International Journal of Population Data Science

Journal Website: www.ijpds.org

Algorithms for assessing person-based consistency among linked records for the investigation of maternal use of medications and safety

Tran, Duong^{1*}, Havard, Alys¹, and Jorm, Louisa¹

¹Centre for Big Data Research in Health, UNSW Australia

Objective

Perinatal records linked to pharmaceutical claims and other administrative data provide a powerful resource to investigate maternal use of medications and safety. In this population-based project, data quality assessment was performed on the consistency of linked records brought together from several data collections to ensure reliable links for data analysis.

Approach

Perinatal data for the Australian states of New South Wales (NSW) and Western Australia (WA) were linked to pharmaceutical claims by a Commonwealth integrating authority, while linkage to hospital admission, emergency department (ED), mortality and congenital notification data was performed by respective State-based data linkage units. All de-identified records belonging to a unique person ID were sorted chronologically. To assess the consistency of unique persons, both within and across States, algorithms were developed based on pregnancy plurality and birth order, gestation, parity, maternal age and sex, date of delivery, dates of health service use, and State where the claim was made.

Results

The dataset included 595,456 NSW and 188,014 WA mothers with respectively 937,344 and 295,095 pregnancies delivered between 2003 and 2012. The information brought together through linkage was highly consistent for the majority of mothers and infants. Inconsistencies are identified in 742 cases, including negative inter-pregnancy period, highly illogical parity, highly inconsistent maternal age, maternal gender being systematically recorded as male, use of health services after date of death, and different infants sharing a common ID. These cases will be excluded from analyses. Date of delivery was corrected for

667 pregnancies, using date of birth recorded in the infant's admission and ED records, and date of the mother's admissions. Among admission and ED records, over 8000 needed correction in infant age due to typographical errors, 1820 were duplicates, while 1000 had discrepancies between dates of birth, date of admission and separation. There were 455 deaths, mostly neonates, identified by status of admission or ED discharge but not recorded in mortality data. The majority of these deaths were confirmed by the status of neonatal discharge at birth. There were 3404 women who had a single unique ID according to Commonwealth linkage but more than one unique IDs according to State-based linkage. 2827 mothers had births recorded in both NSW and WA.

Conclusion

Quality assessment indicated high consistency among linked records. The set of algorithms developed in this project can be applied to similar linked perinatal datasets to promote a consistent approach and comparability across studies.

*Corresponding Author:

Email Address: danielle.tran@unsw.edu.au (D. Tran)