Health IT infrastructure needs to support population health improvements in Colorado

Identity Management

Arthur Davidson, MD, MSPH

*Denver Public Health*

*eHealth Commission*

Office of eHealth Innovation (OeHI)  
Colorado’s State Designated Entity  
Wednesday, May 11, 2016  
303 17th Ave, Conference room 10A  
Denver, CO
Objectives

- Establish the problem(s)
- Review purpose and functions of a statewide
  - Master Patient Index (sMPI)
  - Master Provider Index (sMPri)
- Present some next step options
Identity Management: No sure method to know and uniquely identity a client/patient/provider.

<table>
<thead>
<tr>
<th>Client/Patient</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMV</td>
<td>DMV</td>
</tr>
<tr>
<td>Vital statistics</td>
<td>Institution A ….n</td>
</tr>
<tr>
<td>CORHIO</td>
<td>Payer B ….n</td>
</tr>
<tr>
<td>QHN</td>
<td>RCCO C….n</td>
</tr>
<tr>
<td>APCD</td>
<td>DORA</td>
</tr>
<tr>
<td>CIIS</td>
<td>Medicare</td>
</tr>
<tr>
<td>State OIT (justice, education, social services)</td>
<td>NPI</td>
</tr>
</tbody>
</table>

Statewide identity management is a ‘team’ sport...
No team means ineffective identity management causing:

- Failed attribution
- Inaccurate integration assessments
- Inaccurate monitoring

**Problem**

SIM Advisory Board Orientation Packet (12/22/15)

- **Payment Reform**: Develop and implement *value based payment models* that incent integration and improve quality of care.
- **Practice Transformation**: Support practices as they accept new payment models and *integrate behavioral and physical health care*
- **Population Health**: Engage communities to reduce stigma, *promote prevention, and remove barriers* to accessing care
Estimated Prevalence of Children and Youth with Obesity (2012-2014): City and County of Denver

This map displays the percentage of children and youth 2-17 years of age with obesity, by census tract, based upon body mass index (BMI) measurements contributed to the Colorado BMI Monitoring System from multiple health care organizations. BMI was calculated from height and weight and plotted on the Centers for Disease Control and Prevention (CDC) male or female BMI-for-age growth chart to determine a percentile. Obesity was defined as a BMI at the 95th percentile or higher. The percentage with obesity was calculated for each census tract by dividing the number of children and youth with a BMI at the 95th percentile or higher by the total number of valid BMI measurements available in that census tract.

These data are comprised of children and youth 2-17 years of age who reside in the City and County of Denver. The total number of children and youth represented with valid BMI measurements is 79,493, or 60% of the total children and youth population of 133,508, as reported from the 5-year American Community Survey (2010-2014). Census tracts were designated as having insufficient data for display if: 1) the total population was fewer than 50 individuals; or 2) fewer than 50 valid BMI measurements were available; or 3) coverage was less than 20% of the population 2-17 years of age.

Denver County Average 18%

1% - < 10%
10% - < 15%
15% - < 20%
20% - 26%
Range: 1% - 26%

To request additional maps and/or aggregated data tables, please contact LeeAnn Rohn at: LeeAnn.M.Rohn@kp.org
BMI Data Contributors: Denver Health, Kaiser Permanente Colorado, Children's Hospital Colorado, Salud Family Health Centers and High Plains Community Health Center

Census Tract Boundary Data: US Census 2010
Map Created March 2016
The Colorado Department of Public Health and Environment provides data hosting and cartographic services for the Colorado BMI Monitoring System.

Funded By The Colorado Health Foundation
and Kaiser Permanente Colorado Community Benefit
Matching Across Institutions - 2009

2,471,441 \rightarrow 1,852,396 (75\%)
“The hub will... leverage the existing Master Patient Index (MPI), provider directories and other tools. Building on clinical information, the phased approach will link to administrative claims information via the APCD and other sources as needed, providing a central aggregated clinical and cost data hub.”
Identity Management - Functions

- Regular *automated receipt* of patient/client/provider identifying information from multiple partners
- Data are *standardized* for storage in the statewide master patient/provider index (sMPI/sMPrI)
- *Quality assurance* is performed on data with feedback to the data contributors/partners (e.g., remove duplicates)
- Process to *disambiguate records* is carried out (e.g., resolve potential overlaps across institutions)
- *Tools* are available for managing these processes and feedback to/from the data contributors/partner organizations

© 2015 Denver Public Health
Disambiguation

A process of establishing a single semantic or meaning

- Matching process
- Resolves multiple potential matches
- Uses attributes of individuals (patients or providers) registered at multiple organizations

**GOAL: find all matches for one target individual**
Deterministic indexing: *perfect but inflexible matching*

- *False positives: none*    *False negatives: high*
- search based on an exact match of some combined factors (e.g., name, social security number, date of birth, and/or sex).
  - *Mickey Mouse, 11/18/28, M = Mickey Mouse, 11/18/28, M*

Probabilistic: *improves match by anticipating data entry errors/variance*

- *False positives: adjustable*    *False negatives: adjustable*
- rules-based search mechanism with some subset of exact matching
  - *Mickey Mouse, 11/18/28, M = Mick Mouse, 11/18/28, M*
  - *Mickey Mouse, 11/18/28, M = Mickey Mouse, 11/18/29, M*
  - *Mickey Mouse, 11/18/28, M = Micky Mouse, 12/18/28, M*
String Matching (record comparisons)

- Exact Rules-based
- Rules-based
- Probabilistic (including Statistical Frequency Analysis)
- Machine Learning

String Frequencies (uniqueness)

Record Matching (record comparisons)

- String and Substring
- Character Transposition
- Soundex
- NYSIIS
- String Edit Distance
- Bipartite Graphs

String Matching (field comparisons)

- Exact only

“Fuzzy logic” (via rules and/or data massaging)
Electronic Linking Cause:

Records *seem* to match

Resulting error: **false positive** (overlay)
2 records linked under 1 MRN

Records *should* match

Resulting error: **false negative** (duplicate)
2 MRNs created
Compare the various data sources:

- a step-by-step procedure for solving a mathematical problem that frequently involves repetition of an operation especially using a computer

- mathematical formula using a combination of weighted MPI/MPrI data elements to determine the probability of a duplicate or overlap

Same institution

Different institutions
Definitions

**Duplicate Entry/File:** (undesirable and propagated)

- more than one entry/file for the same patient or person (Rates around 9-15% •; 7-40% ●)
  - *Mickey Mouse incorrectly has both record numbers 001 and 100 at Disneyland Clinic*
  - may represent information capture errors

**Overlay Entry/File:** (undesirable and propagated)

- more than one distinct individual assigned to the same record or identification number in a facility's MPI. (Among 2 hospital [n=5000] samples: 1 or 2 = rate of 0.02 – 0.04%)
  - *Mickey Mouse and Donald Duck incorrectly share record 001 at Disneyland Clinic*

**Overlap Entries/Files:** (function of sMPI/sMPri)

- more than one MPI entry/file for the same patient/provider in two or more facilities across the state
  - *At Disneyland Clinic, Mickey Mouse has record 001 and record 100 at Disneyworld Clinic*
  - algorithm works to identify and resolve overlaps without creating overlays

- Grannis, Overhage, and McDonald 2004
- Initiate Systems, Inc. 2008
Chance of False Positive Matches
Large Demographic Database (80M)

Source: Social Security Death Master File
Note: Numbers in the blocks such as 1/1.02 in the leftmost dark block means there is 1 chance in 1.02 tries of a false positive match in database when this key type (LN) is used. Moving to the right in the diagram ‘DOB 1/3.5’ means 1 chance in 3.5 tries of false positive match when both last name and date of birth are used.
<table>
<thead>
<tr>
<th>Data Element</th>
<th>DMV</th>
<th>CORHIO</th>
<th>QHN</th>
<th>APCD</th>
<th>.....</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Name</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Middle Name</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Last Name</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Gender (M / F / U)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Address 1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>City</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>State</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Zip</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>County</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phone</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>SSN (Last 4 digits only)</td>
<td>X</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Date of birth</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Unique ID</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Payer number</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X+</td>
</tr>
<tr>
<td>MRN</td>
<td></td>
<td>X+</td>
<td>X+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
eMPl Maintenance

- Need reporting mechanisms to maintain sMPl/sMPrI accuracy
  - Cost-effective use of staff time as adds to collective accuracy
- Established procedures:
  - Data sources:
    - receive/respond to potential duplicate reports
    - identify known non-duplicates
      - identify family members (twins, multi-generation name sharing)
    - share “never match” flag
  - sMPl/sMPrI:
    - rapidly remediate incorrectly matched records
  - Resolve intra-partner reports before determining overlaps
“... emphasis on a **collaborative approach** that shares data and insights across boundaries to drive better, more efficient medical practice and patient care.”

“... drive the process of discovery as a natural outgrowth of patient care... to **ensure innovation, quality, safety, and value** in health care “

Support (90/10 match) for HIE Architecture

“The free flow of information is hampered when not all doctors, facilities or other practice areas are able to make a complete circuit. Adding long-term care providers, behavioral health providers, and substance abuse treatment providers, for example, to statewide health information exchange systems will enable seamless sharing of a patients’ health information between doctors or other clinicians when it’s needed.”


Provider Directories: with an emphasis on dynamic provider directories that allow for bidirectional connections to public health and that might be web-based, allowing for easy use by other Medicaid providers with low EHR adoption rates

Development of a Master Patient Index (should be cost allocated)
Conclusions

• To be successful CO needs identity management solutions:
  – for SIM (near-term)
  – for proper attribution with payment reform
  – to accurately measure interventions and population health
  – to be a vibrant learning health system (long-term)

• Identity management is complex
  – need to establish robust tools and procedures

• CO has an opportunity (and has been encouraged by CMS) to use 90/10 funding to build out a statewide:
  – master patient index
  – master provider index
Questions/Discussion

adavidson@dhha.org