

Monthly APCD User Workgroup Webinar

April 22, 2014

Agenda

- I. Release 2.1
- II. Master Patient Index
- III. Applying for Release 2.1
- IV. 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

		Matching Decision	
		Match	Don't Match
Truth	Same Member	Correct Decision	False Negative
	Different Member	False Positive	Correct Decision

Adapted from IBM

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

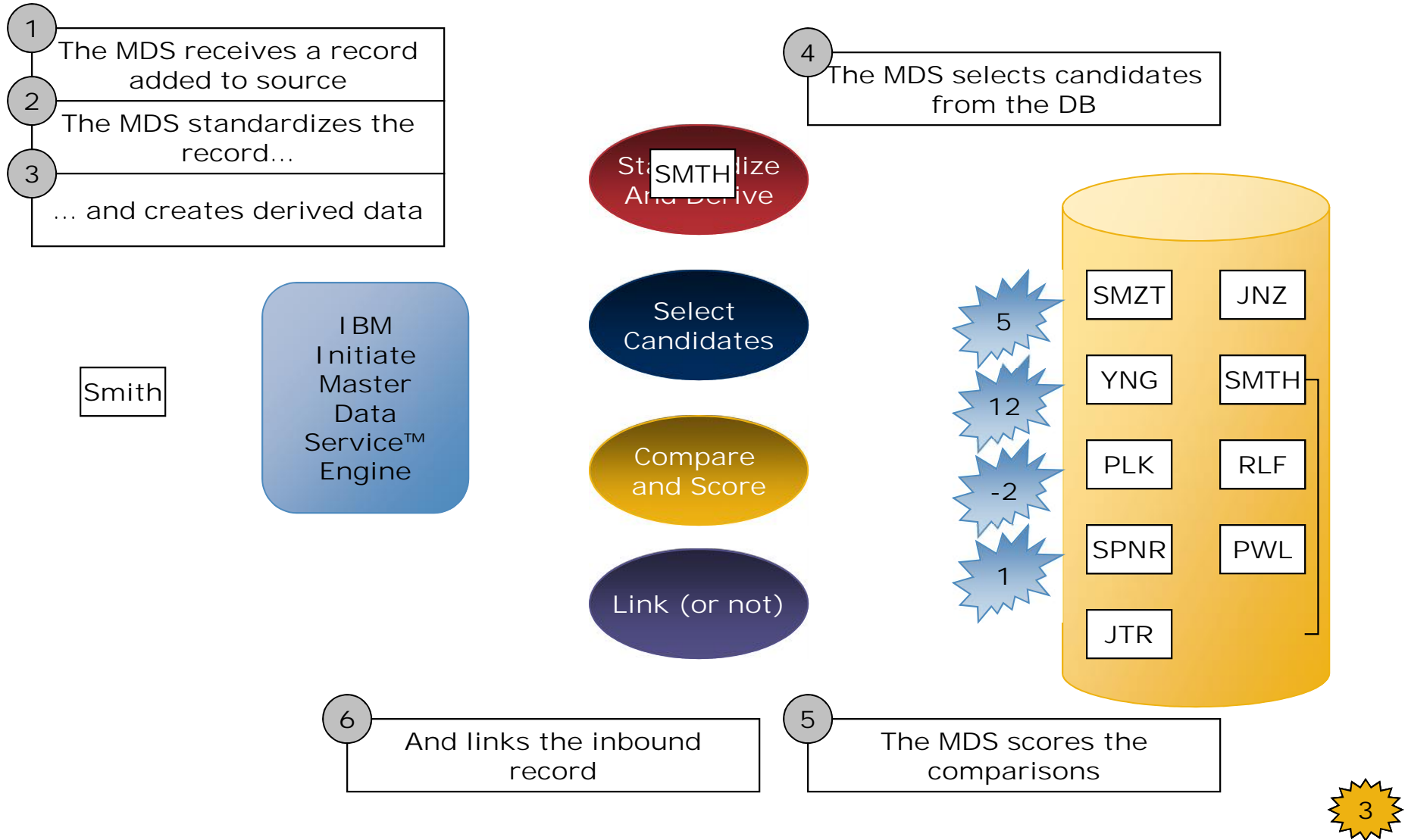
- False negatives are often caused by variation in data – nicknames, spelling error, transpositions, etc.
- The comparison functions compare attribute values from two records and determine how closely they match
- They are empirically derived
 - When you compare attributes, what types of errors do you 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



Source: IBM

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

	As of 12/31/11	As of 12/31/12	
<1 yr	1.3%	0.9%	
1-17	19.9%	19.8%	
18-24	8.4%	8.3%	
25-44	25.6%	25.6%	
45-64	26.3%	26.6%	
65-74	5.9%	6.6%	
75-84	3.4%	3.5%	
85+	2.2%	2.5%	
Unknown	7.0%	6.0%	
Total	100.0%	100.0%	

Master Patient Validation

Number of Org IDs for Member Enterprise ID in 2011 and 2012

Number of Org IDs for each Member EIN	ME	MC	PC	DC		
1	56.9%	73.4%	84.5%	95.6%		
2	19.4%	21.0%	13.2%	4.2%		
3	13.5%	4.7%	1.9%	0.2%		
4	5.9%	0.8%	0.2%	0.0%		
5	2.7%	0.1%	0.0%	N/A		
6	1.1%	0.0%	0.0%	N/A		
7	0.4%	0.0%	0.0%	N/A		
8	0.1%	0.0%	N/A	N/A		
9	0.0%	0.0%	N/A	N/A		
10+	0.0%	N/A	N/A	N/A		

*Excludes carve outs.

Master Patient Validation

Member Enterprise ID Compared to US Census Estimate
by MA 3 Digit Zip Code

Census for MA Compared to Enterprise IDs in Member Elig File				
Note: Excludes Age Unknown				
3 Digit Zip	US Census Estimates for 2012		APCD/ 2012 Census Estimate	
	Total	Percent	All	Under 65
021xx	1,290,378	19.7%	0.97	0.98
018xx	714,744	10.9%	0.98	0.99
027xx	526,927	8.0%	0.91	0.92
019xx	478,051	7.3%	0.97	0.97
010xx	470,242	7.2%	0.91	0.90
023xx	441,700	6.7%	0.96	0.96
024xx	408,494	6.2%	0.97	0.98
020xx	401,255	6.1%	0.97	0.97
017xx	380,720	5.8%	0.97	0.98
015xx	373,327	5.7%	0.99	0.98
014xx	212,623	3.2%	0.95	0.95
016xx	189,019	2.9%	0.99	1.00
011xx	169,081	2.6%	1.07	1.08
026xx	148,466	2.3%	0.96	0.97
012xx	130,447	2.0%	0.94	0.94
025xx	116,149	1.8%	0.96	0.98
013xx	83,344	1.3%	0.94	0.95
022xx	25,624	0.4%	0.71	0.69

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	Singleton Member 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.

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)