International Journal of Population Data Science

Journal Website: www.ijpds.org

Challenges of Analysing Case-Control Datasets: A Health and Homelessness Case Study

Waugh, A^{1*}, Rowley, D¹, and Clarke, A¹

¹National Records of Scotland

Background

We discuss the methodological challenges of analysing a case-control dataset.

Objectives

The study seeks to explore the relationship between health activity and homelessness in Scotland.

Methods

Our study involved 430,000 people with experience of homelessness, each matched with two controls of the same age and sex: one from each of the 20% least deprived and 20% most deprived areas. This gives 1.3 million people in total. We aimed to compare health service usage among the different groups to ascertain whether the health needs of the homeless group exceeded that of the general population, and in particular those of non-homeless deprived people.

However, as the cases were defined individually, and the controls were defined only by proxy (via datazone of residence), observed differences between the groups could simply result from differences in the proportion of the group that were deprived, rather than specifically relating homelessness itself.

To address this we compare the timing of health activity with the timing of the homelessness assessment to more directly isolate the relationship between homelessness and health. This temporal analysis also allows discussion of causation. This method raised further difficulties: effectively being a convolution of complicated functions. However as each group has the same age-sex structure the complexity applies equally. Thus direct comparisons can be made between the groups resulting in more-straightforward analysis.

Findings

Findings of the study are presented in a separate talk (The Relationship Between Health and Homelessness in Scotland) at this conference.

Conclusions

Identifying the causal relationships involved in correlations can be difficult. However by comparing the case cohort to the control cohort, and more specifically comparing the time of activity of these relative to the identifying event of the case group (the homelessness assessment), it is possible to identify some causal relationships.

*Corresponding Author:

Email Address: andrew.waugh@nrscotland.gov.uk (A Waugh)