

International Journal of Population Data Science

Journal Website: www.ijpds.org



Linkage of Pregnant People and Infants - Pseudonymized Hospital and Ambulatory Care Data for Surveillance

Amro Hassan¹, Shelly Sital¹, Andrew Hamilton¹, and Jena Wallander Gemkow¹

¹AllianceChicago, Chicago, USA

As maternal mortality and morbidity continue to worsen in the US, better data is needed to comprehensively understand the problem, especially in underserved populations. Surveillance in safety-net populations is complicated by care fragmentation and interoperability challenges between primary care and hospital settings. AllianceChicago (AC), a Federally Qualified Health Centers (FQHCs) network, supports a data infrastructure in Chicago that is linked to a Federated Data network (CAPriCORN). Leveraging electronic health record (EHR) data collected by AC and CAPriCORN, the objective of this effort was to protect privacy while connecting infant and maternal data. Utilizing multiple approaches and technologies for Pregnant People-Infant (PPI) pair identification, a deterministic matching algorithm was developed. A 5-step method for linkage was used where each step was recorded

within the database and the confidence interval was calculated. The initial pregnant person universe included 122,700 patients. The matching rate for the linkage steps ranged from 7.2% to 86.2%, and the final accuracy rate was 94.3%. The final sample (2014–2023) included a match of N=70,048 pregnant people and N=36,205 infants. Through each step of the process, the confidence interval increased with the increase of matching nodes, highlighting its accumulative accuracy. The result demonstrates the feasibility of linking maternal and baby characteristics using a range of clinical and demographic variables captured in pseudonymized hospital and ambulatory data. Our approach improves the capacity of EHR data for surveillance while protecting patient privacy and preserving data confidentiality. Triangulating outcomes recorded in different settings can help improve data quality for research and surveillance.

