

International Journal of Population Data Science

Journal Website: www.ijpds.org



Swansea University
Prifysgol Abertawe

Addressing Missing Paternal Linkage in Administrative Data – A Candidate Father Approach

Amani Hamad¹, Lin yan¹, Barret Monchka¹, Oleguer Plana-Ripoll², Olawale Ayilara¹, and Lisa Lix¹

¹University of Manitoba, Winnipeg, Canada

²Barcelona Institute for Global Health – ISGLOBAL, Barcelona, Spain; Aarhus University, Aarhus, Denmark

Objective

Paternal linkages are often missing in administrative data. We propose and evaluate a method to identify candidate fathers using residence and demographic information.

Approach

We conducted a population-based cohort study of Manitoba adults born 1977–2017 with maternal linkages. We modelled the probability of missing paternal linkage using key covariates for births post-1980, when missingness was highest. We applied this model to births pre-1980, when missingness was lowest, to exclude those with the highest predicted probability of missing linkage, at missingness levels of 10–50%. We identified candidate fathers in the same postal code area; up to three candidates were randomly selected using parental age-gap weights. We measured the percentage of candidate father sets containing the true father and estimated hazard ratios (HR) and 95% confidence intervals for paternal transfer of mental disorder and ischemic heart disease (IHD) for true versus candidate fathers.

Results

The cohort included 140,001 births; 25.8% lacked paternal linkages, and 26.8% were born before 1980. Only 27–28% of candidate sets included the true father across missingness levels. For mental disorders, the true-father HR was 1.27 (1.23–1.31); candidate-father HRs ranged from 1.17 (1.13–1.22) at 50% to 1.24 (1.20–1.29) at 10% missingness. For IHD, the true-father HR was 1.94 (1.35–2.79); candidate-father HRs ranged from 1.64 (0.92–2.89) to 1.91 (1.32–2.75).

Conclusions

Our method had low accuracy in identifying true fathers but produced similar risk estimates for true and candidate fathers, especially at lower missingness levels, suggesting utility for intergenerational research with incomplete paternal linkage.

