

International Journal of Population Data Science

Journal Website: www.ijpds.org



Swansea University
Prifysgol Abertawe

Multiple Correspondence Analysis is a Useful Tool to Visualize Complex Categorical Correlated Data

Quail, Jacqueline^{1*}, Osman, Meric¹, and Teare, Gary¹

¹Health Quality Council (Saskatchewan)

Objectives

We sought to identify the most expensive hospitalized individuals in the Canadian province of Saskatchewan in fiscal year 2012/13, and determine the primary cause of their high use of health services. Our aim was to identify health problems that can be prevented or better managed in a non-hospital health care setting. Comorbid conditions are an important and confounding covariate in this population and so we used multiple correspondence analysis (MCA) to investigate the association of these conditions with each other and the most responsible diagnosis for each hospitalization. MCA is a multivariable descriptive statistical technique that displays the relationship between categorical variables in 2-dimensional graphical form.

Approach

We identified the most expensive 5% of people hospitalized between 01APR2012 and 31MAR2013. Hospital costs accounted for the majority of costs, but physician, drug, long-term care, and home care costs were added. Comorbid conditions in any of the 25 hospital diagnostic fields were identified and grouped into categories based upon ICD-10-CA subcategories. For example, category 1 was ICD-10-CA codes F10-F19: Mental and behavioural disorders due to psychoactive drug use, while category 2 was ICD-10-CA codes F20-F29: Schizophrenia, schizotypal, and delusional disorders. SASTM v9.3 was used to conduct MCA and generate graphs displaying the correlation between each comorbid condition category, where the distance of each dot from the other represents the strength of the association between the disease categories (i.e., diseases that are correlated cluster together.) The frequency of each category of comorbid condition was represented by the size of the dots on the graph (e.g., the more people with the disease, the larger the dot.) Categories of comorbid conditions were redefined based upon data findings and clinical expertise.

*Corresponding Author:

Email Address: jacqueline.quail@gmail.com (J. Quail)

Results

Three patient groups emerged as being amenable to intervention and thus cost savings, specifically (1) individuals of advanced age who are no longer able to live at home and are hospitalized while waiting for a bed in a long-term care facility, (2) individuals with a mental health and/or addiction problem, and (3) individuals who experienced medical harm during their time in hospital.

Conclusion

MCA is a valuable graphical tool that is easy to learn and, in conjunction with other statistical techniques, can be used to elucidate the relationship between complex correlated categorical variables.

