

International Journal of Population Data Science

Journal Website: www.ijpds.org



Swansea University
Prifysgol Abertawe

Using family relationships to improve consistency of identification of Aboriginal people in linked administrative data

Gibberd, Alison^{1*}, Simpson, Judy¹, and Eades, Sandra²

¹University of Sydney

²Baker IDI Heart and Diabetes Institute

Objectives

With the life expectancy of Aboriginal Australians 10 years less than non-Aboriginal Australians, administrative health datasets are essential tools to identify disparities and track changes over time. However, the Aboriginal status of individuals is often inconsistently recorded within, and across, datasets and Aboriginal people are frequently undercounted. Therefore, information from multiple linked records is often used to resolve inconsistencies and undercount. However, very different algorithms are used and their performance is poorly understood.

We aimed to (i) assess the relative accuracy of three common algorithms by comparing how consistently family members were categorised; (ii) develop a new algorithm that incorporates information about the Aboriginal status of family members; and (iii) investigate the impact of these algorithms on estimates of health outcomes.

Approach

Our sample was people born from 1980 onwards recorded as Aboriginal at least once in 6 Western Australian datasets, as well as their Aboriginal and non-Aboriginal family members.

We applied the most inclusive algorithm, 'ever-Aboriginal' (EA, any record of being Aboriginal), and two more specific algorithms, 'multi-stage median' (MSM) and 'last record' (LR), to linked birth, birth registration, birth defects and hospital inpatient records. We also supplemented these algorithms with information from any children's birth records (EA+, MSM+ and LR+). Finally, we used records from all family members to derive a final Aboriginal status (MSM+ with family), using a range of decision rules, such as classifying someone as Aboriginal if they had two full siblings categorised as Aboriginal.

Results

EA+ classified 61% of our sample of 156,426 people as Aboriginal, compared with only 51% for LR.

MSM+ resulted in the most consistent classification of full siblings; only 8% of pairs had one sibling identified as non-Aboriginal and the other as Aboriginal, compared with 14% using EA+. MSM+ also performed best with parent-child triads; the classifications of the child and parents were inconsistent in only 8% of cases, compared with 27% using EA.

The proportions of low birthweight babies were similar for EA+, MSM+ and MSM+ with family, (10.6%, 10.8% and 10.8%, respectively). Stillbirth rates were also similar. However, an estimated 46% of mothers smoked during pregnancy for EA+, compared with 50% for both MSM+ and MSM+ with family.

Conclusion

Of the 6 algorithms examined, MSM+ assigned Aboriginal status most consistently and EA was the least reliable. Estimates of health outcomes were worse for Aboriginal people when the more consistent algorithms were used. However, the differences were generally small.

*Corresponding Author:

Email Address: agib5822@uni.sydney.edu.au (A. Gibberd)

