Modelling of Cancer Morbidity Risk in a Bayesian Framework

Ayşe Arık

Heriot-Watt University, Edinburgh

Research Fellow

joint work with George Streftaris and Erengül Dodd

under Modelling, Measurement and Management of Longevity and Morbidity Risk research programme funded by the ARC

EAJ 2018 Conference, 10 September 2018
Outline

1. Aim of the Study
2. Data
3. All Cancer Incidences
4. Lung Cancer
5. Prostate Cancer
6. Breast Cancer
7. A Comparison bw Population Incidence Rates and CII Rates
8. Summary
Aim of the Study

The main purpose of the study

1. Investigate the dynamics of various types of population cancer risk
   - under a Bayesian setting by employing generalised linear model structures
   - using cancer registration numbers in England

2. Compare population incidence rates with the CII rates
Cancer registration data for England provided by Office for National Statistics (ONS)

1. International Statistical Classification of Diseases (ICD): ICD 8, ICD 9 and ICD 10
2. age from zero to 95+
3. year from 1971 through 2015
4. region ?
5. gender
All Cancer Incidences in England

⇒ Bayesian model

\[ C_i | \theta_i \sim \text{Poisson}(E_i \theta_i) \]
\[ \theta_i \sim \text{Lognormal}(\mu_i, \sigma^2) \]
\[ \mu_i = \beta_0 + \beta_{1,j}x_j + \beta_{2,k}y_k + \beta_{3,g}x_g + \beta_{4,j,g}x_jx_g \]
\[ \sigma^2 \sim \text{Inv.Gamma}(1, 0.001) \]
\[ \beta's \sim \text{Normal}(0, 10^4), \]

\[ i = 1, \ldots, 1080; \ k = 1, \ldots, 36; \ j = 1, \ldots, 15 \text{ and } g = 1, 2. \]
Actual and fitted incidence rates of all cancer types without carcinoma in situ fitted incidence rates as solid lines & crude rates as dots

- An increasing trend for all ages.
incidence rates for females from ages 47 to 62 for years 1980, 1995, 2000 and 2015 fitted incidence rates as solid lines

- Ageing is a fundamental factor for cancer.
Lung Cancer in England

⇒ Bayesian model

\[ C_i | \theta_i \sim \text{Poisson}(E_i \theta_i) \]
\[ \theta_i \sim \text{Lognormal}(\mu_i, \sigma^2) \]
\[
\mu_i = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 y + \beta_5 xy + \beta_{6,g} x_g + \beta_{7,g} x_g y + \beta_{8,g} x_g x
\]
\[ \sigma^2 \sim \text{Inv.Gamma}(1, 0.001) \]
\[ \beta's \sim \text{Normal}(0, 10^4), \]

\( i = 1, \ldots, 648 \) and \( g = 1, 2. \)

<table>
<thead>
<tr>
<th>Model</th>
<th>age</th>
<th>year</th>
<th>age:year</th>
<th>gender</th>
<th>gender:year</th>
<th>( \text{age}^2 )</th>
<th>( \text{age}^3 )</th>
<th>gender:age</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Cont.</td>
<td>Cont.</td>
<td>✔</td>
<td>Factor</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>7148</td>
</tr>
</tbody>
</table>
Lung Cancer in England

Actual and fitted incidence rates of malignant neoplasm of trachea, bronchus and lung

fitted incidence rates as solid lines & crude rates as dots

- A narrowing gap bw males & females with different trends due to different smoking patterns.
fitted incidence rates as solid lines

- An opposite trend for males and females.
Prostate Cancer in England

⇒ Bayesian model

\[ C_i|\theta_i \sim \text{Poisson}(E_i\theta_i) \]
\[ \theta_i \sim \text{Lognormal}(\mu_i, \sigma^2) \]

\[
\mu_i = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 y + \beta_5 y^2 + \beta_6 y^3 + \beta_7 y^4 + \beta_8 xy
\]

\[ \sigma^2 \sim \text{Inv.Gamma}(1, 0.001) \]
\[ \beta's \sim \text{Normal}(0, 10^4), \]

\[ i = 1, \ldots, 324. \]

<table>
<thead>
<tr>
<th>Model</th>
<th>age</th>
<th>year</th>
<th>age:year</th>
<th>year^2</th>
<th>year^3</th>
<th>year^4</th>
<th>age^2</th>
<th>age^3</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>cont.</td>
<td>cont.</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>3580</td>
</tr>
</tbody>
</table>
Prostate Cancer in England

Actual and fitted incidence rates of malignant neoplasm of prostate fitted incidence rates as solid lines & crude rates as dots

- PSA testing started to be used in the UK in the early 1990s.
- The second increase after 2006 with a peak in 2013.
Breast Cancer in England

⇒ Bayesian model

\[ C_i | \theta_i \sim \text{Poisson}(E_i \theta_i) \]
\[ \theta_i \sim \text{Lognormal}(\mu_i, \sigma^2) \]
\[ \mu_i = \beta_0 + \beta_{1.j} x_j + \beta_2 y + \beta_{3.j} x_j y \]
\[ \sigma^2 \sim \text{Inv.Gamma}(1, 0.001) \]
\[ \beta \text{'s} \sim \text{Normal}(0, 10^4), \]

\( i = 1, \ldots, 540 \) and \( j = 1, \ldots, 15. \)

<table>
<thead>
<tr>
<th>Model</th>
<th>age</th>
<th>year</th>
<th>age:year</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>factor</td>
<td>cont.</td>
<td></td>
<td>5616</td>
</tr>
</tbody>
</table>
Actual and fitted incidence rates of malignant neoplasm of breast
fitted incidence rates as solid lines & crude rates as dots

- NHSBSP began in 1988, targeted women between ages 50 and 64.
- After 2005, screening offered for women between ages 50 and 70.
A Comparison between Population Incidence Rates and CII Rates

- incidence rates > CII rates
- Cancer rates are higher for socio-economically deprived groups.
Summary

- An increasing trend for all cancer types for all age groups.
- An adverse trend in lung cancer for different genders.
- An increasing trend both in prostate and breast cancers.
- Population incidence rates are higher than the insurance rates.
Sources

1. CII claims settled bw 1999 and 2005 by Continuous Mortality Investigation (CMI) in the UK.
Thank You!

Questions?

e-mail: A.ARIK@hw.ac.uk