Statistical Models for Bias and Overdiagnosis in Prostate Cancer Screening

Tony Hsiu-Hsi Chen

2007/05/09
Evaluation by cumulative mortality curve: Swedish Two-county trial

(1) RR 0.68 (0.59 – 0.80)

(2) Lead-time bias problem
Evaluation by survival curve

Cumulative Survival

Years of follow-up

Clinically-detected cases
Screen-detected cases
(A) Lead-time Bias

Case 1: Screen-detected case

Birth 45 LT 55 RST 60 70 No benefit!

Case 2: Clinically-detected case

Birth 45 RST 60 70

Case 3: Screen-detected case

Birth 45 LT 55 RST 60 80 Benefit by screening!

◇: Date of biological onset (entering preclinical phase);
☆: date of surfacing to clinical phase due to occurrence of symptoms or signs;
☼: hypothetical date of surfacing to clinical phase if no screen had taken place;
☠: date of death from prostate cancer;
■: date of detected by screen;
□: hypothetical date of screen if no clinical symptoms or signs of prostate cancer had taken place;
■: date of screen as normal;
OST: observed survival time; LT: lead-time; RST: real survival time.
Length Bias

(B)

Case 4: Clinically-detected case

Birth 45 50 55

Surface to clinical phase before 1st screen!

Case 5: Screen-detected case

Birth 45 55 65

(C)

Case 6: Clinically-detected case

Birth 55 57 58 59

Surface to clinical phase before subsequent screen!

Case 7: Screen-detected case

Birth 55 57 59 61
Survival of Screen- and Clinically-detected Prostate Cancer in a Population-based Screening Trial, Adjusted for Lead-time and Length Bias

(Part I)
Study Aims

• To compare survival between cancers detected clinically and at screening
  – Both uncorrected and corrected for major sources of bias

• To estimate the magnitude of these biases
Patients and Methods

• Data Sources
  ■ Finnish Prostate Cancer Screening Trial
  ■ 1st round: 1996-1999
    Screening arm: 32,000 men aged 55-67
    Control arm: 48,458 men aged 55-67
  ■ 2nd round: 2000-2003
  ■ Prostate cancer cases deriving from 1996-2002

• Follow-up
  ■ All patients were followed up until death, migration, or until end of 2002
Patients and Methods

• Case Definition
  ■ Screen-detected cases:
    ■ prostate cancers detected
      ■ In the 1\textsuperscript{st} round of screening (prevalence)
      ■ In the 2\textsuperscript{nd} round of screening (incidence)
  ■ Clinically-detected cases:
    ■ Cancers from control arm
Uncorrected Survival Analysis

- Prostate cancer specific survival analysis was performed by life-table method without adjustment for biases
- The cumulative risk of death from prostate cancer and its 95% CI were calculated
Lead-time Bias Adjustment

Cumulative survival curve after correcting lead-time can be calculated as

\[ S(t) = \exp \left( \int_0^t - \lambda_2(s) \, ds \right) \]
Lead-time Bias Adjustment--Likelihood

• Censored:
\[ P_0(t) = P_{11}(t) + P_{12}(t) \]
\[ = \exp\left[-\left(\hat{\lambda}_1 + \lambda_3\right)t\right] + \int_0^t \hat{\lambda}_1 \cdot \exp\left[-\left(\hat{\lambda}_1 + \lambda_3\right)s\right] \cdot \exp\left[-\lambda_{20}(t-s)\right] ds \]

• Death from prostate cancer:
\[ P_1(t) = dP_{13}(t) \]
\[ = \int_0^t \lambda_1 \cdot \exp\left[-\left(\hat{\lambda}_1 + \lambda_3\right)s\right] \cdot \exp\left[-\lambda_{20}(t-s)\gamma\right] \cdot \lambda_{20} \gamma^{(t-s)\gamma-1} ds \]

• Death from other causes of death:
\[ P_2(t) = dP_{14}(t) \]
\[ = \lambda_3 \cdot \exp\left[-\left(\hat{\lambda}_1 + \lambda_3\right)t\right] \]
Adjustment for Length Bias

State 0: Normal

$\lambda_3$

State 4: Other Causes of Death

$I(t) \rightarrow$ State 1: Detectable Preclinical Phase

$\lambda_1(t) \rightarrow$ State 2: Clinical Phase

$\lambda_2(t) \rightarrow$ State 3: Prostate Cancer Death
Adjustment for Length Bias (1)

(1) Short PCDP: incorporate interval cancer

\[ dP_{02}(t_a, t_b) = \int_{t_a}^{t_b} \lambda_1(s) \cdot S_0(t_a, s) \cdot S_1(s, t_b) \cdot \lambda_1(t_b - s) ds \]

\( dP_{02} \): derivative of probability of being normal at \( t_a \) and being diagnosed clinically due to symptoms and signs at \( t_b \).

\( t_a \): age at previous normal screen

\( t_b \): age at clinical detected as interval cases

\( S_0 \): survival functions of staying in state 0 (normal)

\( S_1 \): survival function of staying in state 1 (preclinical phase)
Adjustment for Length Bias (2)

(2) Long PCDP: 1st screen detected cases has to be conditional on time of surfacing to clinical phase > time to screen

- 1st screen: Normal:

\[ P_{f1}(t_a, t_b) = \frac{P_{01}(t_a, t_b)}{P_{00}(t_a, t_b) + P_1(t_a, t_b)} \]

- 1st screen: Prostate Cancer:

\[ P_{f0}(t_a, t_b) = \frac{P_{00}(t_a, t_b)}{P_{00}(t_a, t_b) + P_1(t_a, t_b)} \]
Adjustment for Length Bias

State 0: Normal

State 1: Detectable Preclinical Phase

Weibull($\lambda_{00}, \gamma_0$)

$P_{00}(t_a, t_b) = \exp(\lambda_{00}t_a^{\gamma_0} - \lambda_{00}t_b^{\gamma_0})$

$P_{01}(t_a, t_b) = \int_{t_a}^{t_b} \left[ \lambda_{00} \gamma_0 s^{\gamma_0-1} \cdot \exp(\lambda_{00}t_a^{\gamma_0} - \lambda_{00}s^{\gamma_0}) \cdot \left[ \exp - \int_s^{t_b} (\lambda_{10} \gamma_1 v^{\gamma_1-1} + \lambda_{3}) dv \right] ds \right]$  

$dP_{02}(t_a, t_b) = \int_{t_a}^{t_b} \left[ \lambda_{00} \gamma_0 s^{\gamma_0-1} \cdot \exp(\lambda_{00}t_a^{\gamma_0} - \lambda_{00}s^{\gamma_0}) \cdot \left[ \exp - \int_s^{t_{b_1}} (\lambda_{10} \gamma_1 v^{\gamma_1-1} + \lambda_{3}) dv \right] \cdot \lambda_{10} \gamma_1 t_{b_1}^{\gamma_1-1} \right] ds$

• $\lambda_{00}, \gamma_0, \lambda_{10}, \gamma_1, \lambda_{20}$, and $\gamma_2$ can be obtained by MLE
Adjustment for Overdiagnosis with a mover-stayer model

- Mover (M): case with potential for progressing to the symptomatic stage

- Stayer (S): case has no potential of progressing to the symptomatic stage

Weibull ($\lambda_{20}$, $\gamma_2$)

Estimate obtained from previous model adjusted for both lead-time and length bias
Further Adjustment for Overdiagnosis

- Mover (M): Proportion of mover is \((1-f)\)

\[
M = \begin{pmatrix}
- \left( \hat{\lambda}_1 + \hat{\lambda}_3 \right) & \hat{\lambda}_1 & 0 & \lambda_3 \\
0 & - \lambda_2(t) & \lambda_2(t) & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

- Stayer (S): Proportion of stayer is \(f\)

\[
S = \begin{pmatrix}
- \lambda_3 & 0 & 0 & \lambda_3 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

- Given \(f\), \(\lambda_20\) and \(\gamma_2\) were estimated by MLE method.
Results

• 1,784 prostate cancer cases were ascertained in 1996-2002
  ■ Intervention arm: 987 cases
  ■ Control arm: 797 cases

• Follow-up time
  ■ Average: 2.7 years (SD ± 1.8)
  ■ Median: 2.5 years
*Taking the average age at diagnosis among screen-detected cases, 64 years old, as example.
Relative death rate after correcting lead-time, length-bias, and over-diagnosis

<table>
<thead>
<tr>
<th>Years of follow-up</th>
<th>Lead-time Adjustment</th>
<th>Both length bias and lead-time adjustment†</th>
<th>Length bias and lead-time adjustment, given 40% overdiagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR</td>
<td>95% CI</td>
<td>RR</td>
</tr>
<tr>
<td>1</td>
<td>0.67</td>
<td>0.35-1.30</td>
<td>0.61</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>0.47-1.19</td>
<td>0.73</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.67-1.50</td>
<td>0.99</td>
</tr>
<tr>
<td>4</td>
<td>0.76</td>
<td>0.53-1.10</td>
<td>0.76</td>
</tr>
<tr>
<td>5</td>
<td>0.81</td>
<td>0.57-1.16</td>
<td>0.84</td>
</tr>
<tr>
<td>6</td>
<td>0.72</td>
<td>0.51-1.01</td>
<td>0.74</td>
</tr>
<tr>
<td>7</td>
<td>0.69</td>
<td>0.49-0.96</td>
<td>0.71</td>
</tr>
</tbody>
</table>

*RR, relative risk; CI, confidence interval.
†Taking the average age at diagnosis among screen-detected cases, 64 years old, as example.
Final Remark

- In the Finnish PSA-based trial, the crude estimate of 83% benefit was reduced to non-significant 6% gain after correcting for lead-time and length bias as well as overdiagnosis.
- It seems too early to show the benefit of PSA screening.
A Markov Model for Over-diagnosis in Prostate Cancer Screening With PSA: Finnish Prostate Cancer Screening Trial (Part II)
Stop screen design-- Overdiagnosis

Catch-up time = mean sojourn time
= average duration of tumour staying in the preclinical phase
A thorny issue in estimating the magnitude of over-diagnosis in prostate cancer screening
Factors related to over-diagnosis

- Age at screening
- Mean sojourn time
- Sensitivity for progressive and non-progressive tumour
Study aim

• The aim of the current study is to quantify overdiagnosis in the Finnish Prostate Cancer Screening Trial using a novel stochastic model of prostate cancer by taking related factors mentioned above.
Patients and Methods

• Data source
  - Finnish prostate cancer screening trial.
  - During 1996 to 1999, 80,458 men aged 55, 59, 63, and 67 years entered the trial.
  - 32,000 in the screening group.
  - 4 years inter-screening interval until age 71 years.

• Follow-up
  - We followed the subjects till end of 2002 or till they received their second screen
Patients and Methods

• Case Definition
  ■ Screen-detected cases:
    ■ PSA $\geq 4$: consisting of a DRE, TRUS and prostate biopsy.
    ■ $3.0 \leq PSA \leq 3.9$: suspicious finding

• Control group
  ■ The remaining 48,458 men comprise the control group
  ■ Cancers which occurred between two screening rounds and the clinical cancers in the control group:

=> Identified by record linkage with the Finnish Cancer Registry.
A four-state Markov Model

Figure 1. Four-state Markov model depicting the natural history of prostate cancer.
• Sensitivity parameter

• The likelihood for each event or mode of observation can be expressed by the transition probability, $P_{ij}(t)$, and taking into account the program sensitivity (S).
## Likelihood function

<table>
<thead>
<tr>
<th>Mode of Observations</th>
<th>Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Screening Arm</strong></td>
<td></td>
</tr>
<tr>
<td>First Screen</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>[ \frac{P_{11} \text{(age)}}{P_{11} \text{(age)} + P_{12} \text{(age)} + P_{14} \text{(age)}} + \frac{[P_{12} \text{(age)} + P_{14} \text{(age)}]}{P_{11} \text{(age)} + P_{12} \text{(age)} + P_{14} \text{(age)}} \cdot (1 - S) ]</td>
</tr>
<tr>
<td>Screen-detected Cancers</td>
<td>[ \frac{[P_{12} \text{(age)} + P_{14} \text{(age)}]}{P_{11} \text{(age)} + P_{12} \text{(age)} + P_{14} \text{(age)}} \cdot S ]</td>
</tr>
<tr>
<td>Second Screen</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>[ F_{TN} \cdot {P_{11} \text{(t)} + [P_{12} \text{(t)} + P_{14} \text{(t)}] \cdot (1 - S) } ]</td>
</tr>
<tr>
<td>Screen-detected Cancers</td>
<td>[ F_{TN} \cdot [P_{12} \text{(t)} + P_{14} \text{(t)}] \cdot S + F_{FN1} \cdot P_{22} \text{(t)} + F_{FN2} \cdot P_{44} \text{(t)} ]</td>
</tr>
<tr>
<td>Interval Cancers</td>
<td>[ F_{TN} \cdot dP_{13} \text{(t)} + F_{FN1} \cdot dP_{23} \text{(t)} ]</td>
</tr>
</tbody>
</table>
Control Arm

Normal
\[ \frac{P_{11}(\text{age})}{P_{11}(\text{age}) + P_{12}(\text{age})} \times [P_{11}(t) + P_{12}(t)] + \]
\[ \frac{P_{12}(\text{age})}{P_{11}(\text{age}) + P_{12}(\text{age})} \times P_{22}(t) \]

Clinical Cases
\[ \frac{P_{11}(\text{age})}{P_{11}(\text{age}) + P_{12}(\text{age})} \times dP_{13}(t) + \]
\[ \frac{P_{12}(\text{age})}{P_{11}(\text{age}) + P_{12}(\text{age})} \times dP_{23}(t) \]
Progressive tumour with data from control arm

Figure 2. Three-state Markov model depicting the natural history of prostate cancer.
Expectation-Maximum (E-M) Algorithm

• Maximum Likelihood- estimated the maximum likelihood estimates for $\lambda_2$ and $\lambda_3$ and S in the four-state Markov model (Figure 1).

• Expectation: $\lambda_1$ was obtained by using the expected equation from control arm that non-progressive tumour would not be observed.
**Expectation Step**

Expected \( (E_{CN}) \)

\[
E_{CN} = N_{CN} \times \left[ \frac{P_{11}(\text{age})}{P_{11}(\text{age}) + P_{12}(\text{age})} \times dP_{13}(t) + \frac{P_{12}(\text{age})}{P_{11}(\text{age}) + P_{12}(\text{age})} \times dP_{23}(t) \right] = C_{CN},
\]

\( E_{CN} \): the expected number of prostate cancers in the control arm given time \( t \)
\( N_{CN} \): the total number of subjects in the control arm
\( C_{CN} \): the observed number of prostate cancers in the control arm given time \( t \)

Then we obtained the estimated \( \lambda_{1}^{(1)} \) from this equation and re-estimated \( \lambda_{2}^{(2)}, \lambda_{3}^{(2)}, \text{and } S^{(2)} \).

Repeat the procedure of M-E steps until the convergence of parameters.
Model Validation

• The model goodness of fit was evaluated based on the Pearson’s chi-square test using the formula as follow.

\[ \chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \]
Proportion of overdiagnosis

\[
P_{\text{over}}(t) = \frac{P_{14}(t) \times \text{sen}}{P_{11}(t) + P_{12}(t) + P_{14}(t)} \cdot \frac{P_{14}(t) \times \text{sen}}{P_{11}(t) + P_{12}(t) + P_{14}(t)} + f_1 \times P_{13}(4) + f_2 \times P_{13}(4)
\]

where

\[
f_1 = \frac{P_{11}(t)}{P_{11}(t) + [P_{12}(t) + P_{14}(t)] \times (1 - \text{sen})}
\]

and

\[
f_2 = \frac{P_{12}(t) \times (1 - \text{sen})}{P_{11}(t) + [P_{12}(t) + P_{14}(t)] \times (1 - \text{sen})}
\]

- \(P_{\text{over}}(t)\): The proportion of over-diagnosed prostate cancers, varies with time (age or inter-screening interval),
<table>
<thead>
<tr>
<th>Detection Mode</th>
<th>Number of Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Screening Arm</strong></td>
<td></td>
</tr>
<tr>
<td>First Screen</td>
<td>23,142</td>
</tr>
<tr>
<td>Normal</td>
<td>22,478</td>
</tr>
<tr>
<td>Screen-detected Cancers</td>
<td>664</td>
</tr>
<tr>
<td>Second Screen</td>
<td>16,129</td>
</tr>
<tr>
<td>Normal</td>
<td>15,758</td>
</tr>
<tr>
<td>Screen-detected Cancers</td>
<td>371</td>
</tr>
<tr>
<td>Interval Cancers</td>
<td>103</td>
</tr>
<tr>
<td><strong>Control Arm</strong></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>47,721</td>
</tr>
<tr>
<td>Clinical Cases</td>
<td>569</td>
</tr>
</tbody>
</table>
Table 4 Estimates of the parameters of the natural history.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimates</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>0.003256</td>
<td>0.002989~0.003524</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.1668</td>
<td>0.0980~0.2357</td>
</tr>
<tr>
<td>Mean Sojourn Time</td>
<td>6.00</td>
<td>4.24~10.20</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.0004465</td>
<td>0.0003228~0.0005701</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>61.39%</td>
<td>57.28%~65.49%</td>
</tr>
</tbody>
</table>

$\chi^2 = 1.27$, degree of freedom=2, $p$-value=0.5309.
Proportion of Overdiagnosis by sensitivity, age, and round of screen

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Age</th>
<th>First</th>
<th>Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>30%</td>
<td>40</td>
<td>26.1%</td>
<td>45.3%</td>
</tr>
<tr>
<td>50%</td>
<td>50</td>
<td>30.9%</td>
<td>51.1%</td>
</tr>
<tr>
<td>60%</td>
<td>60</td>
<td>35.4%</td>
<td>56.0%</td>
</tr>
<tr>
<td>70%</td>
<td>70</td>
<td>39.4%</td>
<td>60.1%</td>
</tr>
<tr>
<td>80%</td>
<td>80</td>
<td>43.1%</td>
<td>63.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Age</th>
<th>First</th>
<th>Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>60%</td>
<td>40</td>
<td>37.2%</td>
<td>36.6%</td>
</tr>
<tr>
<td>50%</td>
<td>50</td>
<td>42.9%</td>
<td>41.7%</td>
</tr>
<tr>
<td>60%</td>
<td>60</td>
<td>47.9%</td>
<td>46.2%</td>
</tr>
<tr>
<td>70%</td>
<td>70</td>
<td>52.2%</td>
<td>50.2%</td>
</tr>
<tr>
<td>80%</td>
<td>80</td>
<td>56.0%</td>
<td>53.7%</td>
</tr>
</tbody>
</table>

Crude CI: 1.95
Future work

• Modeling the sensitivity as a function of age, PSA level at baseline, or other markers - identify an individual with high potential of over-diagnosis

By taking over-diagnosis into account,

• extend the model to elucidate the natural history defined by gleason score and stage
• do Cost-effectiveness analysis for PSA screening