., in number of screening rounds). Here, we describe the differences in the absolute effect of screening, with number of men needed to invite to avoid one prostate cancer death (NNI) as an indicator of the screening benefit. For harms, we used absolute risk of overdiagnosis expressed as number needed for overdiagnosis (NNO), calculated as the inverse of the excess risk (cumulative incidence) of prostate cancers in the screening versus control arm. The balance of the benefits and harms (NNI/NNO ratio) is expressed as the number of excess cases needed to detect for averting one prostate cancer death (number needed to detect, NND). The pooled ERSPC results on the basis of the same data were published recently (1), and the present article constitutes a secondary analysis.

Materials and Methods
ERSPC is a randomized multicenter trial, with the primary aim to assess the effect of prostate-specific antigen (PSA)–based screening on prostate cancer mortality. The trial was initiated in the mid-1990s (from 1991 to 1998) in seven centers (the Netherlands, Belgium, Sweden, Finland, Italy, Spain, and Switzerland). France joined later but was excluded from this analysis due to short follow-up.

Men in the intervention arm were invited to PSA-based screening and those in the control arm received no intervention (usual care, Fig. 1). Eligible subjects were identified from population registers and individually randomized on the basis of random numbers (with 1:1 allocation, except in Finland with a fixed size of the screening arm leading to an intervention/control ratio of 1:1.5). Consent was obtained prior to randomization in most countries (volunteer-based efficacy trial), but randomization before consent was used in Sweden, Finland, and Italy.
Screening tests (an ancillary test was used at 3.0–3.9 ng/mL in Finland). Screen-positive men were referred to prostatic biopsies. Screening was discontinued after 3 screening rounds in most centers but continued up to 5 times in the Netherlands and 10 rounds in Sweden. Institutional review board approvals were obtained in each center. The ERSPC trial is registered in Current Controlled Trials as ISRCTN49127736.

Data on deaths and emigrations were obtained from national or regional population registries. Prostate cancer cases detected outside screening (interval cases, cancers in nonparticipants and in the control arm) were identified through cancer registries. Prostate cancer deaths were ascertained in each center by causes of death certifying a uniform algorithm developed for the trial (10). In Finland, death certificates were used after the initial period, on the basis of the outcome and number of men as the offset.

The present analysis is limited to the core age group, 55 to 69 years, to improve comparability between the centers. For the same reason, follow-up was truncated at 13 years, even if data were available through 2010, potentially allowing longer follow-up in some centers. A mean length of follow-up of at least 13 years was available in all included centers, except only 10 years in Switzerland. The intention-to-screen principle was used in the analysis, with groups compared formed by random allocation (regardless of compliance).

NNI to avert one prostate cancer death was calculated as the inverse of the absolute risk difference in prostate cancer mortality between the arms \( NNI = 1/(M_s - M_c) \) to indicate the mortality reduction by screening. As an indicator of overdiagnosis, NNO was calculated as the inverse of the risk difference in cumulative incidence of prostate cancer between the trial arms \( NNO = 1/(I_s - I_c) \), which represents the absolute risk of overdiagnosis (12). Finally, NND as a measure of the overall impact (benefits and harms) is shown as the ratio of the reduction in prostate cancer mortality to the excess prostate cancer incidence (NNI/NNO). In the calculation of number needed indicators, the values are always rounded upward (to the next integer). For their interpretation, it is essential to appreciate that a smaller value indicates a larger effect.

Linear regression with weighting by number of prostate cancer deaths was used to assess the relationship between NNI and NNO between the centers, as well as in the analysis of the relation between the characteristics of the ERSPC centers (baseline risk, number of screens, proportion of positive screening tests, and proportion of screen-positive men biopsied) and the impact measures (NNI, NNO, and NND). The assumption of normal distribution was assessed using Shapiro–Welks test. Poisson regression analysis was conducted with the numbers of cases as the outcome and number of men as the offset.

**Results**

A summary of the screening performance by center is shown in Table 1. The size of the study population (within the core age group) ranged from in 2,197 in Spain to 80,379 in Finland (Table 2). The proportion of men in the screening arm attending at least once was more than 90% in the volunteer-based centers and approximately 75% in the population-based centers. The proportion of positive tests was lowest (~11%) in Italy with a low prostate cancer incidence and in Finland, where an ancillary test (free/total

### Table 1. Description of some key characteristics of the centers in the ERSPC

<table>
<thead>
<tr>
<th>Feature</th>
<th>The Netherlands</th>
<th>Belgium</th>
<th>Sweden</th>
<th>Finland</th>
<th>Italy</th>
<th>Spain</th>
<th>Switzerland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age(^{a})</td>
<td>62</td>
<td>63</td>
<td>60</td>
<td>59</td>
<td>62</td>
<td>60</td>
<td>61</td>
</tr>
<tr>
<td>Screened ≥ once (%)</td>
<td>95</td>
<td>91</td>
<td>76</td>
<td>74</td>
<td>79</td>
<td>100</td>
<td>97</td>
</tr>
<tr>
<td>Screening rounds</td>
<td>2.4</td>
<td>1.6</td>
<td>3.5</td>
<td>2.2</td>
<td>2.2</td>
<td>1.7</td>
<td>2.5</td>
</tr>
</tbody>
</table>

\(^{a}\)At entry, within the core age group, 55–69 years.

\(^{b}\)At first screening round.

(population-based effectiveness trial). Recruitment was completed in all centers by the end of 2003. The screen interval was 4 years, except in Sweden, a 2-year interval was used. PSA determination with a cutoff of ≥3.0 ng/mL was the principal screening test (an ancillary test was used at 3.0–3.9 ng/mL in Finland). Screening was discontinued after 3 screening rounds in most centers but continued up to 5 times in the Netherlands and 10 rounds in Sweden. Institutional review board approvals were obtained in each center. The ERSPC trial is registered in Current Controlled Trials as ISRCTN49127736.

Data on deaths and emigrations were obtained from national or regional population registries. Prostate cancer cases detected outside screening (interval cases, cancers in nonparticipants and in the control arm) were identified through cancer registries. Prostate cancer deaths were ascertained in each center by causes of death certifying a uniform algorithm developed for the trial (10). In Finland, death certificates were used after the initial period, on the basis of the outcome and number of men as the offset.

The present analysis is limited to the core age group, 55 to 69 years, to improve comparability between the centers. For the same reason, follow-up was truncated at 13 years, even if data were available through 2010, potentially allowing longer follow-up in some centers. A mean length of follow-up of at least 13 years was available in all included centers, except only 10 years in Switzerland. The intention-to-screen principle was used in the analysis, with groups compared formed by random allocation (regardless of compliance).

NNI to avert one prostate cancer death was calculated as the inverse of the absolute risk difference in prostate cancer mortality between the arms \( NNI = 1/(M_s - M_c) \) to indicate the mortality reduction by screening. As an indicator of overdiagnosis, NNO was calculated as the inverse of the risk difference in cumulative incidence of prostate cancer between the trial arms \( NNO = 1/(I_s - I_c) \), which represents the absolute risk of overdiagnosis (12). Finally, NND as a measure of the overall impact (benefits and harms) is shown as the ratio of the reduction in prostate cancer mortality to the excess prostate cancer incidence (NNI/NNO). In the calculation of number needed indicators, the values are always rounded upward (to the next integer). For their interpretation, it is essential to appreciate that a smaller value indicates a larger effect.

Linear regression with weighting by number of prostate cancer deaths was used to assess the relationship between NNI and NNO between the centers, as well as in the analysis of the relation between the characteristics of the ERSPC centers (baseline risk, number of screens, proportion of positive screening tests, and proportion of screen-positive men biopsied) and the impact measures (NNI, NNO, and NND). The assumption of normal distribution was assessed using Shapiro–Wilks test. Poisson regression analysis was conducted with the numbers of cases as the outcome and number of men as the offset.

### Table 2. Numbers of men, prostate cancers and prostate cancer deaths by trial arm and center in the ERSPC

<table>
<thead>
<tr>
<th>Center</th>
<th>Men</th>
<th>Prostate cancer cases</th>
<th>Prostate cancer deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Netherlands</td>
<td>17,390</td>
<td>1,070</td>
<td>126</td>
</tr>
<tr>
<td>Belgium</td>
<td>4,255</td>
<td>521</td>
<td>23</td>
</tr>
<tr>
<td>Sweden</td>
<td>5,951</td>
<td>469</td>
<td>62</td>
</tr>
<tr>
<td>Finland</td>
<td>48,409</td>
<td>3,609</td>
<td>284</td>
</tr>
<tr>
<td>Italy</td>
<td>7,251</td>
<td>289</td>
<td>32</td>
</tr>
<tr>
<td>Spain</td>
<td>1,141</td>
<td>52</td>
<td>4</td>
</tr>
<tr>
<td>Switzerland</td>
<td>4,955</td>
<td>297</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Center</th>
<th>Men</th>
<th>Prostate cancer cases</th>
<th>Prostate cancer deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Netherlands</td>
<td>17,443</td>
<td>2,180</td>
<td>85</td>
</tr>
<tr>
<td>Belgium</td>
<td>4,307</td>
<td>417</td>
<td>18</td>
</tr>
<tr>
<td>Sweden</td>
<td>5,901</td>
<td>738</td>
<td>38</td>
</tr>
<tr>
<td>Finland</td>
<td>31,970</td>
<td>3,018</td>
<td>170</td>
</tr>
<tr>
<td>Italy</td>
<td>7,266</td>
<td>369</td>
<td>18</td>
</tr>
<tr>
<td>Spain</td>
<td>1,056</td>
<td>87</td>
<td>2</td>
</tr>
<tr>
<td>Switzerland</td>
<td>4,948</td>
<td>572</td>
<td>16</td>
</tr>
</tbody>
</table>
The highest cumulative incidence at 13 years in the screening arm was seen in Sweden and the Netherlands (12.5%), followed by Switzerland (11.6%; Table 3). The cumulative incidence in the control arm was highest in Sweden (7.9%), followed by Finland and Belgium (7.5%). The excess incidence (difference in cumulative incidence between the arms) ranged from 1.5% in Italy to 6.3% in the Netherlands. NNO ranged from 16 to 22 in the Netherlands, Switzerland, and Sweden (indicating one excess case detected per ~20 screened men) to 47 to 69 in Finland, Belgium, and Italy.

When patients treated with active surveillance were excluded, the extent of overdetection was strongly reduced in Sweden (NNO increased from 22 to 41), Finland (from 51 to 137), and Switzerland (from 18 to 34), where active surveillance was commonly used, and also to some extent in Belgium (from 47 to 62), but the difference was not substantial in the other centers (in the Netherlands from 16 to 19, Italian and Spanish results almost unaltered because of infrequent use of active surveillance).

Prostate cancer mortality within the control arm was highest in Sweden (10 per 1,000 men) followed by the Netherlands, Finland, and Belgium (~5–7 per 1,000; Table 4). Also in the screening arm, prostate cancer mortality was highest in Sweden, Finland, and the Netherlands (5–6 per 1,000). A reduction in prostate cancer mortality was achieved in all centers except Switzerland. The absolute reduction was largest in Sweden (4 per 1,000), followed by the Netherlands and Belgium (1–2 per 1,000). The mortality reduction relative to population size (NNI) ranged from 252 in Sweden to 1,821 in Finland (not defined for Switzerland).

The amount of mortality reduction per number of invited men (mean NNI pooled across the centers was 855 (SD, 576; excluding Switzerland) and the excess incidence relative to population size (mean NNO) was 36 (SD, 20; including Switzerland). The variability between centers was larger in mortality impact than excess detection (ratio of smallest to largest value, 7 for NNI vs. 4 NNO; SD, 67% of the mean for NNI vs. 55% for NNO). When the probability of benefit (absolute reduction in prostate cancer mortality, NNI) was related to the probability of harm (excess incidence, NNO), Sweden and the Netherlands showed both the largest absolute mortality reduction and the highest absolute excess incidence (Fig. 2). The other centers (Finland, Italy, Belgium, and Spain) showed both lower mortality reduction (higher NNI) and less excess incidence (larger NNO). In Switzerland, however, a high excess incidence was not balanced by mortality reduction. The proportion of screen-positive results was significantly associated with both mortality reduction (NNI) and excess detection (NNO; P = 0.001 and r² > 0.8 for both) but other center characteristics such as median age, number of screening rounds, and baseline risk (prostate cancer incidence or mortality in the control arm) were not.

The correlation coefficient for absolute mortality impact and absolute excess incidence (NNI and NNO) was 0.76 (excluding Switzerland) and linear regression coefficient (with weighting of centers by number of prostate cancer deaths) was 33 [95% confidence interval (CI), 5–62; P = 0.03 with r² = 0.65], larger mortality reduction associated with larger overdetection, indicating intensive screening resulting in both large benefits and adverse effects. Adjustment for age in linear regression analysis did not weaken the association between mortality reduction and excess incidence (NNI and NNO). If Finland was excluded, the association was stronger (r = 0.95, linear regression r² = 0.84, P = 0.02), whereas exclusion of the other major centers (Sweden or the Netherlands) did not substantially affect the result.

### Table 3. Cumulative incidence of prostate cancer by arm, with excess incidence and NNO by ERSPC center

<table>
<thead>
<tr>
<th>Center</th>
<th>Cumulative prostate cancer incidence, %</th>
<th>Excess incidence (I_S – I_C), %</th>
<th>NNO (1/excess incidence)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Netherlands</td>
<td>12.5 (ND)</td>
<td>6.2</td>
<td>16</td>
</tr>
<tr>
<td>Belgium</td>
<td>9.7 (ND)</td>
<td>7.5</td>
<td>47</td>
</tr>
<tr>
<td>Sweden</td>
<td>12.5 (ND)</td>
<td>7.9</td>
<td>22</td>
</tr>
<tr>
<td>Finland</td>
<td>9.4 (ND)</td>
<td>7.5</td>
<td>51</td>
</tr>
<tr>
<td>Italy</td>
<td>5.5 (ND)</td>
<td>4.0</td>
<td>69</td>
</tr>
<tr>
<td>Spain</td>
<td>8.2 (ND)</td>
<td>4.6</td>
<td>28</td>
</tr>
<tr>
<td>Switzerland</td>
<td>11.6 (ND)</td>
<td>6.0</td>
<td>18</td>
</tr>
</tbody>
</table>

*Expressed as an integer, with rounding upward.

### Table 4. Prostate cancer mortality by arm, with mortality reduction and NNI by ERSPC center

<table>
<thead>
<tr>
<th>Center</th>
<th>Cumulative prostate cancer mortality, %</th>
<th>Mortality reduction (M_S – M_C), %</th>
<th>NNI (1/mortality reduction)*</th>
<th>NND (NNI/NNO)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Netherlands</td>
<td>0.49 (ND)</td>
<td>0.24</td>
<td>422 (253–1,381)</td>
<td>27</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.42 (ND)</td>
<td>0.12</td>
<td>816 (243-ND)</td>
<td>18</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.64 (ND)</td>
<td>0.40</td>
<td>252 (140–1,534)</td>
<td>12</td>
</tr>
<tr>
<td>Finland</td>
<td>0.53 (ND)</td>
<td>0.05</td>
<td>1,821 (631-ND)</td>
<td>37</td>
</tr>
<tr>
<td>Italy</td>
<td>0.36 (ND)</td>
<td>0.08</td>
<td>1,198 (549-ND)</td>
<td>29</td>
</tr>
<tr>
<td>Spain</td>
<td>0.19 (ND)</td>
<td>0.16</td>
<td>621 (419-ND)</td>
<td>23</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.52 (ND)</td>
<td>−0.04</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

*Expressed as an integer, with rounding upward.

**Abbreviation:** ND, not detected.

*NNI/NNO (1/mortality reduction/1/excess incidence = excess incidence/mortality reduction).

*No mortality reduction, NNI not meaningful.

PSA ratio with a cutoff 0.16) was used at PSA range 3.0 to 3.9. The proportion of positive tests was highest (>20%) in the Netherlands and Switzerland with various side studies involving additional tests. The correlation coefficient for absolute mortality impact and absolute excess incidence (NNI and NNO) was 0.76 (excluding Switzerland) and linear regression coefficient (with weighting of centers by number of prostate cancer deaths) was 33 [95% confidence interval (CI), 5–62; P = 0.03 with r² = 0.65], larger mortality reduction associated with larger overdetection, indicating intensive screening resulting in both large benefits and adverse effects. Adjustment for age in linear regression analysis did not weaken the association between mortality reduction and excess incidence (NNI and NNO). If Finland was excluded, the association was stronger (r = 0.95, linear regression r² = 0.84, P = 0.02), whereas exclusion of the other major centers (Sweden or the Netherlands) did not substantially affect the result.
A confirmatory Poisson regression analysis also showed a significant association between excess incidence of prostate cancer and reduction in prostate cancer mortality by center ($P < 0.001$).

The ratio of mortality reduction to excess incidence (mean NND, calculated as NNI/NNO) for the 6 centers was 22, with a range of 12 to 36. It was smallest for Sweden (NND, 12), indicating that the number of excess cases per averted prostate cancer death was less than in other centers, and largest for Finland (NND, 36), with both smaller mortality reduction and excess incidence. Baseline risk (prostate cancer mortality in the control arm) showed little association with the excess incidence (NNO), nor did number of screening rounds attended, proportion of positive screening tests, or proportion of screening-positive men biopsied (for all $P > 0.2, r^2 < 0.2$ in linear regression).

**Discussion**

Our analysis of absolute measures of screening benefits (prostate cancer mortality reduction, NNI) and harms (excess prostate cancer incidence, NNO) shows that at 13 years of follow-up on average 12 to 36 excess prostate cancer have to be detected to avert one death from the disease. In comparison between the ERSPC centers, a direct correlation was observed between screening benefit and harm. This suggests that with the current screening regimens, any efforts to increase the effectiveness of screening are likely accompanied by unavoidable increase in the harmful effects. This was further supported by the association of the proportion of positive screening tests with both mortality reduction (NNI) and excess incidence (NNO). A likely explanation is the largely overlapping PSA distribution among fatal and overdiagnosed prostate cancer cases, that is, PSA does not allow a sharp distinction of lethal and inconsequential prostate cancer in the early preclinical phase of the disease due to a long lead time (detection several years earlier—and at a younger age—compared with clinical diagnosis) and length bias (increased detection of slowly growing cancers with a long detectable preclinical phase by screening).

Our results can be interpreted as suggesting a substantial overlap between cases that contribute to mortality reduction achievable by screening and those that would remain undetected in the absence of screening. This notion is supported by the fact that prediction of disease outcome in patients with low-risk features is highly uncertain and therefore distinguishing patients that would substantially benefit from treatment is challenging. This means that screen-detected cases include both tumors with indolent and progressive behavior, but we are currently unable to clearly discern the two. It is also consistent with trials showing only modest benefit from prostatectomy compared with expectant management in low-risk prostate cancer (13, 14) and very high cause-specific survival in patients treated with active surveillance (15), both indicating small advantage attainable by active treatment in men with a low-risk prostate cancer.

The wide variation between the ERSPC centers in the outcomes appears to reflect mainly differences in screening protocols, as risks and benefits were not clearly associated with the underlying risks between the populations. Yield of screening depends on the cutoff value for the screening test, number of screening rounds, and participation. These features are very likely to account for the differences in excess incidence and hence the extent of overdetection.

Of the various features, proportion of positive tests (reflecting PSA threshold and additional tests) was most closely associated with the absolute benefit and harm (NNI and NNO). There was less variation between the centers in the ratio of benefits and harms (NND, NNI/NNO) than in the mortality effect (3- vs. 7-fold). This also supports the interpretation that increased mortality reduction is linked to a higher rate of overdiagnosis. Sweden showed the most favorable overall impact (in terms of balance of mortality reduction to excess incidence, NND) and Finland the smallest, despite the fact that the two centers share similar population-based approaches, with comparable participation and the highest prostate cancer mortality (both nationally and in the control arm). Possible explanations include differences in screening intensity or contamination. The 3-fold benefit/harm ratio (NND) in Finland compared with Sweden may indicates that it is more strongly affected by mortality reduction (NNI) than excess incidence (NNO), as NNI showed a larger (7-fold difference) between the two countries. Likewise, substantial differences emerged between the volunteer-based efficacy trials, with both NNI and NNO at least twice as high for Belgium than the Netherlands. This could be attributable to the very long screening interval in Belgium (7 years).

Finding an optimal balance is probably not possible simply by maximizing only detection or minimizing overdiagnosis, but it requires consideration of several aspects of the screening program. The findings indicate that an aggressive screening protocol with a large proportion of screening-positive results is likely to increase both harms and benefits. However, the favorable overall balance of effects shown for Sweden (with an NND of 12) seems to depend mainly on the large absolute mortality reduction. One explanation could be that only an intensive or continuous screening can detect potentially lethal cancers at a curable stage. This likely leads also to increased diagnosis of indolent cases, whereas a less intense program yields fewer such cases but also a small mortality reduction.

The largest mortality reduction (smallest NNI) was seen in Sweden and the Netherlands. As they had the largest number of screening rounds, as well as the longest duration of screening (up to 20 years), this finding suggests that 3 screening rounds used in the other centers may not be sufficient to provide a similar mortality benefit. The lack of a mortality reduction in Switzerland may reflect shorter follow-up, as the screening effect emerges only after 7 to 9 years of follow-up (16), but may also be chance fluctuation (the difference being due to only two additional deaths). The absolute

**Figure 2.**

The relationship between absolute prostate cancer mortality reduction, expressed as number needed to invite (NNI = 1/risk difference) and absolute excess of cumulative prostate cancer incidence, expressed as number needed for overdiagnosis (NNO = 1/risk difference) by center in the ERSPC.
mortality impact is likely to increase with additional follow-up especially in centers continuing screening, owing to the slow progression of the disease and substantial proportion of deaths occurring only 10 to 15 years or more after diagnosis. The extent of overdetection may on the hand decreases, as the catch-up in the control arm is likely to continue for the duration of the lead time gained in the screening arm (after cessation of screening).

The results do not necessarily indicate which screening protocol is superior, because the harms and benefits are not commensurate. Most would agree that an averted prostate cancer death carries a higher impact than an overdiagnosed case. However, assessment of their relative weight is not straight forward, as it reflects value judgment (and various preferences would likely yield contrasting evaluations of the optimal balance). One approach to improving the comparability between the different endpoints is use of quality-adjusted or disability-adjusted life-years (QALYs and DALYs). We have previously modeled the long-term impact of screening and performed a cost-effectiveness analysis using QALYs as indicator of the overall effects of screening (4, 17).

For reduction of overdagnosis, multiparametric MRI in guiding prostate biopsies has shown promise in detecting primarily clinically significant prostate cancer and reducing the yield of insignificant cancers (18). Novel, powerful prognostic markers would also allow identification of cases that would not require treatment and hence reduce overtreatment. Screening is likely to provide largest absolute benefit in the subgroups with the highest risk, and correspondingly, the balance of benefit and harm is likely disadvantageous for low-risk men. Multifactorial risk stratification may allow identifying a target population with a superior screening effect.

There are some caveats in the interpretation of the results. First, the number of cases is small, as each center as a unit of observation results in a material of only 6 or 7 data points (depending on inclusion or exclusion of Switzerland and Spain). Second, all excess incidence does not necessarily constitute over-diagnosis but may also reflect lead time (earlier diagnosis due to screening) with subsequent catch-up by the control arm, as both absolute and relative excess incidence have increased with the duration of follow-up (11). Furthermore, a high prostate cancer incidence in the control arm may reflect contamination (opportunististic screening), and the low excess incidence in such context may be regarded as spurious. This is also one possible explanation for the small mortality benefit and modest excess incidence in Finland. In addition, cause of death attribution is prone to errors, which may affect the outcomes (11, 19). Overall mortality is less prone to bias, but prostate cancer is an uncommon cause of death, which precludes a reduction in overall mortality and limits the gain in life-years, as for most other screening programs (20, 21). This does not, however, preclude any impact on life-years, as modeling studies have shown up to 100 life-years saved per 1,000 screened men (17).

In conclusion, the results indicate a strong interrelation between benefits and harms of prostate cancer screening. Decision-making about prostate cancer screening needs to involve judgment of the importance of averted prostate cancer death relative to unnecessary diagnosis and harms of treatment, to gauge the tradeoffs between benefits and harms. With intensive screening, mortality reduction achievable may be larger, but the optimal balance between benefits and harms depends on weighting and valuation of the gains versus adverse effects of screening.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Disclaimer
The authors are solely responsible for the study design, collection, analysis, and interpretation of the data, writing the report and decision to submit for publication.

Authors’ Contributions
Conception and design: A. Auvinen, T.L.J. Tammela, F.H. Schröder, J. Hugosson
Development of methodology: A. Auvinen
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A. Auvinen, T.L.J. Tammela, K. Taari, M.J. Roobol, C.H. Bangma, S. Carlsson, M. Zappa, L.J. Denis, V. Nelen, M. Kwiatkowski, A. Paez, M. Lujan, J. Hugosson
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A. Auvinen, S.M. Moss, M.J. Roobol, C.H. Bangma, S. Carlsson, M. Zappa, L.J. Denis, V. Nelen, M. Kwiatkowski, A. Paez, M. Lujan, J. Hugosson
Writing, review, and/or revision of the manuscript: A. Auvinen, S.M. Moss, T.L.J. Tammela, K. Taari, M.J. Roobol, F.H. Schröder, C.H. Bangma, S. Carlsson, G. Aus, M. Zappa, L.J. Denis, V. Nelen, M. Kwiatkowski, A. Paez, M. Lujan, J. Hugosson, M. Randazzo
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): T.L.J. Tammela, M.J. Roobol, D. Pulliti, J. Hugosson
Study supervision: A. Auvinen, T.L.J. Tammela, K. Taari, L.J. Denis, M. Kwiatkowski, A. Paez, J. Hugosson
Other (functioned as international ERSPC co-coordinator): F.H. Schröder
Other (interpretation of the data and critical revision of the article for important intellectual content): S. Carlsson

Grant Support
The coordination of the ERSPC trial has been financially supported in part by the 6th Framework programme of the EU, European Commission Europe against Cancer programme; and an unconditional grant from Beckmann-Coulter-Hybrecht. The Finnish center has been financially supported by grants from the Academy of Finland [grant #260 931]; Cancer Society of Finland; Pirkanmaa Hospital District Competitive Research Funding, and Nordic Cancer Union; the Dutch Center by the Dutch Cancer Society, and the Netherland Organisation for Health Research and Development (ZonMW), Sweden by Abbott Pharmaceuticals; AF Jochnick’s Foundation; Catarina and Sven Hagemans Family Foundation, Gunvor and Ivan Svensson’s Foundation; Johannis terorden; King Gustav V Jubilee Clinic Cancer Research Foundation; Sahlgrenska University Hospital; Schering Plough; Swedish Cancer Society; and Wallac Oy; Belgium by Flemish Ministry of Welfare, Public Health and Family; Province and City of Antwerp; Public Centre for Social Welfare Antwerp. Italy by Italian League for the Fight against Cancer (ILILT); Italian Association for Cancer Research (AIRC), and National Research Council (CNR), Tuscany Region; Switzerland by The Horten Foundation; Aargau Cancer League; Swiss Cancer League; Health Department of Canton Aargau; Prostate Cancer Research Foundation; Baugarten Foundation; and the Messerli Foundation; and Spain by The Spanish “Fondo de Investigación Sanitaria.”

Received April 17, 2015; revised July 22, 2015; accepted July 26, 2015; published OnlineFirst August 19, 2015.

References
Correction: Absolute Effect of Prostate Cancer Screening: Balance of Benefits and Harms by Center within the European Randomized Study of Prostate Cancer Screening

In this article (Clin Cancer Res 2016;22:243–9), which was published in the January 1, 2016, issue of Clinical Cancer Research (1), the grant support is listed incorrectly. It should read as follows: "The coordination of the ERSPC trial has been financially supported in part by the 6th Framework programme of the EU; European Commission Europe against Cancer programme; and an unconditional grant from Beckmann-Coulter-Hybritech. The Finnish center has been financially supported by grants from the Academy of Finland (grant #260 931); Cancer Society of Finland; Pirkanmaa Hospital District Competitive Research Funding, and Nordic Cancer Union; the Dutch Center by the Dutch Cancer Society, and the Netherlands Organization for Health Research and Development (ZonMW); Sweden by Abbott Pharmaceuticals; Af Jochnick Foundation; Catarina and Sven Hagstroms Family Foundation, Gunvor and Ivan Svensson Foundation; Johanniterorden; King Gustav V Jubilee Clinic Cancer Research Foundation; Sahlgrenska University Hospital; Schering Plough; Swedish Cancer Society; and Wallac Oy; Belgium by Flemish Ministry of Welfare, Public Health and Family; Province and City of Antwerp; Public Centre for Social Welfare Antwerp; Italy by Italian League for the Fight against Cancer (LILT); Italian Association for Cancer Research (AIRC); and National Research Council (CNR), Tuscany Region; Switzerland by The Horten Foundation; Aargau Cancer League; Swiss Cancer League; Health Department of Canton Aargau; Prostate Cancer Research Foundation; Baugarten Foundation; and the Messerli Foundation; and Spain by The Spanish ‘Fondo de Investigacion Sanitaria.’ This study was also partly funded by the NIH/NCI Cancer Center Support Grant P30 CA008748.’ The authors regret this error.

Reference


Published online July 15, 2016.
doi: 10.1158/1078-0432.CCR-16-1195
©2016 American Association for Cancer Research.
Absolute Effect of Prostate Cancer Screening: Balance of Benefits and Harms by Center within the European Randomized Study of Prostate Cancer Screening

Anssi Auvinen, Sue M. Moss, Teuvo L.J. Tammela, et al.


Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-15-0941

This article cites 20 articles, 1 of which you can access for free at: http://clincancerres.aacrjournals.org/content/22/1/243.full#ref-list-1

This article has been cited by 1 HighWire-hosted articles. Access the articles at: http://clincancerres.aacrjournals.org/content/22/1/243.full#related-urls

Sign up to receive free email-alerts related to this article or journal.

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.