Title

Privacy-preserving record linkage in community settings: merging HMIS & EMR data

Abstract

Background

Optimizing the use of siloed data can facilitate the integration of social and medical care. Privacy-preserving record linkage (PPRL) allows individuals to be matched across datasets using cryptographically hashed information instead of direct identifiers.

Objective

We used PPRL to explore patterns of acute healthcare utilization among a cohort of people experiencing homeless, by merging electronic medical records (EMR) from a regional healthcare system in Oregon with data from the local HUD-designated community action agency's Homeless Management Information System (HMIS).

Methods

The EMR dataset included measures of healthcare utilization for which we had complete coverage in the region (urgent care visits, ED visits, and hospitalizations). The HMIS dataset included assessments of social history and vulnerability. We merged on a hashed ID created from first name, last name, and date of birth.

Results

A total of 775 individuals were matched between both datasets. Individuals in the merged dataset were 58% male and 88% white. Individuals with a high vulnerability index score were more likely to have two or more ED visits in 2022 compared to people with a low vulnerability index score (48% vs 36%, $p=0.090$), and were more likely have two or more urgent care visits as well (11% vs 6%, $p=0.341$). Individuals defined as chronically homeless were more likely to have multiple ED visits (43% vs 29%, $p<0.001$) and multiple hospitalizations (9% vs 3%, $p=0.003$) than patients not classified as chronically homeless.

Conclusion

PPRL offers a practical, low-cost way to merge medical and social data in a deidentified manner. This technique is feasible in community-based settings and has potential to inform targeted interventions to improve population health outcomes.