International Journal of Population Data Science

Journal Website: www.ijpds.org

Building a high quality data linkage spine using a targeted approach to clerical review

Wiggins, Nadine^{1*} and Stokes, Brian¹

¹Menzies Institute for Medical Research, University of Tasmania

Objectives

The Tasmanian Data Linkage Unit (TDLU) was established through the University of Tasmania in 2011 with the first dataset imported to its Master Linkage Map (MLM) during 2014. Tasmania an island state of Australia, has a population of approximately 516,000. From the TDLU's earliest inception, it was deemed important to build a high quality linkage spine comprising key administrative data representative of significant state health and related datasets to support quality population level research.

Approach

The TDLU has embraced a model of continual quality and process enhancement as a determined strategy to support ongoing business improvement. Initial linkage approaches utilised 'traditional' methods of reviewing record pairs within an upper and lower confidence range. This approach resulted in false record pairs with high confidence levels being linked (false positives) and true record pairs at lower confidence levels not linked (false negatives). To improve linkage quality, the TDLU has continually refined and modified its clerical review methodology with a specialist software module developed to identify specific record attributes within groups that require the group to be manually reviewed and resolved. A range of SQL queries have also been developed to identify incorrect links and further enhance the linkage quality of the MLM.

Results

The linkage quality tools implemented have led to improved clerical review and quality assurance processes which in turn have increased the overall quality of the linkage spine. The 'targeted' method of clerical review provides easy identification of false positive records, particularly those with high confidence scores

such as twins and husband/wife combinations. The review of groups at lower confidence levels has minimised the rate of false negative pairs however further refinement of tools is required to minimise the time spent on reviewing these groups. The clerical review software module has equipped staff with the necessary information to make informed and timely decisions when reviewing groups of records. Detailed documentation is maintained for each linkage project providing continual feedback for system and process improvements as the linkage spine increases in size.

Conclusion

The process of clerical review and quality assurance requires a commitment to continual refinement of tools and techniques resulting in a higher quality linkage spine and a reduction in the total time and resource required to link datasets.

*Corresponding Author:

Email Address: nadine.wiggins@utas.edu.au (N. Wiggins)