**Monthly APCD User Workgroup Webinar**

April 22, 2014

**Agenda**

1. Release 2.1
2. Master Patient Index
3. Applying for Release 2.1
4. Questions/SME Office Hours

**Release 2.1 Features**

* Same data as in Release 2.0
	+ Dates of service (2009-12 paid thru June 13)
	+ No new data submissions
* Important data enhancements
	+ Highest version flags for three more payers
* Master Patient Index on CY 2011 and 2012 data (Including 6 months of run out).

**Master Patient Index – Summary**

* Custom probabilistic matching algorithm based on MA APCD data submissions
	+ IBM Initiate Software
	+ Refined by 12,000+ manual reviews of proposed matches by CHIA staff
* Resulting data elements
	+ MEID – Member Enterprise ID
	+ MCL - Matching Confidence Level

**The Problem**

* Over 100 sources of data
* Rolling 24 months of eligibility
* 65M records/month
* Probabilistic model:
	+ Low tolerance for false positives
	+ Minimize false negatives
	+ No manual reviews

**Data Elements Used in Model**Partial List

* + Member Name
	+ Member SSN
	+ Member DOB
	+ Member Address
	+ Member Gender
	+ Subscriber SSN
	+ CarrierSpecificUniqueMemberID and Organization ID

**Comparison Components**

* False negatives are often caused by variation in data – nicknames, spelling error, transpositions, etc.
* The comparison functions compare attribute values from two records and determine how closely they match
* They are empirically derived
	+ When you compare attributes, what types of errors do you observe?
	+ How do you measure “close”?
* IBM Initiate has a library of these functions which include …
	+ General routines which can be applied to arbitrary attributes
		- Edit distance – 555059552 and 554045952 are 2 apart
		- Resolution-based – postal codes agree exactly or on the first n digits
	+ Specific routines based on IBM experience
		- Person names – Exact match, nickname match, phonetic match, name-to-initial match, edit distance – check all alignments
		- Address – Consider different weights for numeric tokens versus alpha tokens
* Source: IBM

**Modeling Process**

* The IBM MDM Weight Generation Process
* Manual Adjustment and Adding of Weights
	+ Based on feedback from Sample Pairs Review process
	+ Over 12,000 samples reviewed by hand
* Creation of filters to reduce false positives, for example to assure the following types of matches can not be linked even with some similar data fields:
	+ Father/Son, Mother/Daughter
	+ Siblings
	+ Twins

**The Matching & Linking Process**

1. **The MDS receives a record added to source**
2. **The MDS standardizes the record…**
3. **… and creates derived data**
4. **The MDS selects candidates from the DB**
5. **The MDS scores the comparisons**
6. **And links the inbound record**

**Result: Member Enterprise ID**

* 19 digit number
* Assigned by the system to each “entity group”
* No embedded intelligence in the number
* Appears on the Member Eligibility and Claims files

**Entity Group Simplified Example**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Name** | **Address** | **DOB** | **Gender** |
| MC Record 1 | Charles M Smith | 18 Main Street | 12/12/94 | M |
| MC Record 2 | Charles Smith | 18 Main St | 12/12/95 | M |
| MC Record 3 | CM Smith | 18MainSt | 12/12/94 | M |
| ME Record 1 | Charlie Smith | 19 Main St | 12/2/94 | Unknown |
| ME Record 2 | Charles Smith | 18 Main | 12/12/94 | F |
| PC Record 1 | Charles N Smith | 18 Main Street | 12/12/84 | M |

**Master Patient Validation**Distribution of Member Enterprise IDs by Age
YE 2011 and 2012

**Master Patient Validation**Number or Org IDs for Member Enterprise ID in 2011 and 2012

**Master Patient Validation**Member Enterprise ID Compared to US Census Estimate
by MA 3 Digit Zip Code

**Matching Confidence Level**

* Assigned by CHIA to each entity group
* Based on lowest match value…weakest link
* Three values

|  |  |
| --- | --- |
| **Value** | **Meaning** |
| 2 | High Confidence Level |
| 1 | Low Confidence Level |
| 0 | SingletonMember Enterprise ID appears in just one file type |

**Among Topics CHIA is Analyzing**

* Whether singletons are more prevalent in certain payers and/or file types
* Linkages between carve out vendors and primary risk holders
* Turnover, persistency, etc.

**Among Topics CHIA is Analyzing**

* Whether singletons are more prevalent in certain payers and/or file types
* Linkages between carve out vendors and primary risk holders
* Turnover, persistency, etc.

**Receiving Release 2.1**

* New documentation and application forms available on/around April 30th
* Fulfillment times are not expected to change (3 months from initial application to delivery of data, on average)
* You can apply for Release 2.1 on current forms and start the review process now

**May Dates**

* – Monthly APCD Technical Assistance Group Meeting
* – Monthly APCD User Workgroup Webinar
* – Data Release Committee (DRC) Meeting

**Questions?**

* General questions about the APCD:

 (CHIA-APCD@state.ma.us)

* Questions related to APCD applications: (apcd.data@state.ma.us)
* Questions related to Casemix: (casemix.data@state.ma.us)