Mortality and Deprivation

Torsten Kleinow
joint work with Jie Wen and Andrew J.G. Cairns

Heriot-Watt University, Edinburgh
Actuarial Research Centre, IFoA

ARC workshop - Edinburgh - April 2018
The IMD is a weighted combination of seven indices of deprivation:

- Income (22.5%)
- Employment (22.5%)
- Education (13.5%)
- Health (13.5%)
- Crime (9.3%)
- Barriers to Housing and Services (9.3%)
- Living environment (9.3%)

source: GOV.UK
Index of Multiple Deprivation (IMD) areas

Most deprived 10%
Least deprived 10%
Data

- We consider mortality rates for males in England for the ten IMD deciles (2015).
- ages: 40-89, years: 2001-2015
- source: Office for National Statistics
Model for the Number of Death in Different Groups

\[ D_{xti} \sim \text{Poisson}(m_{xti}E_{xti}) \]

For each period (calendar year) \( t \), age \( x \) and socio-economic group \( i \) we have

- \( D_{xti} \): Number of deaths,
- \( E_{xti} \): Central exposure-to-risk
- \( m_{xti} \): force of mortality
Model for the Number of Death in Different Groups

\[ D_{xti} \sim \text{Poisson} \left( m_{xti} E_{xti} \right) \]

For each period (calendar year) \( t \), age \( x \) and socio-economic group \( i \) we have

\( D_{xti} \): Number of deaths, \\
\( E_{xti} \): Central exposure-to-risk \\
\( m_{xti} \): force of mortality

Death rate: \( D_{xti} / E_{xti} \)
Model for the Number of Death in Different Groups

\[ D_{xti} \sim \text{Poisson} \left( m_{xti} E_{xti} \right) \]

For each period (calendar year) \( t \), age \( x \) and socio-economic group \( i \) we have

- \( D_{xti} \): Number of deaths,
- \( E_{xti} \): Central exposure-to-risk
- \( m_{xti} \): force of mortality

Death rate: \( D_{xti} / E_{xti} \)

Aim: to compare different models for the force of mortality \( m_{xti} \).
Model for the Number of Death in Different Groups

\[ D_{xti} \sim \text{Poisson}(m_{xti}E_{xti}) \]

For each period (calendar year) \( t \), age \( x \) and socio-economic group \( i \) we have

- \( D_{xti} \): Number of deaths,
- \( E_{xti} \): Central exposure-to-risk
- \( m_{xti} \): force of mortality

Death rate: \( D_{xti}/E_{xti} \)

Aim: to compare different models for the force of mortality \( m_{xti} \).

We define socio-economic groups with reference to the Index of Multiple Deprivation for England.
Death rates by IMD decile

male mortality in year 2015

log death rate

age

Jie Wen, AJG. Cairns, T. Kleinow: Mortality and Deprivation
Death rates by IMD decile

male mortality in year 2015

male mortality in year 2001

log death rate

age

Jie Wen, AJG. Cairns, T. Kleinow: Mortality and Deprivation
Death rates by IMD decile

male mortality at age 65

log death rate

calendar year
All considered models are variants of group specific Lee-Carter type models with the extension to a second age-period effect by Renshaw & Haberman (2003):

$$\log m_{xti} = \alpha_{xi} + \beta_{xi}^{1} \kappa_{ti}^{1} + \beta_{xi}^{2} \kappa_{ti}^{2} + \gamma_{ci}$$

where $c = t - x$ is the cohort (year of birth).
All considered models are variants of group specific Lee-Carter type models with the extension to a second age-period effect by Renshaw & Haberman (2003):

\[
\log m_{xti} = \alpha_{xi} + \beta_{xi}^{1} \kappa_{1}^{1}_{ti} + \beta_{xi}^{2} \kappa_{2}^{2}_{ti} + \gamma_{ci}
\]

where \( c = t - x \) is the cohort (year of birth).
Specific versions include models with:

- **common age effect**: \( \alpha_{xi} = \alpha_{x} \)
All considered models are variants of group specific Lee-Carter type models with the extension to a second age-period effect by Renshaw & Haberman (2003):

\[ \log m_{x t i} = \alpha_{x i} + \beta_{x i}^1 \kappa_{t i}^1 + \beta_{x i}^2 \kappa_{t i}^2 + \gamma_{c i} \]

where \( c = t - x \) is the cohort (year of birth).

Specific versions include models with:

- **common age effect**: \( \alpha_{x i} = \alpha_x \)
- **non-parametric common age effects**: \( \beta_{x i}^k = \beta_x^k \) (Kleinow, 2015)
- **fixed age effects**: constant \( \beta_{x i}^1 \) and linear \( \beta_{x i}^2 = x - \bar{x} \), where \( \bar{x} \) is the mean age in the data set. (Plat, 2009)
- **common period effects**: \( \kappa_{t i}^k = \kappa_{t}^k \) (Li and Lee, 2005)
- **group specific trends in common period effects**: \( \kappa_{t i}^k = \kappa_{t}^k + \eta_i (t - \bar{t}) \) and variations with and without cohort effects.
Models

All considered models are variants of group specific Lee-Carter type models with the extension to a second age-period effect by Renshaw & Haberman (2003):

\[
\log m_{xti} = \alpha_{xi} + \beta_{xi}^{1} \kappa_{ti}^{1} + \beta_{xi}^{2} \kappa_{ti}^{2} + \gamma_{ci}
\]

where \( c = t - x \) is the cohort (year of birth).

Specific versions include models with:

- **common age effect**: \( \alpha_{xi} = \alpha_{x} \)
- **non-parametric common age effects**: \( \beta_{xi}^{k} = \beta_{x}^{k} \) (Kleinow, 2015)
- **fixed age effects**: constant \( \beta_{xi}^{1} = 1 \) and linear \( \beta_{xi}^{2} = x - \bar{x} \), where \( \bar{x} \) is the mean age in the data set. (Plat, 2009)
All considered models are variants of group specific Lee-Carter type models with the extension to a second age-period effect by Renshaw & Haberman (2003):

\[
\log m_{xti} = \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 + \gamma_{ci}
\]

where \( c = t - x \) is the cohort (year of birth).

Specific versions include models with:

- **common age effect**: \( \alpha_{xi} = \alpha_x \)

- **non-parametric common age effects**: \( \beta_{xi}^k = \beta_x^k \) (Kleinow, 2015)

- **fixed age effects**:
  - constant \( \beta_{xi}^1 = 1 \)
  - linear \( \beta_{xi}^2 = x - \bar{x} \), where \( \bar{x} \) is the mean age in the data set. (Plat, 2009)

- **common period effects**: \( \kappa_{ti}^k = \kappa_t^k \) (Li and Lee, 2005)
Models

All considered models are variants of group specific Lee-Carter type models with the extension to a second age-period effect by Renshaw & Haberman (2003):

\[ \log m_{xti} = \alpha_{xi} + \beta_{xi}^{1} \kappa_{ti}^{1} + \beta_{xi}^{2} \kappa_{ti}^{2} + \gamma_{ci} \]

where \( c = t - x \) is the cohort (year of birth).

Specific versions include models with:

- **common age effect**: \( \alpha_{xi} = \alpha_{x} \)
- **non-parametric common age effects**: \( \beta_{xi}^{k} = \beta_{x}^{k} \) (Kleinow, 2015)
- **fixed age effects**: constant \( \beta_{xi}^{1} = 1 \) and linear \( \beta_{xi}^{2} = x - \bar{x} \), where \( \bar{x} \) is the mean age in the data set. (Plat, 2009)
- **common period effects**: \( \kappa_{ti}^{k} = \kappa_{t}^{k} \) (Li and Lee, 2005)
- **group specific trends in common period effects**: \( \kappa_{ti}^{k} = \kappa_{t}^{k} + \eta_{i}(t - \bar{t}) \)

and variations with and without cohort effects.
Models

\[ m1: \quad \log m_{xti} = \alpha_{xi} + \kappa_{1t_i}^1 + (x - \bar{x})\kappa_{1t_i}^2 \]  
(Plat, 2009)

\[ m2: \quad \log m_{xti} = \alpha_{xi} + \beta_{1x}^1 \kappa_{1t_i}^1 + \beta_{2x}^2 \kappa_{1t_i}^2 \]  
(Kleinow, 2015)

\[ \vdots \]

\[ m6: \quad \log m_{xti} = \alpha_{x} + \kappa_{1t_i}^1 + (x - \bar{x})\kappa_{1t_i}^2 \]  
m1 + common \( \alpha \)

\[ \vdots \]

\[ m9: \quad \log m_{xti} = \alpha_{x} + \eta_i(x - \bar{x}) + \kappa_{1t}^1 \]  
\[ + d_i^0 + d_i^1(t - \bar{t}) \]  
\[ +(x - \bar{x})(\kappa_{2t}^1 + d_i^2(t - \bar{t})) \]

\[ \vdots \]

\[ m12: \quad \log m_{xti} = \alpha_{xi} + \beta_{1xi}^1 \kappa_{1t_i}^1 + \beta_{2xi}^2 \kappa_{1t_i}^2 \]  
(Renshaw & Haberman, 2003)

\[ \vdots \]

\[ m14: \quad \log m_{xti} = \alpha_{x} + \beta_{1x}^1 \kappa_{1t}^1 + \beta_{2x}^2 \kappa_{1t}^2 \]  
m2 + common \( \alpha \)

\[ m15: \quad \log m_{xti} = \alpha_{x} + \beta_{1x}^1 \kappa_{t}^1 + \beta_{2x}^2 \kappa_{t}^2 \]  
(Li & Lee, 2005)

+ variants with common or group-specific cohort effect, \( \gamma_c \) or \( \gamma_{ci} \).

Jie Wen, AJG. Cairns, T. Kleinow: Mortality and Deprivation
Maximum Likelihood estimation based on\n\[ D_{x,t,i} \sim \text{Poisson}\left(\mu_{x,t,i} E_{x,t,i}^c\right) \] is applied to obtain estimated parameter values.

All suggested models have some identifiability issues, that is, different parameter values lead to the same fitted mortality rates \( m_{xti} \), and, therefore to the same value of the likelihood function.

To obtain unique parameter values we apply model-specific constraints.
Questions for this talk

\[ \log m_{x,t} = \alpha_x + \beta_{x,1} \kappa_{t,1} + \beta_{x,2} \kappa_{t,2} + \gamma_c \]

- What parameters should be chosen to be group specific and which parameters are common?
- Should age-effects be estimated?
- Should we include cohort effects (common or group specific)?
- What parameters show the greatest differences between IMD groups?
- Are the groups clustered?
Parameter estimates - m12 - the most general model

\[ \log m_{xti} = \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 \] (Renshaw & Haberman, 2003)

MLE estimated alpha(x) – m12

![Graph showing MLE estimated alpha(x) for m12 model]
Parameter estimates - m12 - the most general model

\[
\log m_{xti} = \alpha_{xi} + \beta_{xi}^{1} \kappa_{ti}^{1} + \beta_{xi}^{2} \kappa_{ti}^{2} \quad (\text{Renshaw}&\text{Haberman, 2003})
\]
Parameter estimates - m12 - the most general model

\[ \log m_{x,t} = \alpha_x + \beta_{x,t}^1 \kappa_{t}^1 + \beta_{x,t}^2 \kappa_{t}^2 \] (Renshaw & Haberman, 2003)

MLE estimated beta1(x) – m12

Jie Wen, A.J.G. Cairns, T. Kleinow: Mortality and Deprivation
Parameter estimates - m12 - the most general model

\[ \log m_{xti} = \alpha_{xi} + \beta_{xi}^{1}\kappa_{ti}^{1} + \beta_{xi}^{2}\kappa_{ti}^{2} \] (Renshaw & Haberman, 2003)

**MLE estimated kappa1(t) – m12**

![Graph showing MLE estimated kappa1(t) over years from 2002 to 2014.](image)
\[ \log m_{x_{ti}} = \alpha_{x_i} + \beta_{x_i}^1 \kappa_{t_i}^1 + \beta_{x_i}^2 \kappa_{t_i}^2 \] (Renshaw & Haberman, 2003)
Parameter estimates - m12 - the most general model

\[ \log m_{x ti} = \alpha_{xi} + \beta_{xi}^{1} \kappa_{ti}^{1} + \beta_{xi}^{2} \kappa_{ti}^{2} \] (Renshaw & Haberman, 2003)
Parameter estimates - m12 - the most general model

Bayesian Information Criterion: \( k \log n - 2 \log(L) \)

<table>
<thead>
<tr>
<th>Log-likelihood</th>
<th>parameters</th>
<th>constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
</tr>
</tbody>
</table>
Parameter estimates - m12 - the most general model

Bayesian Information Criterion: \( k \log n – 2 \log(L) \)

<table>
<thead>
<tr>
<th>Log-likelihood</th>
<th>parameters</th>
<th>constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
</tr>
</tbody>
</table>

- good fit
- very large number of parameters
Parameter estimates - m2 - common beta

\[ \log m_{x+t} = \alpha_x + \beta^1_x \kappa^1_{t+i} + \beta^2_x \kappa^2_{t+i} \quad \text{(Kleinow, 2015)} \]

MLE estimated alpha(x) – m2
Parameter estimates - m2 - common beta

\[ \log m_{xti} = \alpha_x + \beta_1 x \kappa_1 + \beta_2 x \kappa_2 \] (Kleinow, 2015)

MLE estimated beta1(x) – m2 vs m12

![Graph showing MLE estimated beta1(x) vs age]

Jie Wen, A.J.G. Cairns, T. Kleinow: Mortality and Deprivation
$\log m_{x,t} = \alpha_x + \beta_1^1 \kappa_{t,1} + \beta_2^2 \kappa_{t,2}$ (Kleinow, 2015)
\[ \log m_{xti} = \alpha_{xi} + \beta_{x1}^{1} \kappa_{ti}^{1} + \beta_{x2}^{2} \kappa_{ti}^{2} \] (Kleinow, 2015)
\[ \log m_{x ti} = \alpha_{xi} + \beta_1^{x_i} \kappa_1^{ti} + \beta_2^{x_i} \kappa_2^{ti} \] (Kleinow, 2015)

Fitted mortality in 2015 – m2

male mortality in year 2015
Bayesian Information Criterion:  \[ k \log n - 2 \log(L) \]

<table>
<thead>
<tr>
<th>Log-likelihood</th>
<th>parameters</th>
<th>constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
</tr>
</tbody>
</table>
### Parameter estimates - m2 - common beta

Bayesian Information Criterion: \( k \log n - 2 \log(L) \)

<table>
<thead>
<tr>
<th>Log-likelihood</th>
<th>parameters</th>
<th>constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
</tr>
</tbody>
</table>

- reasonable fit
- smaller number of parameters than m12
- BIC improved
Parameter estimates - m1

\[ \log m_{x+ti} = \alpha_{xi} + \kappa_{1}^{ti} + (x - \bar{x})\kappa_{2}^{ti} \] (Plat, 2009)

**male mortality in year 2015**

**MLE estimated alpha(x) – m1**
Parameter estimates - m1

\[ \log m_{x \tau i} = \alpha_{\xi i} + \kappa_{\tau i}^{1} + (x - \bar{x})\kappa_{\tau i}^{2} \quad (\text{Plat, 2009}) \]

MLE estimated \( \kappa_{1}(t) \) - m1

![MLE estimated \( \kappa_{1}(t) \) - m1](image)
\[ \log m_{x\tau i} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \] (Plat, 2009)
\[
\log m_{xti} = \alpha x_i + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \quad \text{(Plat, 2009)}
\]
Parameter estimates - m1

$$\log m_{xi} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$$ (Plat, 2009)

Fitted mortality in 2015 – m1

male mortality in year 2015
## Parameter estimates - m1

Bayesian Information Criterion: \( k \log n - 2 \log(L) \)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Log-likelihood</th>
<th>Parameters</th>
<th>Constraints</th>
<th>D.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-31,403.22</td>
<td>800</td>
<td>20</td>
<td>780</td>
<td>69,766.10</td>
</tr>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
<td>69,016.49</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
<td>75,966.82</td>
</tr>
</tbody>
</table>
Parameter estimates - m1

Bayesian Information Criterion: $k \log n - 2 \log(L)$

<table>
<thead>
<tr>
<th></th>
<th>Log-likelihood</th>
<th>parameters</th>
<th>constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-31,403.22</td>
<td>800</td>
<td>20</td>
<td>780</td>
<td>69,766.10</td>
</tr>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
<td>69,016.49</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
<td>75,966.82</td>
</tr>
</tbody>
</table>

- model does not fit as well as m2 and m12 (lower likelihood)
- number of parameters reduced further
- but BIC does not improve
Next steps

\begin{align*}
\text{m1: } \log m_{xti} &= \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \quad \text{(Plat, 2009)} \\
\text{m2: } \log m_{xti} &= \alpha_{xi} + \beta_{x1}^1 \kappa_{ti}^1 + \beta_{x2}^2 \kappa_{ti}^2 \quad \text{(Kleinow, 2015)} \\
\text{m12: } \log m_{xti} &= \alpha_{xi} + \beta_{x1}^1 \kappa_{ti}^1 + \beta_{x2}^2 \kappa_{ti}^2 \quad \text{(Renshaw & Haberman, 2003)}
\end{align*}
Next steps

\[ \begin{align*}
\text{m1: } \log m_{x|ti} &= \alpha_{x_i} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \\
\text{m2: } \log m_{x|ti} &= \alpha_{x_i} + \beta_{x|ti}^1 + \beta_{x|ti}^2 \\
\text{m12: } \log m_{x|ti} &= \alpha_{x_i} + \beta_{x|ti}^1 + \beta_{x|ti}^2 
\end{align*} \]  
(Plat, 2009)  
(Kleinow, 2015)  
(Renshaw & Haberman, 2003)

Try a common alpha for m1 and m2.
Parameter estimates - m14 - common alpha and beta

\[ \log m_{xi} = \alpha_x + \beta_x \kappa_{ti}^1 + \beta_x \kappa_{ti}^2 \]

MLE estimated alpha(x) – m14

![Graph showing MLE estimated alpha(x) for ages 40 to 90, with a linear trend line]

Jie Wen, AJG. Cairns, T. Kleinow: Mortality and Deprivation
Parameter estimates - m14 - common alpha and beta

\[ \log m_{x\tau_i} = \alpha_x + \beta_{1x} \kappa_{1\tau_i} + \beta_{2x} \kappa_{2\tau_i} \]

MLE estimated beta1(x) – m14

![Graph showing the MLE estimated beta1(x) for m14 age distribution.](image)
### Parameter estimates - m14 - common alpha and beta

\[
\log m_{x\tau_i} = \alpha_x + \beta_x^1 \kappa_{\tau_i}^1 + \beta_x^2 \kappa_{\tau_i}^2
\]

**MLE estimated kappa1(t) – m14**

![Graph of MLE estimated kappa1(t) - m14](image)
Parameter estimates - m14 - common alpha and beta

\[ \log m_{xti} = \alpha_x + \beta^1_x \kappa^1_{ti} + \beta^2_x \kappa^2_{ti} \]
Parameter estimates - m14 - common alpha and beta

\[ \log m_{xti} = \alpha_x + \beta_1^{x} \kappa_{ti}^1 + \beta_2^{x} \kappa_{ti}^2 \]

Fitted mortality in 2015 – m14

male mortality in year 2015
Parameter estimates - m14 - common alpha and beta

Bayesian Information Criterion: \( k \log n - 2 \log(L) \)

<table>
<thead>
<tr>
<th></th>
<th>Log-likelihood</th>
<th>parameters</th>
<th>constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-31,403.22</td>
<td>800</td>
<td>20</td>
<td>780</td>
<td>69,766.10</td>
</tr>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
<td>69,016.49</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
<td>75,966.82</td>
</tr>
<tr>
<td>m14</td>
<td>-30,852.96</td>
<td>450</td>
<td>4</td>
<td>446</td>
<td>65,685.42</td>
</tr>
</tbody>
</table>

Higher likelihood than m1 with fewer parameters
Best BIC
Bayesian Information Criterion: \( k \log n - 2 \log(L) \)

<table>
<thead>
<tr>
<th></th>
<th>Log-likelihood</th>
<th>parameters</th>
<th>constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-31,403.22</td>
<td>800</td>
<td>20</td>
<td>780</td>
<td>69,766.10</td>
</tr>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
<td>69,016.49</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
<td>75,966.82</td>
</tr>
<tr>
<td>m14</td>
<td>-30,852.96</td>
<td>450</td>
<td>4</td>
<td>446</td>
<td>65,685.42</td>
</tr>
</tbody>
</table>

- higher likelihood than m1 with fewer parameters
- best BIC
Parameter estimates - m6 - M1 with common alpha

\[ \log m_{xti} = \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \]

MLE estimated alpha(x) – m6
Parameter estimates - m6 - M1 with common alpha

\[ \log m_{x\tau_i} = \alpha_x + \kappa_{1\tau_i} + (x - \bar{x})\kappa_{2\tau_i} \]

MLE estimated alpha(x) – m6

MLE estimated alpha(x) – m1
Parameter estimates - m6 - M1 with common alpha

\[ \log m_{x_{ti}} = \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \]

MLE estimated \( \kappa_1(t) - m6 \)

Jie Wen, AJG. Cairns, T. Kleinow: Mortality and Deprivation
Parameter estimates - m6 - M1 with common alpha

\[ \log m_{xti} = \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \]
Parameter estimates - m6 - M1 with common alpha

\[ \log m_{xti} = \alpha_x + \kappa_{1ti} + (x - \bar{x})\kappa_{2ti} \]

MLE estimated kappa2(t) – m6

<table>
<thead>
<tr>
<th>year</th>
<th>kappa2(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>0.000</td>
</tr>
<tr>
<td>2006</td>
<td>0.000</td>
</tr>
<tr>
<td>2010</td>
<td>0.000</td>
</tr>
<tr>
<td>2014</td>
<td>0.000</td>
</tr>
</tbody>
</table>
\[
\log m_{x,t_i} = \alpha_x + \kappa^1_{t_i} + (x - \bar{x})\kappa^2_{t_i}
\]
\[
\log m_{x_t} = \alpha_x + \kappa_{t_i}^1 + (x - \bar{x})\kappa_{t_i}^2
\]
Parameter estimates - m6 - M1 with common alpha

\[
\log m_{x_{ti}} = \alpha_x + \kappa_{x_{ti}}^1 + (x - \bar{x})\kappa_{x_{ti}}^2
\]
## Parameter estimates - m6 - M1 with common alpha

Bayesian Information Criterion: \( k \log n – 2 \log(L) \)

<table>
<thead>
<tr>
<th>Model</th>
<th>Log-likelihood</th>
<th>Parameters</th>
<th>Constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-31,403.22</td>
<td>800</td>
<td>20</td>
<td>780</td>
<td>69,766.10</td>
</tr>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
<td>69,016.49</td>
</tr>
<tr>
<td>m6</td>
<td>-32,052.10</td>
<td>350</td>
<td>2</td>
<td>348</td>
<td>67,209.29</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
<td>75,966.82</td>
</tr>
<tr>
<td>m14</td>
<td>-30,852.96</td>
<td>450</td>
<td>4</td>
<td>446</td>
<td>65,685.42</td>
</tr>
</tbody>
</table>

The smallest number of parameters corresponds to the second best model in terms of BIC, which is m6. m6 has a better BIC than m1.
Parameter estimates - m6 - M1 with common alpha

Bayesian Information Criterion: $k \log n – 2 \log(L)$

<table>
<thead>
<tr>
<th>Log-likelihood</th>
<th>parameters</th>
<th>constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-31,403.22</td>
<td>800</td>
<td>20</td>
<td>780</td>
</tr>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
</tr>
<tr>
<td>m6</td>
<td>-32,052.10</td>
<td>350</td>
<td>2</td>
<td>348</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
</tr>
<tr>
<td>m14</td>
<td>-30,852.96</td>
<td>450</td>
<td>4</td>
<td>446</td>
</tr>
</tbody>
</table>

- smallest number of parameters
- second best model in terms of BIC
- better BIC than m1
### Ranking the Models - Goodness of Fit - Bayesian Information Criterion: $k \log n - 2 \log(L)$

<table>
<thead>
<tr>
<th>Model</th>
<th>Log-likelihood</th>
<th>Parameters</th>
<th>Constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-31,403.22</td>
<td>800</td>
<td>20</td>
<td>780</td>
<td>69,766.10</td>
</tr>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
<td>69,016.49</td>
</tr>
<tr>
<td>m3</td>
<td>-31,962.82</td>
<td>530</td>
<td>2</td>
<td>528</td>
<td>68,636.80</td>
</tr>
<tr>
<td>m4</td>
<td>-31,700.90</td>
<td>665</td>
<td>11</td>
<td>654</td>
<td>69,237.21</td>
</tr>
<tr>
<td>m5</td>
<td>-31,474.15</td>
<td>675</td>
<td>12</td>
<td>663</td>
<td>68,864.03</td>
</tr>
<tr>
<td>m6</td>
<td>-32,052.10</td>
<td>350</td>
<td>2</td>
<td>348</td>
<td>67,209.29</td>
</tr>
<tr>
<td>m7</td>
<td>-31,467.01</td>
<td>665</td>
<td>11</td>
<td>654</td>
<td>68,769.44</td>
</tr>
<tr>
<td>m8</td>
<td>-31,460.51</td>
<td>675</td>
<td>12</td>
<td>663</td>
<td>68,836.75</td>
</tr>
<tr>
<td>m9</td>
<td>-32,184.56</td>
<td>120</td>
<td>6</td>
<td>114</td>
<td>65,386.30</td>
</tr>
<tr>
<td>m10</td>
<td>-31,536.52</td>
<td>550</td>
<td>4</td>
<td>546</td>
<td>67,944.81</td>
</tr>
<tr>
<td>m11</td>
<td>-30,323.70</td>
<td>1350</td>
<td>22</td>
<td>1328</td>
<td>72,496.68</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
<td>75,966.82</td>
</tr>
<tr>
<td>m13</td>
<td>-30,558.20</td>
<td>1150</td>
<td>20</td>
<td>1130</td>
<td>71,199.00</td>
</tr>
<tr>
<td>m14</td>
<td>-30,852.96</td>
<td>450</td>
<td>4</td>
<td>446</td>
<td>65,685.42</td>
</tr>
<tr>
<td>m15</td>
<td>-30,462.15</td>
<td>1215</td>
<td>22</td>
<td>1193</td>
<td>71,569.04</td>
</tr>
</tbody>
</table>
## Ranking the Models - Goodness of Fit - Bayesian Information Criterion: \( k \log n - 2 \log(L) \)

<table>
<thead>
<tr>
<th>Model</th>
<th>Log-likelihood</th>
<th>Parameters</th>
<th>Constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-31,403.22</td>
<td>800</td>
<td>20</td>
<td>780</td>
<td>69,766.10</td>
</tr>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
<td>69,016.49</td>
</tr>
<tr>
<td>m3</td>
<td>-31,962.82</td>
<td>530</td>
<td>2</td>
<td>528</td>
<td>68,636.80</td>
</tr>
<tr>
<td>m4</td>
<td>-31,700.90</td>
<td>665</td>
<td>11</td>
<td>654</td>
<td>69,237.21</td>
</tr>
<tr>
<td>m5</td>
<td>-31,474.15</td>
<td>675</td>
<td>12</td>
<td>663</td>
<td>68,864.03</td>
</tr>
<tr>
<td>m6</td>
<td>-32,052.10</td>
<td>350</td>
<td>2</td>
<td>348</td>
<td>67,209.29</td>
</tr>
<tr>
<td>m7</td>
<td>-31,467.01</td>
<td>665</td>
<td>11</td>
<td>654</td>
<td>68,769.44</td>
</tr>
<tr>
<td>m8</td>
<td>-31,460.51</td>
<td>675</td>
<td>12</td>
<td>663</td>
<td>68,836.75</td>
</tr>
<tr>
<td>m9</td>
<td>-32,184.56</td>
<td>120</td>
<td>6</td>
<td>114</td>
<td>65,386.30</td>
</tr>
<tr>
<td>m10</td>
<td>-31,536.52</td>
<td>550</td>
<td>4</td>
<td>546</td>
<td>67,944.81</td>
</tr>
<tr>
<td>m11</td>
<td>-30,323.70</td>
<td>1350</td>
<td>22</td>
<td>1328</td>
<td>72,496.68</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
<td>75,966.82</td>
</tr>
<tr>
<td>m13</td>
<td>-30,558.20</td>
<td>1150</td>
<td>20</td>
<td>1130</td>
<td>71,199.00</td>
</tr>
<tr>
<td><strong>m14</strong></td>
<td><strong>-30,852.96</strong></td>
<td><strong>450</strong></td>
<td><strong>4</strong></td>
<td><strong>446</strong></td>
<td><strong>65,685.42</strong></td>
</tr>
<tr>
<td>m15</td>
<td>-30,462.15</td>
<td>1215</td>
<td>22</td>
<td>1193</td>
<td>71,569.04</td>
</tr>
</tbody>
</table>
### Ranking the Models - Goodness of Fit - Bayesian Information Criterion: $k \log n − 2 \log(L)$

<table>
<thead>
<tr>
<th></th>
<th>Log-likelihood</th>
<th>parameters</th>
<th>constraints</th>
<th>d.o.f.</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-31,403.22</td>
<td>800</td>
<td>20</td>
<td>780</td>
<td>69,766.10</td>
</tr>
<tr>
<td>m2</td>
<td>-30,591.20</td>
<td>900</td>
<td>22</td>
<td>878</td>
<td>69,016.49</td>
</tr>
<tr>
<td>m3</td>
<td>-31,962.82</td>
<td>530</td>
<td>2</td>
<td>528</td>
<td>68,636.80</td>
</tr>
<tr>
<td>m4</td>
<td>-31,700.90</td>
<td>665</td>
<td>11</td>
<td>654</td>
<td>69,237.21</td>
</tr>
<tr>
<td>m5</td>
<td>-31,474.15</td>
<td>675</td>
<td>12</td>
<td>663</td>
<td>68,864.03</td>
</tr>
<tr>
<td>m6</td>
<td>-32,052.10</td>
<td>350</td>
<td>2</td>
<td>348</td>
<td>67,209.29</td>
</tr>
<tr>
<td>m7</td>
<td>-31,467.01</td>
<td>665</td>
<td>11</td>
<td>654</td>
<td>68,769.44</td>
</tr>
<tr>
<td>m8</td>
<td>-31,460.51</td>
<td>675</td>
<td>12</td>
<td>663</td>
<td>68,836.75</td>
</tr>
<tr>
<td><strong>m9</strong></td>
<td><strong>-32,184.56</strong></td>
<td><strong>120</strong></td>
<td><strong>6</strong></td>
<td><strong>114</strong></td>
<td><strong>65,386.30</strong></td>
</tr>
<tr>
<td>m10</td>
<td>-31,536.52</td>
<td>550</td>
<td>4</td>
<td>546</td>
<td>67,944.81</td>
</tr>
<tr>
<td>m11</td>
<td>-30,323.70</td>
<td>1350</td>
<td>22</td>
<td>1328</td>
<td>72,496.68</td>
</tr>
<tr>
<td>m12</td>
<td>-30,131.47</td>
<td>1800</td>
<td>40</td>
<td>1760</td>
<td>75,966.82</td>
</tr>
<tr>
<td>m13</td>
<td>-30,558.20</td>
<td>1150</td>
<td>20</td>
<td>1130</td>
<td>71,199.00</td>
</tr>
<tr>
<td><strong>m14</strong></td>
<td><strong>-30,852.96</strong></td>
<td><strong>450</strong></td>
<td><strong>4</strong></td>
<td><strong>446</strong></td>
<td><strong>65,685.42</strong></td>
</tr>
<tr>
<td>m15</td>
<td>-30,462.15</td>
<td>1215</td>
<td>22</td>
<td>1193</td>
<td>71,569.04</td>
</tr>
</tbody>
</table>
\[
\log m_{x+i} = \alpha_x + \eta_i(x - \bar{x}) + \kappa_1^t + d_0^i + d_1^i(t - \bar{t}) + (x - \bar{x})(\kappa_2^t + d_2^i(t - \bar{t}))
\]
\[
\log m_{xti} = \alpha_x + \eta_i(x - \bar{x}) + \kappa_1 t_i + d_0^i + d_1^i (t - \bar{t}) + (x - \bar{x})(\kappa_2 t_i + d_2^i (t - \bar{t}))
\]
$$\log{m_{xti}} = \alpha_x + \eta_i(x - \bar{x}) + \kappa_1^1 + d_i^0 + d_i^1(t - \bar{t}) + (x - \bar{x})(\kappa_2^1 + d_i^2(t - \bar{t}))$$

- Lowest BIC, few parameters
- introduces constant group specific improvement rates
- for ever increasing mortality differentials
Models

\[ m1: \quad \log m_{xti} = \alpha_x + \kappa^1_{ti} + (x - \bar{x})\kappa^2_{ti} \quad \text{(Plat, 2009)} \]
\[ m2: \quad \log m_{xti} = \alpha_x + \beta^1_x \kappa^1_{ti} + \beta^2_x \kappa^2_{ti} \quad \text{(Kleinow, 2015)} \]

\[ \vdots \]

\[ m6: \quad \log m_{xti} = \alpha_x + \kappa^1_{ti} + (x - \bar{x})\kappa^2_{ti} \quad \text{m1 + common } \alpha \]

\[ \vdots \]

\[ m9: \quad \log m_{xti} = \alpha_x + \eta_i(x - \bar{x}) + \kappa^1_t \\
+ d^0_i + d^1_i(t - \bar{t}) \\
+ (x - \bar{x})(\kappa^2_t + d^2_i(t - \bar{t})) \]

\[ \vdots \]

\[ m12: \quad \log m_{xti} = \alpha_x + \beta^1_x \kappa^1_{ti} + \beta^2_x \kappa^2_{ti} \quad \text{(Renshaw&Haberman, 2003)} \]

\[ \vdots \]

\[ m14: \quad \log m_{xti} = \alpha_x + \beta^1_x \kappa^1_{ti} + \beta^2_x \kappa^2_{ti} \quad \text{m2 + common } \alpha \]
\[ m15: \quad \log m_{xti} = \alpha_x + \beta^1_x \kappa^1_{ti} + \beta^2_x \kappa^2_{ti} \quad \text{(Li&Lee, 2005)} \]
Conclusions

- There are clear differences between the mortality rates in the ten IMD deciles.
- The improvement rates (from 2001 – 2015) are also different.
Conclusions

- There are clear differences between the mortality rates in the ten IMD deciles.
- The improvement rates (from 2001 – 2015) are also different.
- Models with common age effects seem to perform better than models with group specific age effects.
There are clear differences between the mortality rates in the ten IMD deciles.

The improvement rates (from 2001 – 2015) are also different.

Models with common age effects seem to perform better than models with group specific age effects.

For a wider age range, models with common non-parametric age effects (Kleinow (2015) + common $\alpha$) produce a good fit in terms of BIC, heatmaps ...

However, for a narrower age range (65-89), models with constant/linear $\beta$’s, (Plat (2009) + common $\alpha$) are better.
Conclusions

- There are clear differences between the mortality rates in the ten IMD deciles.
- The improvement rates (from 2001 – 2015) are also different.
- Models with common age effects seem to perform better than models with group specific age effects.
- For a wider age range, models with common non-parametric age effects (Kleinow (2015) + common $\alpha$) produce a good fit in terms of BIC, heatmaps ...
- However, for a narrower age range (65-89), models with constant/linear $\beta$’s, (Plat (2009) + common $\alpha$) are better.
- Cohort effect do not improve the fit for those models
- If a cohort effect is included it should be a common cohort effect