Benefits and Challenges of Linking Electronic Medical Record and Claims Data for Perinatal Research

Alison M. Stuebe, Christine Tucker, Renée Ferrari, Sarah Verbiest, Virginia Pate, Michele Jonsson-Funk

Introduction

Mothers of medically fragile infants (MMFI) must recover from birth while attending to an infant in the neonatal intensive care unit (NICU). There is limited information on the postpartum health needs and access to services for MMFI.

Methods

- An honest broker identified a subset of members with a qualifying claim, defined as a childbirth-related procedure code.
- Context-based blocking and fuzzy matching on combinations of first and last name were applied.

Results

- A 97% match rate (n=1,642) was achieved
  - Exact match = 94%
  - Fuzzy match on combinations of first and last name = 3%
- Contributors to successful linkage
  - Having a finite time-period
  - Adding anesthesia procedure codes
- Challenges with linkage
  - Changes in the mother’s medical record number (MRN) when querying EMR data on the same woman longitudinally
  - Data entry errors on the delivery record
  - Care for baby showing up on mother’s claim

Lessons Learned

- Dynamic nature of EMR data
  - During the two-year study period, medical record numbers of study participants changed as they were merged and updated, resulting in missing data and changes in sample size.
- When errors such as wrong date of delivery were corrected in the analytic data set, date of delivery could not be used to merge additional EMR queries.
- Multiple nested levels of granularity
  - When querying the EMR for perinatal data, there are multiple levels of granularity.
    - Women receiving care in a hospital system
      - > Pregnancy in the lifetime of each woman
      - > Babies born from the same pregnancy
        - Twins, triplets, etc. add complexity to linking records.
    - Time of delivery will differ
    - Mode of delivery may differ
    - Birth outcome (liveborn, still birth, preterm) and date of delivery may differ for multiples

Levels of granularity in perinatal data

- A single pregnancy may result in more than one birth, with more than one mode of delivery. EMR queries at the baby level may result in duplicate documentation at the pregnancy level.

Conclusions

- Linking claims to EMR data can elucidate women’s health utilization in the postpartum period by allowing ascertainment of encounters outside of the hospital.
- Our findings contributed to improved data quality initiatives in the hospital’s EMR system and laid the groundwork for future studies using EMR data.

This study was supported by grant 5R00HD071022-02 from the National Institutes of Health, National Institute of Child Health and Human Development. We acknowledge the editorial assistance of the NC Translational and Clinical Sciences (NC TraCS) Institute, which is supported by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, through Grant Award Number UL1TR002375.