Unlocking a national adult cardiac surgery audit registry with R

The R User Conference 2013
University of Castilla-La Mancha, Albacete, Spain

GL Hickey\textsuperscript{1,2,3}, SW Grant\textsuperscript{2,3} & B Bridgewater\textsuperscript{1,2,3}

\textsuperscript{1}Northwest Institute of BioHealth Informatics, University of Manchester
\textsuperscript{2}University Hospital of South Manchester
\textsuperscript{3}National Institute of Cardiovascular Outcomes Research, UCL
BACKGROUND
Bristol Inquiry

Contributory factors that led to the failings included:

1. Inadequate collection of data
2. Inadequate monitoring of data
National Adult Cardiac Surgery Audit registry

• Up to 166 clinical variables collected on each patient: administrative, demographics, comorbidities, operative factors, outcomes

• 15 years of data

• 465,000 records

• 44 hospitals + >400 consultant surgeons
Flow of data

NICOR → NIBHI → DATABASE CLEANING → ANALYSES

AUDIT & GOVERNANCE TOOLS

CLINICAL RESEARCHERS

RESEARCH

NATIONAL DEATH REGISTER*

* Ability to link with many other national registries
UNLOCKING THE REGISTRY

MESSY DATA
Cleaning the registry in R

- Script per each variable
- Some dependencies

E.g. duplicates

Scripts to add:
- Risk scores
- Combined variables
- ‘Resolve’ conflicting variables

Rapidly reproducible
```r
> with(SCTS, table(X4.04.Discharge.Destination, X4.05.Status.at.Discharge))

<table>
<thead>
<tr>
<th>X4.04.Discharge.Destination</th>
<th>X4.05.Status.at.Discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td>. Another dept within the trust</td>
<td>0. Alive 1. Dead</td>
</tr>
<tr>
<td>0</td>
<td>828</td>
</tr>
<tr>
<td>0. Not applicable - patient deceased</td>
<td>48296</td>
</tr>
<tr>
<td>1 Home</td>
<td>2453</td>
</tr>
<tr>
<td>1. Home</td>
<td>674</td>
</tr>
<tr>
<td>2 Convalescence</td>
<td>370763</td>
</tr>
<tr>
<td>2. Convalescence</td>
<td>374</td>
</tr>
<tr>
<td>2. Convalescence (Non acute Hospital)</td>
<td>7347</td>
</tr>
<tr>
<td>3 Other hospital</td>
<td>8</td>
</tr>
<tr>
<td>3 Other Hospital</td>
<td>2164</td>
</tr>
<tr>
<td>3 Other Hospital - wd 6</td>
<td>304</td>
</tr>
<tr>
<td>3 Other Hospital wd 2</td>
<td>1</td>
</tr>
<tr>
<td>3 Other ward</td>
<td>1</td>
</tr>
<tr>
<td>3 Other Acute hospital</td>
<td>7680</td>
</tr>
<tr>
<td>3 Other hospital</td>
<td>22935</td>
</tr>
<tr>
<td>4 Patient deceased</td>
<td>1</td>
</tr>
<tr>
<td>4. Not applicable - patient deceased</td>
<td>412</td>
</tr>
<tr>
<td>4. Patient Deceased</td>
<td>13286</td>
</tr>
<tr>
<td>5</td>
<td>173</td>
</tr>
<tr>
<td>5. Transferred to different Consultant - NGH</td>
<td>412</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>51</td>
</tr>
<tr>
<td>9</td>
<td>114</td>
</tr>
<tr>
<td>Second op</td>
<td>3820</td>
</tr>
<tr>
<td></td>
<td>518</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>
```

Transcriptional discrepancies:
- 2 Convalescence
- 3 Other Hospital
- 3 Other Hospital - wd 6
- 3 Other Hospital wd 2
- 3 Other ward
- 3 Other Acute hospital
- 4. Not applicable - patient deceased
- 5. Transferred to different Consultant - NGH

Illegal options:
- Second op

Conflicts:
- 2 Convalescence
- 3 Other Hospital
- 3 Other Hospital - wd 6
- 3 Other Hospital wd 2
- 3 Other ward
- 3 Other Acute hospital
- 4. Not applicable - patient deceased
- 5. Transferred to different Consultant - NGH

Missing data:
Cleaning the registry in

• Errors are **difficult to find** and not all can be resolved
• Excluding all imperfect data not an option
• Balance between a ‘research ready’ dataset and robust audit capability
• Needs to be **reproducible**

• It is **locked** to clinicians & researchers without being cleaned
Warning: cleaning clinical registries without experts is dangerous*

* Applies to analysing healthcare data also
UNLOCKING THE REGISTRY
MONITORING
Publication of named healthcare provider outcomes

http://www.scts.org/patients/
Publication of named healthcare provider outcomes

FILTER DATA
subset

RISK ADJUSTMENT
\texttt{glm}, \texttt{glmer} \{\texttt{lme4}\}, \texttt{mfp} \{\texttt{mfp}\}, \texttt{predict}, \texttt{auc} \{\texttt{pROC}\},

CLASSIFICATION & PRESENTATION
\texttt{ggplot} \{\texttt{ggplot2}\}, \texttt{write.csv}

AGGREGATION
\texttt{summaryBy} \{\texttt{doBy}\}, \texttt{merge}, \texttt{arrange} \{\texttt{plyr}\}
Exploratory analyses

summaryBy \{doBy\} + gvisMotionChart \{googleVis\}

http://www.scts.org/DynamicCharts/
Monitoring medical devices

• Currently does not happen in UK
• Data: 200 valve types entered 13,000 ways (free text)
• But R is good with regular expressions
UNLOCKING THE REGISTRY
RESEARCH
Evidence based medicine

Octogenarians having Mitral Valve Surgery ± CABG ± TV repair over 10-year window

`survfit + Surv {survival}`
`kmplot {by Tatsuki Koyama}`

Mean 4 patients per unit / year

No. at risk 1415 991 779 559 398 276 180 114 64 23 6
Contemporary statistical methodology for retrospective data

Probability of receiving a mechanical valve

Mechanical valve  Biological valve

Matchit \{MatchIt\}
Risk prediction: *status quo*

- Observed
- Expected
- Actual
- Overall average
- Trend

Ratio = 0.73

Ratio = 0.37
## Risk prediction: with R

### Risk prediction

<table>
<thead>
<tr>
<th>Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6.00</td>
<td></td>
</tr>
<tr>
<td>-5.75</td>
<td></td>
</tr>
<tr>
<td>-5.50</td>
<td></td>
</tr>
<tr>
<td>-5.25</td>
<td></td>
</tr>
</tbody>
</table>

|------|------|------|------|------|------|

### Methods

- No update
- Rolling 24-month window (12-months)
- Rolling 24-month window (1-month)
- Piecewise recalibration (12-months)
- Piecewise recalibration (24-months)
- Dynamic logistic regression

### Code

```r
library(dma)
```
CONCLUSIONS
Conclusions

• We need to **unlock** healthcare registries to:
  ▪ Monitor quality & avoid a repeat of Bristol
  ▪ Revalidation of professional credentials
  ▪ Facilitate patient choice
  ▪ Develop & validate evidence based medicine
  ▪ Increase in demand

• We can do it all in R!
Acknowledgements

• Funded by Heart Research UK [Grant Number RG2583]

• Dr Norman Stein, North West e-Health

Comments & suggestions

graeme.hickey@manchester.ac.uk