Patients Who Visit Multiple Sites in the HealthLNK Clinical Data Research Network

Margaret Madden[1] and Marc Rosenman, MD[1,2]

[1]Center for Health Information Partnerships, Northwestern University, Chicago, Illinois
[2]Department of Pediatrics, Feinberg School of Medicine, Northwestern University, Chicago, Illinois

Background

- A clinical data research network (CDRN) with several EHR-data-contributing institutions in a single city offers the opportunity to describe healthcare use across institutions and to mitigate the incomplete follow-up seen in single-institution studies.
- CDRNs thereby also have been used to study "care fragmentation," based on a patient's having EHR records from >1 institution. But a CDRN itself may reflect incomplete data, either sent in that way by the contributors and/or as a result of the CDRN's procedures for merging clinical data elements and patient identities across the institutions.
- And no CDRN comprises every healthcare facility in a city—every CDRN is, ipso facto, incomplete. Such limitations will become increasingly appreciated.

Data Sources

- HealthLNK repository
  - A de-identified assembly of EHR data of adults age 18-89
  - Seven large health care institutions in Chicago
  - A forerunner to the Chicago Area Patient-Centered Outcomes Research Network (CAPriCORN) PCORnet CDRN

Objectives

- Compare a cohort of patients who visited >2 health care institutions with a cohort who visited 1 health care institution
- Describe characteristics associated with a higher volume of cross-over between health care institutions
- Contribute to the discussion of *the metadata of clinical data research networks*

Methods

- Retrospective (2006-2012) cohort study of 6 large health care institutions from HealthLNK. Figure 1 shows the number of institutions visited per person for the 6 institutions analyzed.
- We calculated the number of diagnosis records by patient by institution. Among those with records in exactly two institutions, we calculated each patient's record counts in her/his two institutions.
- We similarly compared patients with ≥12 records in each of ≥3 institutions versus those who visited exactly 1 institution with ≥36 records there.
- International Classification of Diagnosis, Ninth Revision (ICD-9) diagnosis data from six institutions (four academic medical systems, one county hospital and clinic system, and one community health center network) in the HealthLNK
- At its inception HealthLNK had merged and had de-duplicated patient identifiers across the institutions.
- HeathLNK's method for individual-level linkage via de-identified hashing and matching now has been adopted by CAPriCORN and other PCORnet CDRNs.

Figure I. Number of Institutions Visited per Patient Based on the Diagnosis Table

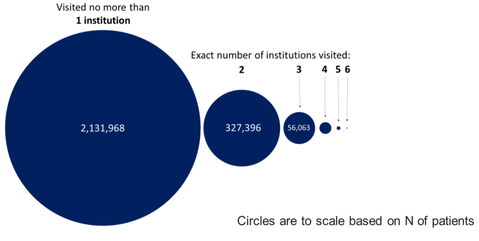
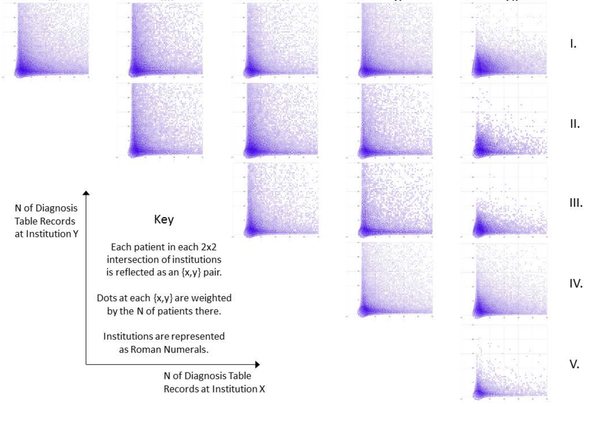


Figure II. Number of Institutions Visited per Patient Based on the Diagnosis Table

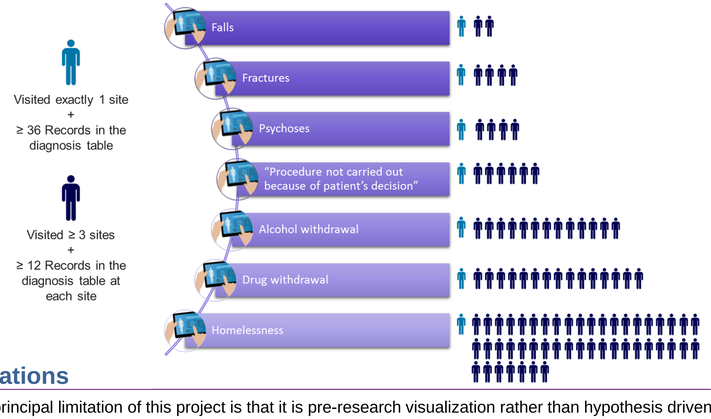


Results

- 2,524,019 patients with ≥1 record in the Diagnosis Table in the 7-year period:
  - 82% had data in no more than 1 institution.
  - 16% (392,051) had records from ≥2 institutions
    - Most of these (327,396 [84% of 392,051; 13% of 2,524,019]) had records in exactly 2 institutions.
      - Among the 327,396, 40% had ≤2 records in the less-frequently-visited institution in her/his pair (see Figure 2).
  - 2.6% (64,655 of 2,524,019) had records in ≥3 institutions.
- Compared to patients who visited exactly one institution and had ≥36 records there, those with ≥12 records in each of ≥3 institutions had higher rates of various ICD-9 diagnosis codes or categories (see Figure 3).
- Among the small population with an encounter at all six institutions, more than 50% had an ICD-9 code for depression, and more than 40% for lack of housing, drug abuse, chronic alcoholism, and suicidal ideation.

Figure III. Comparison of Patients in 1 versus ≥3 Institutions

Ratio of the prevalence of various diagnoses between the two groups (patients in exactly 1 site as the reference group)



Limitations

- The principal limitation of this project is that it is pre-research visualization rather than hypothesis driven work. Nevertheless, we think that this type of work is necessary in order to support the more specific research projects that will follow in city-wide networks.
- Future analyses will consider covariates, and reasons for visiting >1 institution, to further examine these findings

Conclusions

- A high volume of records in ≥3 institutions was associated with psychosis, trauma, and homelessness.
- Based in part on this work, we have begun participating in a city-wide discussion of a multi-institution collaboration for addressing the health of the homeless population.