Unlocking a national adult cardiac surgery audit registry with R

Graeme L. Hickey1*, Stuart W. Grant2, Ben Bridgewater1,2

1. Northwest Institute of BioHealth Informatics, University of Manchester, UK
2. Department of Cardiothoracic Surgery, University Hospital of South Manchester, UK
*Contact author: graeme.hickey@manchester.ac.uk

Keywords: Cardiac surgery, Healthcare registry, Audit, Performance

Following the Bristol Royal Infirmary heart scandal, the Society of Cardiothoracic Surgery in Great Britain & Ireland (SCTS) established a world-leading clinical registry to collect data on all adult cardiac surgery procedures. To date this registry contains >480,000 records and 163 fields. The data includes patient demographics, comorbidities and clinical measurements, cardiac and operative details, and post-operative outcomes. We will describe examples of how R has been used recently to interrogate the SCTS registry and run a national governance programme for performance monitoring.

Understanding the data is vital to making decisions. The SCTS have recently used the googleVis package by Gesmann and de Castillo (2011) to visualize hospital- and surgeon-level data longitudinally over time as Google Motion Charts (SCTS, 2013a). This can be used to interrogate, for example, the risk-adjusted mortality rate of healthcare providers, whilst gaining an understanding of the variation due to sample size or inherent natural variability. It can also be used to understand the multivariate relationships between data; for example is postoperative length-of-stay related to patient age and the number of operations performed by each hospital? This tool has already been the instigator of a number of clinical and care-quality investigations.

Monitoring performance of surgeons requires a broad portfolio of tools. First, statistical modelling tools, for example glm or glmer, are required to appropriately ‘risk-adjust’ outcomes. Second, functions to aggregate and summarize the data in different ways over healthcare providers are required. Finally, graphical tools are required to present the results as funnel plots and case mix charts to patients for scrutiny of their healthcare provision (SCTS 2013b).

“Real-world” databases are messy – the SCTS registry is no exception. Cleaning data can be complicated, especially if there interdependencies between data frame rows and columns. Synonyms and homonyms required homogenizing; numerical, temporal and clinical conflicts required resolving; and duplicate records required accurate identification and removal. Previously this was a terminal obstacle facing cardiac surgeons in their bid to unlock the potential of this data. A registry-specific R package has been written to fully automate the cleaning in a transparent and reproducible manner, thus enabling analyses of the data.

References

