Enhancing State-Based Data with Federally-Available Data: Linking Massachusetts Hospital Discharge Data to ResDAC

Zi Zhang, MD, MPH

Catherine Nwachukwu, MPH

Huong Trieu, PhD

Mark Paskowsky, MPP

NAHDO Annual Conference 2020

CENTER FOR HEALTH INFORMATION AND ANALYSIS

CHIA.

Agenda

- **Background**

- Methods

- Next Steps

- Questions & Answers

CHIA.

Background

- State and local health data organizations have widely used state-specific hospital discharge records and available claims data to analyze and report on health system performance

- In Massachusetts, discharge summary records from inpatient, observation, and emergency department visits are used to measure key performance measures, such as hospital readmissions and revisits

- Limitations of discharge summary records:
    - Expected payer source(s) and potential miscoding
    - Determination of primary vs. secondary payer source(s)
    - Unreliable or incomplete charge information
    - Limited or no patient enrollment information

CHIA.

Background – Patient Identifier / Payer Issues

Missing or Invalid SSN by Payer Type
SFY 2011 - 2018

Medicare
- 1.4%
- 1.7%
- 2.1%
- 2.7%
- 3.1%
- 4.5%
- 3.6%
- 4.3%

Medicaid
- 8.3%
- 8.0%
- 8.8%
- 9.5%
- 10.7%
- 13.1%
- 12.9%
- 13.4%

Commercial
- 5.8%
- 6.8%
- 8.2%
- 10.1%
- 11.0%
- 12.9%
- 13.2%
- 15.1%

(Y-axis: Missing or Invalid SSN, 0% to 14%; X-axis: Year SFY, 2010 – 2018)

Social Security Number is increasingly missing from acute care hospital discharge data.

CHIA.

Background – Patient Identifier / Payer Issues

- Issues with linking patients by Social Security Number between Case Mix (acute care hospital data) and ResDAC:

  - 82% of patients with Medicare indicated as their payer (based on MA acute care hospital case mix data) matched a ResDAC Medicare beneficiary

    Note: expected about 95+%

  - 59% of patients, aged 65+, who did not indicate Medicare as their payer (based on MA acute care hospital data) were actually found in the ResDAC Medicare data

    Note: expected low %

CHIA.

Background – CHIA's Linkage Project

- To address these challenges, the Massachusetts Center for Health Information and Analysis (CHIA) has undertaken a three-phase linkage process:

Phase 1: Complete
- Link patients within the Massachusetts Acute Hospital Case Mix Databases (Case Mix), including inpatient, observation, and emergency department data

Phase 2: Near Completion*
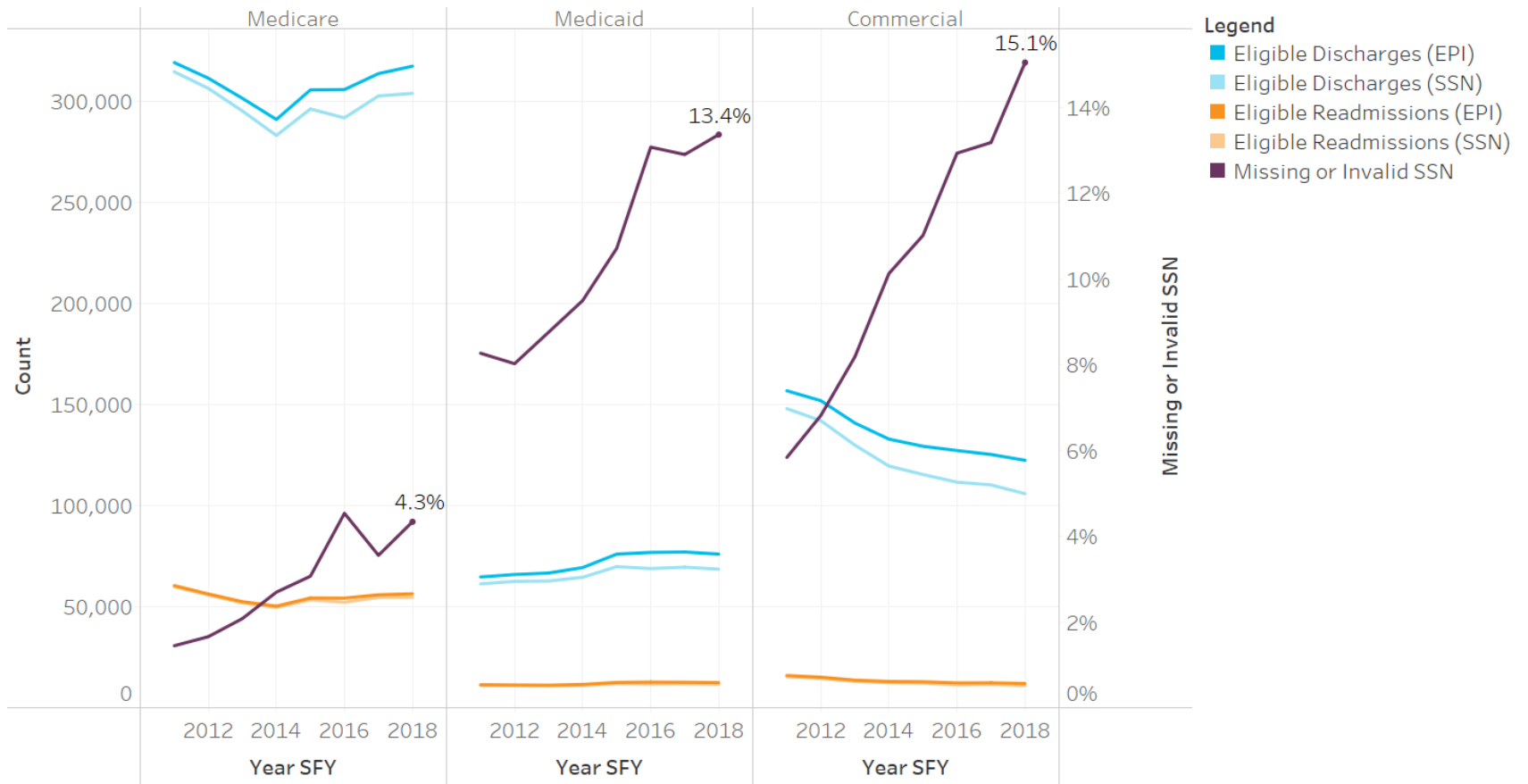- Link Case Mix to the Massachusetts All Payer Claims Database (APCD)

Phase 3: Beginning Soon*
- Link the federally-available ResDAC data to the linked Case Mix-APCD data from Phase 2

*Delayed due to COVID-19

CHIA.

Background – CHIA's Linkage Project Phase 1

Trend in Discharges and Readmissions by Payer Type using EPI vs. SSN
SFY 2011- 2018

Methods

- **<u>Goal:</u>** To create a Master Patient Index (MPI) by linking individuals across data sources (Case Mix, APCD, and ResDAC) and across time.

- **<u>Linking Fields:</u>** Social Security Number, date of birth, gender, zip code, name and address (if available), and other internal patient IDs such as medical record number, health plan subscriber ID, beneficiary ID.

- **<u>Steps:</u>**
    1. Begin with the APCD-Case Mix MPI data.

    2. Adapt the APCD-Case Mix probabilistic matching algorithm from Phase 2 by developing a new scoring matrix that satisfies all three data sources' requirements for linking individuals.

    3. Add ResDAC Medicare Beneficiary Summary Data and run the algorithm to match patients across all three databases.

CHIA.

Methods – MPI Data Elements

The combined MDM algorithm will use these elements from the Case Mix, APCD eligibility, and ResDAC data:

Case Mix Elements	APCD Elements	ResDAC Elements
Patient SSN	Member SSN	Beneficiary SSN
Patient Date of Birth	Member Date of Birth	Beneficiary Date of Birth
Patient Gender	Member Gender	Beneficiary Gender
Patient Zip Code	Member Zip Code	Beneficiary Zip Code
Patient Healthplan ID (from membership card)	Member Healthplan ID (Carrier-specific member ID)	Beneficiary Healthplan ID (Beneficiary ID)
Patient First Name	Member First Name	
Patient Last Name	Member Last Name	
Patient Address		

CHIA.

Methods – Match Rules (1)

First, patients are matched deterministically within data sources:

Data Source	Match Description
Case Mix	Records with the same OrgID and Medical Record Number
APCD	Records with the same OrgID and Carrier-Specific Member ID (CSUMID)
ResDAC	Records with the same Beneficiary ID

CHIA.

Methods – Match Rules (2)

Then, all other data elements are used to probabilistically match records:

Scenario	Description
1	All elements agree.
2	**Any single** element disagrees/is missing, all other elements agree.
3	SSN and DOB agree, all other elements agree but any two elements (not NAME) disagree.
4	DOB missing, all others agree but any one element (not SSN) disagrees.
5	SSN and DOB are missing, all other elements agree.
6	SSN, DOB, and HealthPlanID agree, all other elements are missing.

CHIA.

Methods – Special Considerations

Some insights we learned from Phase 2 Linkage:

Scenario	Description	Comments
1	Different First Names, DOB missing, all others match	Could be same gender twins were born, given mom's SSN, DOB missing
2	Different Last Names, SSN missing, all others match	Without SSN present can't tell if LN change is the same person
3	Different Last Names, SSN disagrees, all others match	If SSN disagrees, can't tell if LN change is the same person
4	Different First and Last Names, all others match	Two different names, same town, same gender, same DOB, SSN transposed
5	Different First and Last Names and addresses, all others match	Not enough to tell if same person, transposed SSN in a diff. location

CHIA.

Methods – Linkage Quality Checks

- Count and percentage of hospitalizations that match between case mix and ResDAC
  - What percentage of demographic characteristics match?
  - Does the payer / dual eligibility status match across the data sources?

- Count and percentage of hospitalizations that do not match between case mix and ResDAC
  - What are the demographic and payer characteristics of those that don't match?

- Explore data issues for subgroups with low match rates

CHIA.

Next Steps

- Start the implementation of the Phase 3 linkage process next month

- Use the linked database from Phase 3 to explore and produce more reliable and meaningful reporting on health system performance:
  - More accurate payer information

  - More accurate patient eligibility/enrollment information, including dual-eligibility status

  - Better payment data for various services

  - Improve/refine CHIA's current reporting, such as annual reporting of hospital readmissions and revisits

  - More opportunity to produce additional and better reporting on health system performance and health care cost and trends

CHIA.

Questions?

Enhancing State-Based Data with Federally Available Data | NAHDO 2020

CHIA.

Contact Information

- For additional questions, please contact:

  Zi Zhang, MD, MPH

  Senior Director of Research

  Center for Health Information and Analysis

  zi.zhang@state.ma.us

  Catherine Nwachukwu, MPH

  Associate Manager of Research

  Center for Health Information and Analysis

  catherine.nwachukwu@state.ma.us