International Journal of Population Data Science

Using Bayesian Model Averaging to Analyse Hierarchical Health Data: model implementation and application to linked health service use data

Hassani-Mahmooei, Behrooz[1*], Janneke, Berecki-Gisolf[2], and Collie, Alex[1]

[1]Institute For Safety, Compensation and Recovery Research

[2]Monash University Accident Research Centre

Objectives

The majority of standard coding systems applied to health data are hierarchical: they start with several major categories and then each category is broken into subcategories across multiple levels. Running statistical models on these datasets, may lead to serious methodological challenges such as multicollinearity between levels or selecting suboptimal models as model space grows exponentially by adding each new level. The aim of this presentation is to introduce an analytical framework that addresses this challenge.

Approach

Data was from individuals who claimed Transport Accident Commission (TAC) compensation for motor vehicle accidents that occurred between 2010 and 2012 in the state of Victoria, Australia and provided consent for Pharmaceutical Benefits Scheme (PBS) and Medicare Benefits Schedule (MBS) linkage (n=738). PBS and MBS records dating from 12 months prior to injury were provided by the Department of Human Services (Canberra, Australia). Pre-injury use of health service items and pharmaceuticals were considered to indicate pre-existing health conditions. Both MBS and PBS listings have a hierarchical structure. The outcome was the cost of recovery; this was also hierarchical across four level (e.g. total, medical, consultations, and specialist). A Bayesian Model Averaging model was embedded into a data mining framework which automatically created all the cost outcomes and selected the best model after penalizing for multicollinearity. The model was run across multiple prior settings to ensure robustness. Monash University's High Performance Computing Cluster was used for running approximately 5000 final models.

Results

The framework successfully identified variables at different levels of hierarchy as indicators of pre-existing conditions that affect cost of recovery. For example, according to the results, on average, patients who received prescription pain or mental health related medication before the injury had 31.2% higher short-term and 36.9% higher long-term total recovery cost. For every anaesthetic in the year before the accident, post-injury hospital cost increased by 24%, for patients with anxiety it increased by 35.4%. For post-injury medical costs, every prescription of drugs used in diabetes (Category A10 in ATC) increased the cost by 8%, long term medical costs were affected by both pain and mental health.

Conclusion

Bayesian model averaging provides a robust framework for mining hierarchically linked health data helping researchers to identify potential associations which may not have been discovered using conventional technique and also preventing them from identifying associations that are sporadic but not robust.

*Corresponding Author:
 Email Address: behrooz.hassani.mahmooei@monash.edu (B. Hassani-Mahmooei)