Underestimation of Mortality Reductions in Cancer Screening Studies: The ERSPC as a Case Study

Sous-estimation des réductions de la mortalité dans les éudes de dépistage du cancer: le ERSPC comme exemple

James Hanley
Dept. of Epidemiology, Biostatistics & Occ. Health, McGill University

Annual Meeting of the Statistical Society of Canada
Congrès annuel de la Société statistique du Canada

Mercredi 26 mai — Wednesday, May 26
Outline

• Background

• *European Randomized Study of Screening for Prostate Cancer*

• Re-analysis of ERSPC data

• Methodologic issues applicable to all screening studies
Background

- PSA-based prostate cancer (Pr Ca) screening: media coverage
  
  **NPR, 2009.10.21:** A Rethink On Prostate and Breast Cancer Screening  
  **Time, 2009.10.23:** Rethinking the benefits of breast and prostate cancer  
  **Globe and Mail, 2010.2.08:** Prostate cancer dilemma  
  **New York Times Mar 10, 2010.3.10:** The Great Prostate Mistake  
  **cyberpresse: 2010.3.13:** Cancer de la prostate: le test de détection remis en doute  
  **BMJ 2010.3.17:** Is the tide turning against the test?

- 1995 CETS (Québec) Report*: uncertain benefit / certain harms


- 2005 RCT: Radical prostatectomy > but ✓ watchful waiting in early Pr Ca

- 2009: European Randomized Study of Screening for Pr Ca (ERSPC)

In all, 5 RCTs of Screening for Prostate Cancer

<table>
<thead>
<tr>
<th>Trial:</th>
<th>Québec</th>
<th>Sweden¹</th>
<th>Sweden²</th>
<th>USA</th>
<th>Europe</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. men</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening arm</td>
<td>31,000</td>
<td>1,500</td>
<td>2,400</td>
<td>38,000</td>
<td>73,000</td>
</tr>
<tr>
<td>Control arm</td>
<td>15,000</td>
<td>7,500</td>
<td>24,000</td>
<td>38,000</td>
<td>89,000</td>
</tr>
<tr>
<td>Frequency of testing</td>
<td>?1y</td>
<td>3y</td>
<td>once</td>
<td>1y × 6</td>
<td>4y</td>
</tr>
<tr>
<td>Duration of follow-up (y)</td>
<td>11</td>
<td>15</td>
<td>15</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Actually Screened ≥ 1 time(s)</td>
<td>24%</td>
<td>78%</td>
<td>74%</td>
<td>85%</td>
<td>82%</td>
</tr>
<tr>
<td>No. Pr Ca deaths</td>
<td>153</td>
<td>20</td>
<td>53</td>
<td>92</td>
<td>214</td>
</tr>
</tbody>
</table>

¹ Norrköping ² Stockholm
“During a median follow-up of 9 years, the rate ratio in the screening group, as compared with the control group, was 0.80 (95% confidence interval [CI], 0.65 to 0.98; adjusted P=0.04). The absolute risk difference was 0.71 death per 1000 men.”

“The analysis of men who were actually screened during the first round (excluding subjects with noncompliance) provided a rate ratio of 0.73 (95% CI, 0.56 to 0.90).”
As of December 31, 2006, with an average follow-up time of 8.8 years, there were 214 prostate-cancer deaths in the screening group and 326 in the control group. Deaths that were associated with interventions were categorized as being due to prostate cancer. The \textit{adjusted rate ratio} for death from prostate cancer in the screening group was \textbf{0.80} (95\% CI, 0.65 to 0.98; \textit{P}=0.04). The Nelsen-Aalen method was used for the calculation of cumulative hazard.

*NEJM, March 2009.*
Expected ‘Response function’: Guidance from 1985 textbook

Screening in Chronic Disease

Alan S. Morrison
Figure 2-5. Changes in the disease-specific mortality rate brought about by postponement of death and by "cure" of screen-detected cases.
Cumulative & Year-specific results, if screen 0, 1, ..., 4 times, q 4y  [HYPOTHETICAL]

(A) Yearly No. of Prostate Cancer Deaths

<table>
<thead>
<tr>
<th>Year of F.U.</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Screens*</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Each arrow indicates the timing of a screen for prostate cancer.

(B) Percentage Reduction in Yearly Cause-Specific Mortality Rate

<table>
<thead>
<tr>
<th>Year of F.U.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage Reduction in Yearly Cause-Specific Mortality Rate</td>
<td>100%</td>
<td>75%</td>
<td>50%</td>
<td>25%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Fig 2

1, 2, 3, 4: No. of Screens for Prostate Cancer

One Screen for Abdominal Aortic Aneurysm
RE-ANALYSIS, with emphasis on time-specificity

- **Year-by-year mortality rate ratios**
  - pdf file containing Fig 2 → encapsulated postscript (eps) file format;
  - eps file → exact information (coordinates of line segments and dots) that statistical program, Stata, had used to draw two Nelson-Aalen cumulative hazard curves. eps file contained exact coordinates of each of 89,308 and 72,837 line segments or dots, one per man.
  - horizontal/vertical coordinates of each segment/dot → exact numbers of men being followed at each point in follow-up time, and thus at exact times of the vertical steps in curves (prostate cancer deaths).
  - size of step × number being followed → number of prostate cancer deaths at each time point
  - Numbers aggregated by year (each of 1st 12) and study arm → counts listed in new Figure.

- **Moving averages** to reduce the statistical noise (deaths in moving 3-year intervals)
- **Smooth curve** for rate ratio function (data bins 0.2 y wide).
Year-specific prostate cancer mortality ratios

(A) Cumulative Prostate Cancer Mortality

(B) Prostate Cancer Mortality Rate Ratio ($S \div C$)

Yearly Numbers of Prostate Cancer Deaths in Control (C) and Screening (S) Arms:

- C: 2 5 6 21 27 26 28 29 27 8
- S: 69K 69K 68K 62K 79K 76K 64K 59K 61K 57K

Numbers of Men Being Followed at Mid-Year in Control (C) and Screening (S) Arms:

- C: 2 5 6 21 27 26 28 29 27 8
- S: 73K 72K 71K 68K 66K 64K 64K 61K 61K 59K

Percentage Reduction in Year-Specific Prostate Cancer Mortality Rate

(\%) 

- 0% 
- 25% 
- 50% 
- 75% 
- 67% 

Follow-Up Year: 1 2 3 4 5 6 7 8 9 10 11 12
Interpretation

- After an expected delay (data indicate $\approx$ 7 years), the prostate mortality reductions that become evident in years 9 and beyond are statistically significant and considerably greater than the reported 20% reduction in the rate of prostate cancer deaths.
- The best (ML) estimate is that, although the rate ratio became non-null starting at $\approx$ 7 years, the steady state reduction has not yet been reached: the point estimate so far is a sustained 67% reduction (80%CI 30% to 89%) beginning at year 12.
- Numbers of deaths are not sufficient to establish its timing and magnitude more precisely. (Data cutoff: Dec 2006)
Implications - substantive

- ** downsides** of PSA-based prostate cancer screening: well documented and long since agreed upon.

  Even if screening could achieve a sustained reduction of 67%, (or even 77 or 87%!) the very low prostate mortality rates in the control group means that the small absolute reductions would be achieved at what some people would consider to be an unacceptable cost. (So far, only 326 or 0.36% of the 89,353 men in control group have died of prostate cancer; the number will approximately triple by follow-up year 20.)

- ‘upsides’: 5 RCTs; 23 years; 321,000 men; 10 countries average f.-u. ranging from 7-15 years.
  
  - 4 have virtually no resolving power.
  
  - **ERSPC**: much larger Δ in screening activity b/w 2 arms → considerably greater **resolving power**.
  
  - Must measure signal in f.-u. window where probably strongest → collect additional data.

  Casual reader of ERSPC report should not conclude that best we can expect from PSA screening is a reduction in prostate cancer mortality of 20%.

  Re-analysis: if screening is carried out for several years, and **if f.-u. pursued into window where reduction in mortality becomes manifest**, reduction to be seen there will be 50-60%.

  ERSPC report published March 2009, but f.-u. ended in Dec 2006, just when pattern had begun to emerge. **Not possible to put precise statistical bounds** on this reduction.

  Prostate cancer deaths from 2007 onwards crucial to more precisely measure the reduction achieved.
Implications - Methodologic

Time-specificity...

- Avoids dilution caused by averaging
  - 7 years of (expected) non-reductions with
  - 5 years of progressively larger reductions
- With current data, imprecise estimates: fixable.
- Follows intention to treat principle
- With objective curve-fitting...
  - avoid need to “pre-specify” when reduction reaches steady state
  - data themselves inform us about two critical parameters that determine ‘response curve’ (i.e., timing & extent of prostate cancer mortality reduction caused by screening).
Only an ineffective cancer screening program can yield proportional hazards!

- **Time-specific** analysis only necessary when effect of intervention is **delayed**, as in case of Pr Ca screening.
- Screening for abdominal aneurysms produces an **immediate and sustained reduction** in mortality from ruptured aneurysms; **cumulative** mortality, in this case, **fully captures benefit** of screening.
- Recognition of **difference between interventions with immediate and delayed effects** should prompt similar re-analyses of data from trials of screening in other cancers, and similar analyses in yet-to-be reported cancer screening trials.
IMPLICATIONS: data-analysis, meta-analyses, public health

- ‘Response Curve’ in any one RCT is a function of the number and timing of screens [& compliance]
- Time-specificity in data-analysis is paramount
- No common parameter (response curve) to meta-analyze: trials not uniform w.r.t. number and timing of screens
- REAL Q: reduction with SUSTAINED SCREENING?
- How about using nadir of response curve?
Timing of Screening Effects
(as seen in cumulative cause-specific mortality curves)

Cumulative Cause-Specific Mortality

Abdominal Aortic Aneurysms
(One-off Screening, MASS)

Prostate Cancer
(q 4y, ERSPC )

Supp Fig. A
Timing of cholesterol reductions produced by statins

3 dogs at 20 mg/kg/day; 3 at 50 mg/kg/day

3 monkeys at 50
Timing of cholesterol reductions produced by statins

Humans
Acknowledgments
Mammographic screening: no reliable supporting evidence?

Olli S Miettinen, Claudia I Henschke, Mark W Pasmantier, James P Smith, Daniel M Libby, David F Yankelevitz

Much confusion is being generated by the conclusion of a recent review that “there is no reliable evidence that screening for breast cancer reduces mortality.” In that review, however, there was no appreciation of the appropriate mortality-related measure of screening’s usefulness; and correspondingly, there was no estimation of the magnitude of this measure. We take this measure to be the proportional reduction in case-fatality rate, and studied its magnitude on the basis of the only valid and otherwise suitable trial. We found reliable evidence of fatality reduction, apparently substantial in magnitude.

Lancet 2002; 359: 404–06
FUNDING & CO-ORDINATES

Natural Sciences and Engineering Research Council of Canada

Le Fonds québécois de la recherche sur la nature et les technologies

James.Hanley@McGill.CA

http://www.biostat.mcgill.ca/hanley
References

1. Hanley JA. Mortality reductions produced by sustained prostate cancer screening have been underestimated. In press, Journal of Medical Screening.


