Producing Small Area Estimates of the Need for Hip and Knee Replacement Surgery

ANDY JUDGE

Nicky Welton
Mary Shaw
Yoav Ben-Shlomo
Background

- Fairness in access to healthcare a founding NHS principle
- Health needs vary across different socio-demographic groups and geographical areas
- PCTs now responsible for planning, commissioning and delivery of NHS services
- PCTs should conduct health equity audits (HEA) to ensure health services are distributed in relation to health needs
Why Is Hip & Knee Replacement Important

- Musculoskeletal Services Framework
  - Describes needs assessment in the context of understanding the prevalence and incidence of musculoskeletal disorders, where patients are and their use of services

- National Audit Office 2003
  - Commission for Health Improvement should include, in their clinical governance reviews, the equity with which patients are offered hip replacements

- Department of Health
  - Commissioned the Modernisation Agency to focused on improving equity of access on grounds of age and other criteria in orthopaedics

- National Strategic Framework for Older People
  - Fair access to joint replacement surgery is singled out
  - “even very complex treatment, used appropriately, can benefit older people and should not be denied them solely on the basis of age”
Previous Approaches to Generate Small Area Predictions

- PBS Diabetes Population Prevalence model
  - Developed by the Yorkshire and Humber Public Health Observatory (YHPHO) to predict the prevalence of diabetes in small areas of England.
- Multilevel Modelling approach using MLwiN (Twigg, Moon, Jones et al)
  - Estimate prevalence of smoking and problem drinking behaviour across small areas of England using data from the Health Survey for England
- Bayesian approach using WinBUGS (Peter Congdon)
  - Bayesian modelling approach used by Congdon to estimate prevalence of Diabetes, CHD and Psychiatric conditions in small areas across England
Desirable Features of a Model to Generate Small Area Predictions

- Estimate prevalence using a nationally representative population based sample
- The dataset used to estimate prevalence should be regularly updated
- Use a Multilevel modelling approach
  - Control for overdispersion and clustering
  - Incorporate estimate of residual variation into the predictions process
- Bayesian modelling allows incorporation of prior evidence in regression model
- Estimates of precision are required around small area predictions of prevalence
Datasets Used to Generate Small Area Predictions

- **Census dataset**
  - Population counts by 5-year age groups and sex for each of 7969 CAS wards in England
  - For each CAS ward:
    - IMD2004 deprivation quintiles (Least, 2, 3, 4, Most)
    - Rurality (Urban $\geq 10k$, Town&Fringe, Village/Isolated)
    - Ethnic mix of area (White, Non-white)
  - 7969 CAS wards are nested within 354 districts in England

- **ELSA dataset**
  - Binary outcome of whether patient is in need of hip/knee replacement
  - Exposure variables are: Age, Sex, IMD2004 deprivation quintiles, Rurality, Ethnic mix of area
  - Hierarchical structure consists of 11,392 individuals, within 2,913 wards (anonymised), within 348 districts
  - Multilevel Poisson regression model fitted to data – separate models fitted for hips and knees
Method to Generate Small Area Predictions

Step 1 – Manipulate ELSA Dataset

\[ \text{needTHRad}_{ij} \sim \text{Poisson}(\pi_i) \]

\[ \log(\pi_i) = -3.804(0.195)\text{CONS} + 0.156(0.119)\text{agenew\_60}_i + 0.466(0.118)\text{agenew\_70}_i + 0.704(0.154)\text{agenew\_80}_i + 0.924(0.177)\text{agenew\_85}_i + \]
\[ -0.338(0.091)\text{"Male"}_i + 0.336(0.164)\text{W1\_ind2004\_cat\_eng\_2}_i + 0.621(0.158)\text{W1\_ind2004\_cat\_eng\_3}_i + 0.818(0.158)\text{W1\_ind2004\_cat\_eng\_4}_i + \]
\[ 1.088(0.158)\text{W1\_ind2004\_cat\_eng\_5}_i + -0.089(0.177)\text{"village/isolated"}_i + 0.123(0.148)\text{"town \& fringe"}_i + -0.013(0.122)\text{"non-white"}_i \]

- Estimate predicted log rate in each possible age, sex, deprivation, rurality and ethnic group (there are 300 such groups)
- Forms a dataset of predictions \( P \)

\[
P_j = \begin{bmatrix}
P_1 \\
P_2 \\
\vdots \\
P_{300}
\end{bmatrix}
\]
Method to Generate Small Area Predictions

Step 1 – Manipulate ELSA Dataset

\[ \text{needTHRadj}_i \sim \text{Poisson}(\pi_i) \]
\[ \log(\pi_i) = -3.804(0.195) \text{CONS} + 0.156(0.119) \text{agenew}_60_i + 0.466(0.118) \text{agenew}_70_i + 0.704(0.154) \text{agenew}_80_i + 0.924(0.177) \text{agenew}_85_i + \\
-0.338(0.091) \text{"Male"}_i + 0.336(0.164) \text{W1_ind2004cat_eng}_2_i + 0.621(0.158) \text{W1_ind2004cat_eng}_3_i + 0.818(0.158) \text{W1_ind2004cat_eng}_4_i + \\
1.088(0.158) \text{W1_ind2004cat_eng}_5_i + -0.089(0.177) \text{"village/isolated"}_i + 0.123(0.148) \text{"town & fringe"}_i + -0.013(0.122) \text{"non-white"}_i \]

- Estimate predicted log rate in each possible age, sex, deprivation, rurality and ethnic group (there are 300 such groups)
- Forms a dataset of predictions \( P \)
- If a Multilevel model was used, would have a dataset of district residuals \( D \)

\[ P_j = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_{300} \end{bmatrix} \]

\[ D_k = \begin{bmatrix} D_1 & D_2 & \cdots & D_{354} \end{bmatrix} \]
Method to Generate Small Area Predictions
Step 1 – Manipulate ELSA Dataset

\[ \text{needTHRadi}_i \sim \text{Poisson}(\pi_i) \]

\[ \log(\pi_i) = -3.804(0.195)\text{CONS} + 0.156(0.119)\text{agenew}_60_i + 0.466(0.118)\text{agenew}_70_i + 0.704(0.154)\text{agenew}_80_i + 0.924(0.177)\text{agenew}_85_i + \\
-0.338(0.091)\text{"Male"}_i + 0.336(0.164)W1\text{_ind2004cat}\_\text{eng}_2_i + 0.621(0.158)W1\text{_ind2004cat}\_\text{eng}_3_i + 0.818(0.158)W1\text{_ind2004cat}\_\text{eng}_4_i + \\
1.088(0.158)W1\text{_ind2004cat}\_\text{eng}_5_i + -0.089(0.177)\text{"village/isolated"}_i + 0.123(0.148)\text{"town \& fringe"}_i + -0.013(0.122)\text{"non-white"}_i \]

- Estimate predicted log rate in each possible age, sex, deprivation, rurality and ethnic group (there are 300 such groups)
- Forms a dataset of predictions \( P \)
- If a Multilevel model was used, would have a dataset of district residuals \( D \)
- For each of the 354 districts, we add the district residual to each of the 300 predicted log rates
- The log-rates are then exponentiated to convert them into rates

\[ P_j = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_{300} \end{bmatrix} \]

\[ D_k = \begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_{354} \end{bmatrix} \]

\[ P_jD_k = \begin{bmatrix} P_1D_1 & P_1D_2 & \ldots & P_1D_{354} \\ P_2D_1 & P_2D_2 & \ldots & P_2D_{354} \\ \vdots & \vdots & \ddots & \vdots \\ P_{300}D_1 & P_{300}D_2 & \ldots & P_{300}D_{354} \end{bmatrix} \]
Method to Generate Small Area Predictions
Step 2 – Manipulate Census Dataset

- The census dataset is then reshaped so that we have the estimated population count in each of the 300 age, sex, deprivation, rurality and ethnic groups, for each ward forming a (7969*300) data matrix \( C \)

\[
C_{\text{ward, pred}} = \begin{bmatrix}
C_{1,1} & C_{1,2} & \ldots & C_{1,300} \\
C_{2,1} & C_{2,2} & \ldots & \ldots \\
\vdots & \vdots & \ddots & \vdots \\
C_{79691} & C_{79692} & \ldots & C_{7969300}
\end{bmatrix}
\]
Method to Generate Small Area Predictions
Step 2 – Manipulate Census Dataset

The census dataset is then re-shaped so that we have the estimated population count in each of the 300 age, sex, deprivation, rurality and ethnic groups, for each ward forming a (7969*300) data matrix \( C \)

\[
C_{\text{ward,pred}} = \begin{bmatrix}
C_{1,1} & C_{1,2} & \ldots & C_{1,300} \\
C_{2,1} & C_{2,2} & \ldots & \cdot \\
\cdot & \cdot & \ldots & \cdot \\
C_{79691} & C_{79692} & \ldots & C_{7969300}
\end{bmatrix}
\]

For each ward, we also know which district it belongs to

\[
C_{\text{ward,pred}}D_k = \begin{bmatrix}
C_{1,1}D_1 & C_{1,2}D_1 & \ldots & C_{1,300}D_1 \\
C_{2,1}D_1 & C_{2,2}D_1 & \ldots & \cdot \\
\cdot & \cdot & \ldots & \cdot \\
C_{79691}D_{354} & C_{79692}D_{354} & \ldots & C_{7969300}D_{354}
\end{bmatrix}
\]
Method to Generate Small Area Predictions

Step 3 – Generate Ward Level Predictions

- Apply the predicted rates in each of the 300 groups in the ‘ELSA’ dataset, to the population counts in each of the 300 groups in the ‘census’ dataset

\[
P_{jD_k} \cdot C_{\text{ward, pred}} \cdot D_k = \begin{bmatrix}
P_1D_1 \cdot C_{1,1} \cdot D_1 & P_2D_1 \cdot C_{1,2} \cdot D_1 & \ldots & P_{300}D_1 \cdot C_{1,300} \cdot D_1 \\
P_1D_1 \cdot C_{2,1} \cdot D_1 & P_2D_1 \cdot C_{2,2} \cdot D_1 & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots \\
P_1D_{354} \cdot C_{7969,1} \cdot D_{354} & P_2D_{354} \cdot C_{7969,2} \cdot D_{354} & \ldots & P_{300}D_{354} \cdot C_{7969,300} \cdot D_{354}
\end{bmatrix}
\]
Method to Generate Small Area Predictions

Step 3 – Generate Ward Level Predictions

• Apply the predicted rates in each of the 300 groups in the ‘ELSA’ dataset, to the population counts in each of the 300 groups in the ‘census’ dataset

\[
P_j D_k \times C_{\text{ward, pred}} D_k = \begin{bmatrix}
P_1 D_1 \times C_{1,1} D_1 & P_2 D_1 \times C_{1,2} D_1 & \cdots & P_{300} D_1 \times C_{1,300} D_1 \\
P_1 D_1 \times C_{2,1} D_1 & P_2 D_1 \times C_{2,2} D_1 & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots \\
P_1 D_{354} \times C_{7969,1} D_{354} & P_2 D_{354} \times C_{7969,2} D_{354} & \cdots & P_{300} D_{354} \times C_{7969,300} D_{354}
\end{bmatrix}
\]

• Sum across the rows to get the total expected numbers of people in need in each census ward

\[
E_{\text{ward}} D_k = \begin{bmatrix}
E_1 D_1 \\
E_2 D_1 \\
\cdots \\
E_{7969} D_{354}
\end{bmatrix}
\]
Method to Generate Small Area Predictions

Step 3 – Generate Ward Level Predictions

- Apply the predicted rates in each of the 300 groups in the ‘ELSA’ dataset, to the population counts in each of the 300 groups in the ‘census’ dataset

\[
P_j D_k \times C_{\text{ward, pred}} D_k = \begin{bmatrix}
P_1 D_1 \times C_{1,1} D_1 & P_2 D_1 \times C_{1,2} D_1 & \cdots & P_{300} D_1 \times C_{1,300} D_1 \\
P_1 D_1 \times C_{2,1} D_1 & P_2 D_1 \times C_{2,2} D_1 & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots \\
P_1 D_{354} \times C_{7969,1} D_{354} & P_2 D_{354} \times C_{7969,2} D_{354} & \cdots & P_{300} D_{354} \times C_{7969,300} D_{354}
\end{bmatrix}
\]

- Sum across the rows to get the total expected numbers of people in need in each census ward

\[
E_{\text{ward}} D_k = \begin{bmatrix}
E_1 D_1 \\
E_2 D_1 \\
\vdots \\
E_{7969} D_{354}
\end{bmatrix}
\]

- From the ‘census’ dataset, sum across the rows, to get the total number of people in the population in each census ward

\[
T_{\text{ward}} D_k = \begin{bmatrix}
T_1 D_1 \\
T_2 D_1 \\
\vdots \\
T_{7969} D_{354}
\end{bmatrix}
\]
Method to Generate Small Area Predictions
Step 3 – Generate Ward Level Predictions

- Apply the predicted rates in each of the 300 groups in the ‘ELSA’ dataset, to the population counts in each of the 300 groups in the ‘census’ dataset

\[
P_j D_{k} \ast C_{\text{ward, pred}} D_k = \begin{bmatrix}
P_{D1} \ast C_{1,1} D_1 & P_{D1} \ast C_{1,2} D_1 & \cdots & P_{D1} \ast C_{1,300} D_1 \\
P_{D1} \ast C_{2,1} D_1 & P_{D1} \ast C_{2,2} D_1 & \cdots & \\
& \cdots & \cdots & \\
P_{D354} \ast C_{7969,1} D_{354} & P_{D354} \ast C_{7969,2} D_{354} & \cdots & P_{D354} \ast C_{7969,300} D_{354}
\end{bmatrix}
\]

- Sum across the rows to get the total expected numbers of people in need in each census ward

- From the ‘census’ dataset, sum across the rows, to get the total number of people in the population in each census ward

- Divide expected count by total population count to get predicted rate of need
Generate Estimates of Precision Around Small Area Predictions (1)

- Solution is to use the simulation environment provided in the statistical software package WinBUGS
- Fixed-effects Poisson regression models fitted in WinBUGS to estimate rates of need for hip/knee replacement by socio-demographic groups – flat non-informative priors used
- Convergence assessed using BGr diagnostic tool. Considered adequate after 5000 iterations monitoring 2-chains for both hip and knee models. A further sample of 2000 simulations run on which results based
- The estimates from the 4000 iterations run for each variable (2000x2-chains), form a posterior distribution for that variable
Generate Estimates of Precision Around Small Area Predictions (2)

- Logical expressions are used to extend the model code in WinBUGS to generate the predicted log-rates of need in each of the 300 possible age, sex, deprivation, rurality and ethnic groups (step 1).
- Predicted rates are then applied to population counts from the census (step 2) to get the expected numbers of people in need in each of the 354 districts (step 3).
- Expected counts are summed together to get the total number of people in need in each district.
- Expected counts are divided by the total number of people in the population in each district, to get the predicted rate of need in each district (step 4).
- Computational problems (solution to use R on supercomputer).
## Top 5 Highest and Lowest Rates of Need for Hip and Knee Replacement, Adjusted for Socio-demographic Variables

<table>
<thead>
<tr>
<th>District</th>
<th>Adjusted rate per 1000 (95%CI)</th>
<th>District</th>
<th>Adjusted rate per 1000 (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hip replacement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five Lowest</td>
<td></td>
<td>Five Highest</td>
<td></td>
</tr>
<tr>
<td>Wokingham (00MF)</td>
<td>22.98 (17.43, 29.22)</td>
<td>Manchester (00BN)</td>
<td>74.06 (61.99, 87.28)</td>
</tr>
<tr>
<td>Hart (24UG)</td>
<td>23.06 (17.42, 29.37)</td>
<td>Newham (00BB)</td>
<td>75.32 (62.18, 89.77)</td>
</tr>
<tr>
<td>East Hertfordshire (26UD)</td>
<td>24.55 (19.10, 30.61)</td>
<td>Easington (20UF)</td>
<td>75.43 (60.42, 93.18)</td>
</tr>
<tr>
<td>Uttlesford (22UQ)</td>
<td>24.71 (18.61, 31.73)</td>
<td>Islington (00AU)</td>
<td>75.56 (62.45, 90.04)</td>
</tr>
<tr>
<td>South Northamptonshire (34UG)</td>
<td>24.94 (18.81, 32.07)</td>
<td>Hackney (00AM)</td>
<td>75.60 (62.42, 90.12)</td>
</tr>
<tr>
<td><strong>Knee replacement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five Lowest</td>
<td></td>
<td>Five Highest</td>
<td></td>
</tr>
<tr>
<td>Wokingham (00MF)</td>
<td>31.08 (24.09, 38.64)</td>
<td>Manchester (00BN)</td>
<td>104.60 (90.22, 119.86)</td>
</tr>
<tr>
<td>Hart (24UG)</td>
<td>31.24 (24.31, 38.80)</td>
<td>Islington (00AU)</td>
<td>107.15 (91.78, 124.02)</td>
</tr>
<tr>
<td>East Hertfordshire (26UD)</td>
<td>33.14 (26.43, 40.34)</td>
<td>Newham (00BB)</td>
<td>107.25 (91.71, 124.00)</td>
</tr>
<tr>
<td>Uttlesford (22UQ)</td>
<td>33.53 (26.10, 41.87)</td>
<td>Hackney (00AM)</td>
<td>107.31 (91.82, 124.30)</td>
</tr>
<tr>
<td>South Cambridgeshire (12UG)</td>
<td>33.62 (25.68, 42.80)</td>
<td>Easington (20UF)</td>
<td>116.54 (95.87, 139.97)</td>
</tr>
</tbody>
</table>
Rate of Need for Hip Replacement Across the 354 Districts in England

Rate per 1000
- 22.98 - 32.41
- 32.41 - 39.78
- 39.78 - 46.57
- 46.57 - 55.69
- 55.69 - 75.6
Rate of Need for Knee Replacement Across the 354 Districts in England

Rate per 1000
- 31.08 - 42.99
- 42.99 - 52.92
- 52.92 - 61.96
- 61.96 - 76.39
- 76.39 - 116.54
Discussion

- For the first time small area estimates of need for hip and knee replacement surgery have been produced – data should be of use to local health planners
- Rates of need are adjusted for the socio-demographic profile of each district and include estimates of precision
- Multilevel modelling approach allows the incorporation of residual variation across districts into the predictions process
- Limitation of analysis is that estimates of need were not adjusted for willingness and fitness for surgery
- Future reproducibility of ELSA study
- Further research is required to see whether provision matches need (across socio-demographic groups) in order to determine whether it is equitable