Use of Unique Beneficiary IDs in Medicaid Data Analyses

Medicaid Innovation Accelerator Program - National Webinar

January 25, 2018
3:00 PM – 4:00 PM EDT

1

Logistics for the Webinar

- All lines will be muted
- Use the chat box on your screen to ask a question or leave a comment
  - Note: chat box will not be seen in "full screen" mode
- Slides will be posted online

Welcome!

- Jessie Parker, GTL and Analyst on Medicaid IAP Data Analytic Team, Data and Systems Group, CMCS
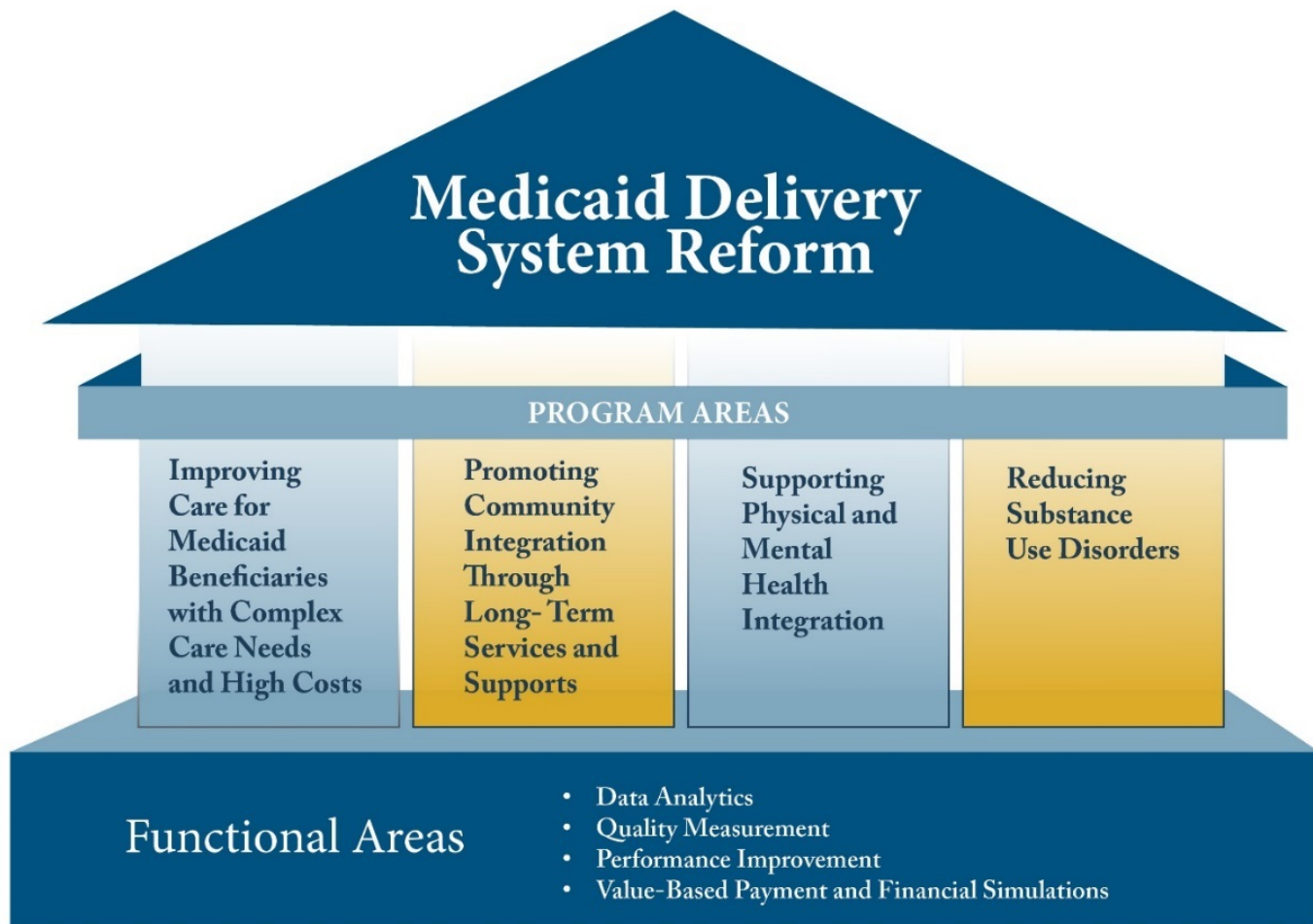
Today's Speakers

- Manjusha Gokhale, Senior Data Scientist, Truven Health Analytics, an IBM Company

- Bruce Greenstein, Chief Technology Officer, U.S. Department of Health and Human Services

Agenda for Today's Webinar

- Introduction

- Overview of the Medicaid Innovation Accelerator Program (IAP)

- Working with Beneficiary Identifiers (IDs)

- Linkage Across Data Sources

- National Death Index

- Takeaways from Today's Webinar

Medicaid Innovation Accelerator Program (IAP)

Medicaid Delivery System Reform

PROGRAM AREAS

Improving Care for Medicaid Beneficiaries with Complex Care Needs and High Costs

Promoting Community Integration Through Long-Term Services and Supports

Supporting Physical and Mental Health Integration

Reducing Substance Use Disorders

Functional Areas

- Data Analytics
- Quality Measurement
- Performance Improvement
- Value-Based Payment and Financial Simulations

Goals for Today's Webinar

In this interactive webinar, states will learn about:

- challenges in working with Medicaid enrollment data

- linkage methods

- linking to the National Death Index (NDI)

- examples of other linkages with state data

Use of Unique Beneficiary IDs in Medicaid Data Analyses

Challenges and Strategies

Manjusha Gokhale, Senior Data Scientist, Truven Health Analytics, an IBM Company

Beneficiary IDs in Medicaid Data

- Accurate identification of unique individuals is important for program administration, oversight, and analytics

- Analyses which require correct denominator information include:

    - utilization analysis and comparison to benchmarks

    - assessment of expenditures

    - population health analysis

Medicaid Enrollee Identifier Assignment

- Medicaid enrollee identifiers are assigned by each state Medicaid agency.

- This identifier is assigned during enrollment along with highly identifiable information including:

    - social security number (SSN)

    - date of birth (DOB)

    - first name

    - last name

    - gender

    - address

Medicaid Enrollee Identifier Issues

- If you simply count the number of unique Medicaid enrollees identifiers in a year, you would likely get a number which was different than the total number of Medicaid enrollees.

- This is due to known issues with enrollment which include:

    - carve-outs for managed care, behavioral health, pharmacy coverage

    - combined mother/baby claims at birth

    - disenrollment / re-enrollment

Medicaid Enrollee Identifier Issues: Specialty Carve-outs

- <u>Specialty carve-outs</u> are arrangements where the state has contracted a third-party entity to administer the care given for certain services.

  ➤ <u>Issue</u>: Presence of multiple enrollee identifiers.

  ➤ <u>Recommendation</u>: Maintain a crosswalk of specialty carveout enrollee identifiers to state Medicaid enrollee identifiers.

Medicaid Enrollee Identifier Issues: Vertical Carve-outs

- <u>Vertical carve-outs</u> are where the state has contracted with an organization to administer care, such as Medicaid Managed care plans.

  ➤ <u>Issue</u>: Individual is listed in Medicaid enrollment, but they could also be assigned another internal enrollment identifier by the health plan. The individual's utilization is not in the Medicaid claims.

  ➤ <u>Recommendation</u>: Maintain a crosswalk of vertical carveout enrollee identifiers to state Medicaid enrollee identifiers. Exclude these individuals in any cost or use analyses with Medicaid claims.

Medicaid Enrollee Identifier Issues: Combined Mother/Baby Enrollment

- <u>Mother/Baby</u>: Healthy babies are usually not enrolled in Medicaid at the time of birth.

  ➢ <u>Issue</u>: Some current enrollment methods undercount healthy babies in Medicaid enrollment.

  ➢ <u>Recommendation</u>: Confirm the number of infant enrollees by augmenting figures with information from birth records and hospital discharge claims.

Medicaid Enrollee Identifier Issues: Disenrollment/Re-enrollment

- <u>Disenrollment/Re-enrollment</u>: Some individuals will disenroll from Medicaid and later re-enroll and get assigned a different Medicaid enrollee ID.

  ➢ <u>Issue</u>: The same individual is represented several times in the enrollment data.

  ➢ <u>Recommendation</u>: Use Social Security Number to confirm that an individual does not have prior Medicaid enrollee ID.

Master Patient Index Definition

- Master patient index is a method of aggregating the information from disparate sources.

- The master patient index should contain only those fields which uniquely identify an individual (e.g. Medicaid ID, SSN, date of birth, gender).

- Ideally, Medicaid enrollee information should be consolidated into a master patient index.

Deterministic vs Probabilistic Matching

- Deterministic matches are exact matches

- Probabilistic matching uses a statistical approach and calculates the likelihood of a match as in the examples below:

Phonetics	Synonyms	Abbreviations	Concatenation	Misalignment
Katherine vs. Catherine	Andrew = Andy George = Jorge 1st = First Jr = Junior	Jerome David Salinger ~ J.D. Salinger	Van de Velde = Vandevelde	Robert John = John Robert

Edit Distance	Transliteration	Date Similarity	Proximity	Noise Words
876-5309 ~ 876-5390	Toyota = トヨダ	01/01/1973 ~ 01/02/1973	Geocodes and great-circle distance	IBM Co. = IBM

Deterministic Matching

- Advantages

  - Confidence of match

  - Easy to understand and explain

  - Not dependent on knowledge of data file

  - Can use all matching fields and then drop criteria one-by-one to capture remaining non-matches

- Disadvantages

  - Rigid in structure

  - May undercount denominator

  - Can exclude common errors such as contractions of name, address changes

Probabilistic Matching

- Advantages

  - Can match across fields which may contain transcriptions, multiple spellings, address changes

  - Ability to maintain a longer longitudinal enrollment file

- Disadvantages

  - Difficult to describe

  - Can be hard / expensive to implement

  - May have false matches / non-matches

  - Highly dependent on patterns in database

Master Patient Index
Sample Enrollment File

ID	SSN	DOB	First Name	Last Name	Address	City	Date	Plan
A948	123-45-6789	7/7/2007	Anita	Chen	5 Lexington Ave	Watertown, MA	1/1/17	Medical
M948	123-45-6789	7/7/2007	Anita	Chen	5 Lexington Ave	Watertown, MA	1/1/17	RX
A948	123-45-6789	7/7/2007	Anita	Chen	5 Lexington Ave	Watertown, MA	2/1/17	Medical
M948	123-45-6789	7/7/2007	Anita	Chen	5 Lexington Ave	Watertown, MA	2/1/17	RX
A948	123-45-6789	7/7/2007	Anita	Chen	5 Lexington Ave	Watertown, MA	3/1/17	Medical
M948	123-45-6789	7/7/2007	Anita	Chen	5 Lexington Ave	Watertown, MA	3/1/17	RX
A948	123-45-6789	7/7/2007	Anita	Hines	5 Lexington Ave	Watertown, MA	4/1/17	Medical
M948	123-45-6789	7/7/2007	Anita	Hines	5 Lexington Ave	Watertown, MA	4/1/17	RX
M948	123-45-6789	7/7/2007	Anita	Hines	6 Watertown Ave	Lexington, MA	5/1/17	Medical
A948	123-45-6789	7/7/2007	Anita	Hines	6 Watertown Ave	Lexington, MA	5/1/17	RX
M948	123-45-6789	7/7/2007	Anita	Heinz	6 Watertown Ave	Lexington, MA	5/1/17	Mgd Care

Deterministic vs Probabilistic Matching
cont.

- Deterministic matching on SSN would result in a single person – e.g., Anita Heinz.

- However, if we did not have SSN or DOB and matched on last name, first name, address and city, we would end up with three people – Anita Chen, Anita Hines, and Anita Heinz.

- Probabilistic linkage would allow us to have one person – Anita Heinz – even without SSN or DOB.

Polling Question

Has your state agency used any of the following when working with beneficiary IDs?

- Probabilistic matching

- Deterministic matching

- We have used both probabilistic and deterministic matching methods

- We have not used either approach

Recommendations

- Establish a hierarchy of linkage

- Examine the matches for confirmation and examine a set of non-matches to view patterns in errors

- Loosen the match criteria and check to see whether correct people matched

Recommendations (cont'd)

- Enrollment data which is prone to transcription errors and name changes would be a good candidate for probabilistic linkage

- When creating a master patient index, think about the purpose of creating such a file (e.g. longitudinal analysis)

- Compare results to previously reported benchmarks

Linkage to Other Sources

- Once the Master Patient Index is created, one can use this to link to a number of different sources including administrative health claims, electronic health records, vital statistics and others.

- Both deterministic and probabilistic techniques can be used to linkage between data sources.

Linkage to National Death Index (NDI)

- The CDC National Death Index is a nationwide compilation of state death records

- Researchers can apply to use the Index

- If approved, the researchers send the National Center of Health Statistics (NCHS) an password-protected encrypted file with identifying fields using the structure specified

- NCHS matches the state research file with the NDI and returns results files

Results from NDI Linkage

- Results from NDI Linkage will be complex with multiple files and multiple linkages to a single person.

- CDC provides guidance on how to interpret your results

- Reference: https://www.cdc.gov/nchs/ndi/index.htm

Example 1: Selecting NDI Data
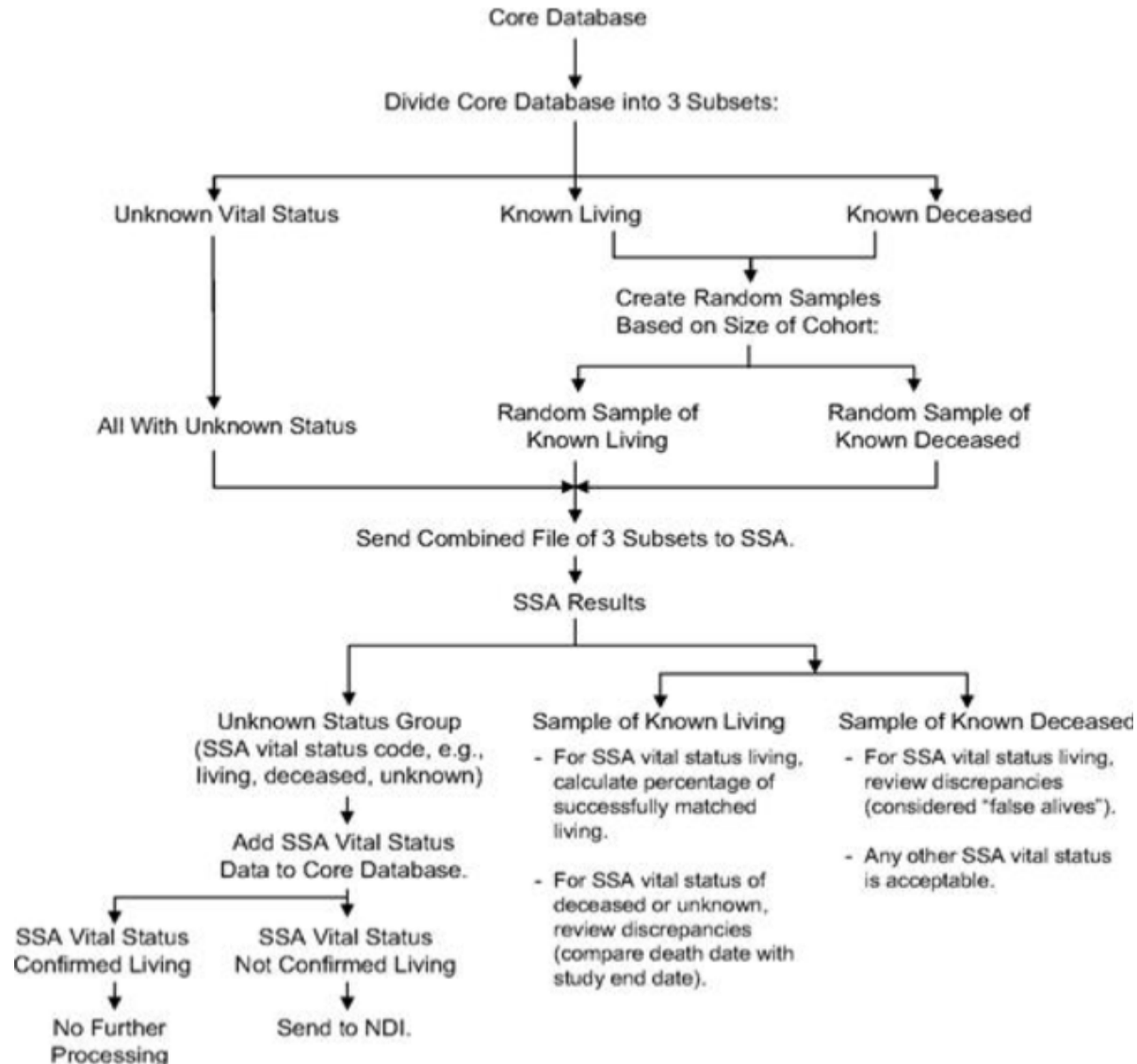
Example 1: Assessing Match

Strata	Criteria
1	Exact match
2	Exact SSN and sex match
3	Exact SSN match
4	8-digit SSN and sex match
5	7-digit SSN and sex match
6	6-digit SSN and sex match
7	5-digit SSN and sex match
8	Valid user SSN, missing NDI SSN, name/DOB/sex match
9	Valid user SSN, missing NDI SSN, name, sex, DOB month and day match, DOB year within 1
10	Valid user SSN, missing NDI SSN, phonetic name, DOB, and sex match
11	Name, DOB, and sex match
12	Name, sex, DOB month and day match, DOB year within 1
13	Phonetic name, DOB, and sex match
14	Exact DOB match
15	Last name, first name, DOB month, sex match, DOB year within 10

Example 1: Results

1 (Strata)	2 (Criteria)	3 (No. of NDI Records Meeting Criteria)	4 (No. of Keeper Death Certificates Requested From States)	5 (No. of Death Certificates Received From States)	6 (No. of Death Certificates of Individuals Kept After Manual Screen)	7 (No. of Death Certificates of Individuals Accepted Into Study Cohort)
1	Exact match	1,778	1,778	1,778	1,778	1,778
2	Exact SSN and sex match	1,173	1,130	1,103	1,103	1,035
3	Exact SSN match	159	116	99	74	30
4	8-digit SSN and sex match	71	71	71	71	69
5	7-digit SSN and sex match	26	26	26	26	24
6	6-digit SSN and sex match	21	18	17	8	8
7	5-digit SSN and sex match	105	88	80	6	6
8	Valid user SSN, missing NDI SSN, name/DOB/sex match	23	23	23	21	21

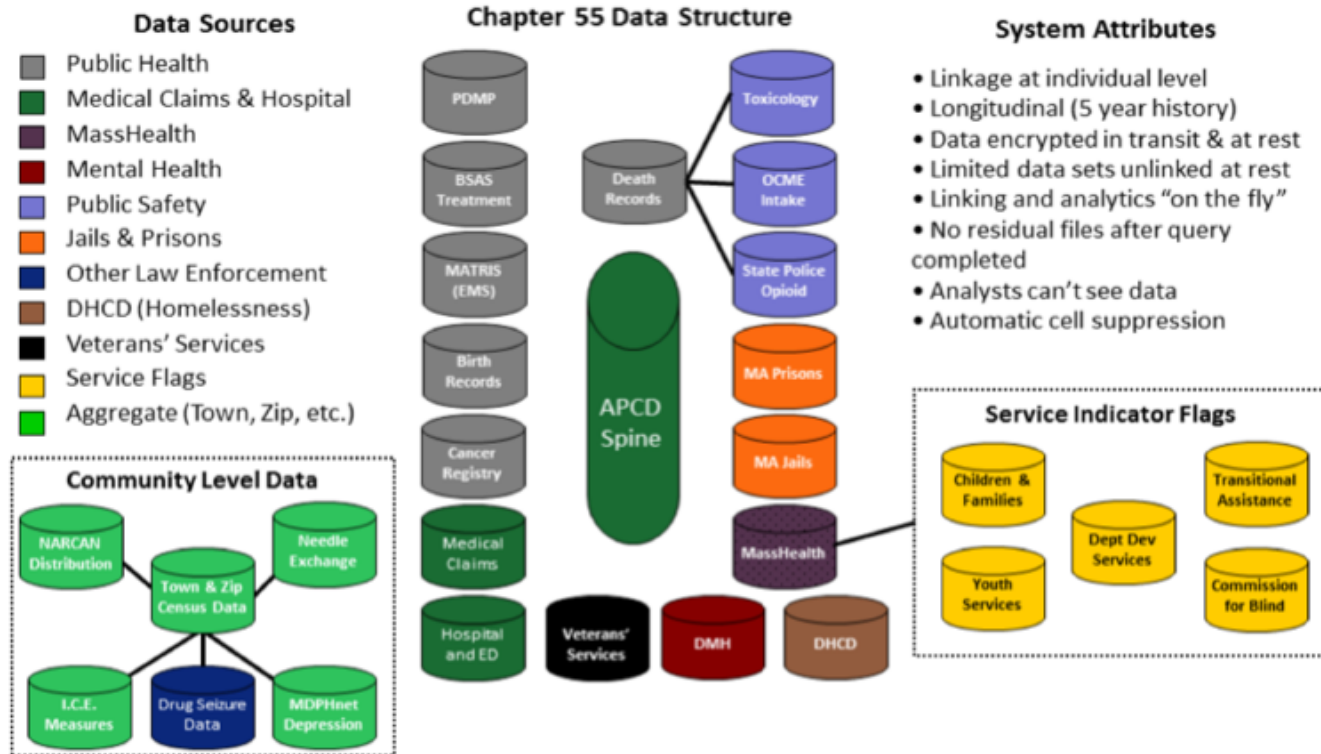# Example 1: Results (continued)

1 (Strata)	2 (Criteria)	3 (No. of NDI Records Meeting Criteria)	4 (No. of Keeper Death Certificates Requested From States)	5 (No. of Death Certificates Received From States)	6 (No. of Death Certificates of Individuals Kept After Manual Screen)	7 (No. of Death Certificates of Individuals Accepted Into Study Cohort)
9	Valid user SSN, missing NDI SSN, name, sex, DOB month and day match, DOB year within 1	1	1	1	1	1
10	Valid user SSN, missing NDI SSN, phonetic name, DOB, and sex match	6	6	5	3	2
11	Name, DOB, and sex match	45	39	37	30	30
12	Name, sex, DOB month and day match, DOB year within 1	43	31	29	3	3
13	Phonetic name, DOB, and sex match	80	70	56	12	11
14	Exact DOB match	167	123	105	13	9
15	Last name, first name, DOB month, sex match, DOB year within 10	8,396	1,060	956	8	6
Total		12,094 records	4,580 records	4,386 death certificates	3,157 death certificates	3,033 death certificates

Example 2: Opioid Data Mapping in Massachusetts

Chapter 55 Data Mapping

Data Sources
- Public Health
- Medical Claims & Hospital
- MassHealth
- Mental Health
- Public Safety
- Jails & Prisons
- Other Law Enforcement
- DHCD (Homelessness)
- Veterans' Services
- Service Flags
- Aggregate (Town, Zip, etc.)

Community Level Data
- NARCAN Distribution
- Needle Exchange
- Town & Zip Census Data
- I.C.E. Measures
- Drug Seizure Data
- MDPHnet Depression

Chapter 55 Data Structure
- PDMP
- BSAS Treatment
- MATRIS (EMS)
- Birth Records
- Cancer Registry
- Medical Claims
- Hospital and ED
- Death Records
- APCD Spine
- Toxicology
- OCME Intake
- State Police Opioid
- MA Prisons
- MA Jails
- MassHealth
- Veterans' Services
- DMH
- DHCD

System Attributes
- Linkage at individual level
- Longitudinal (5 year history)
- Data encrypted in transit & at rest
- Limited data sets unlinked at rest
- Linking and analytics "on the fly"
- No residual files after query completed
- Analysts can't see data
- Automatic cell suppression

Service Indicator Flags
- Children & Families
- Transitional Assistance
- Dept Dev Services
- Youth Services
- Commission for Blind

Example 2: Opioid Linkage

Match Level	Identifiers To Be Matched
1	Exact match on first name, last name, Social Security number, gender, birth date, street address #1, street address #2, town of residence, and zip code.
2	Exact match on last name, Social Security number, gender, birth date, town of residence, and zip code.
3	Exact match on Social Security number, gender, and birth date.
4	Exact match on first name, last name, gender, birth date, street address #1, street address #2, town of residence, and zip code.
5	Exact match on first name, last name, gender, birth date, town of residence, and zip code.
6	Exact match on first name, last name, gender, and birth date.
7	Exact match on first name, last name, gender, and birth date
8	First and third letters of first name, first and third letters of last name, gender, birth date
9	Street address #1, street address #2, town of residence and zip code
10	Exact match on first name, last name, and birth date

References for Examples

- Example 1:

  Nancy C. Wojcik, Wendy W. Huebner, Gail Jorgensen. Strategies for Using the National Death Index and the Social Security Administration for Death Ascertainment in Large Occupational Cohort Mortality Studies, *American Journal of Epidemiology*, Volume 172, Issue 4, 15 August 2010, Pages 469–477. Available at:

  https://academic.oup.com/aje/article/172/4/469/84682

- Example 2:

  Commonwealth of Massachusetts Chapter 55 Opioid Report. Available at:

  https://www.mass.gov/service-details/chapter-55-overdose-report

About HHS

- The U.S. Department of Health and Human Services (HHS) is the nation's principal agency for protecting the health of all Americans and providing essential human services. The Department includes CMS, CDC, FDA, NIH, AHRQ, HRSA, SAMHSA

- ~ 79,540 employees

- The Office of the Chief Technology Officer, located in the Immediate Office of the Secretary, provides leadership and direction on **data**, **technology**, **innovation** and **strategy** across the Department of Health and Human Services

Source: hhs.gov

Office of the Chief Technology Officer
Innovation, Design, Entrepreneurship, Action

The Office of the Chief Technology Officer

Data Insights Institute

Leveraging data to generate better health and human services insights

- HHS ReImagine Initiative
- Code-a-thons
- Enterprise Data Strategy

Industry Partnerships

Driving partnerships across external, international, and HHS

- Reimagine RO
- Health Datapalooza
- Kidney Innovation Accelerator (ASN)
- Global Digital Health Partnership
- Start Up, Entrepreneurial, Private-Public Partnerships

Innovation IDEA Lab

Leadership, Council, and Execution

- HHS Ignite Accelerator
- Secretary's Ventures Fund
- Entrepreneur-in-Residence
- HHS Open Innovation

Office of the Chief Technology Officer
Innovation. Design. Entrepreneurship. Action

HHS.gov
U.S. Department of Health & Human Services

Louisiana State Overview

- Population is approximately 4.6 million

- Strong Governor Model – Cabinet Agencies

- Department of Health includes Medicaid, Medicaid Eligibility, Public Health, Behavioral Health, Aging

- Doesn't Include Human Services, Eligibility

- Constant struggle to keep up with budget pressures

State Priorities

- Reduce provider burden
- Minimize human capital requirements to state agencies
- Rising to challenges of new system procurements, and regulatory complexities
- Implementing new laws and programs

Programmatic Structure

- Medicaid Program – move to managed care
- Programs within the department vs. programs in other departments
- Programs within Medicaid
- Programs run by Public Health Department
- Programs in other departments

Program Data vs Enterprise Data

- Beyond Eligibility Data – A dualopoly

- How far does the Medicaid ID go?

- What is the common client index?

- Health Plans – where do they fit? How is the data handled?

- Same people, many programs

Practical Applications

- Death data, opioid reporting, and validity
- Birth data, claims data, and outcomes
- Immunizations, health plans – registries and claims

Death and Beyond

- Vital records and data
- Capturing and sharing death data
- Making vital record's death module data more actionable
- Constructing a positive Return on Investment (ROI) for enhancing vital records systems

Try to Keep Up with Us

- @HHSCTO
- @HHSIDEALab
- www.hhs.gov/IDEALab
- IDEALab@hhs.gov

Office of the Chief Technology Officer
Innovation. Design. Entrepreneurship. Action

HHS.gov

U.S. Department of Health & Human Services

Questions?

Takeaways

- An awareness of administrative issues with Medicaid IDs can improve the success of an analytic project

- Probabilistic matching is often most appropriate for linking Medicaid IDs across administrative records

- Constructing a positive ROI for enhancing records systems may support evidence-based policy decisions

Thank You

Thank you for joining today's webinar!

Please take a moment to complete

the post-webinar survey.

For more information & resources, contact
MedicaidIAP@cms.hhs.gov