

## HL7 Immunization User Group

Monthly Meeting September 13, 2018 2:00 PM ET

#### Agenda

- Welcome
  - Poll: Which perspective do you primarily identify yourself with?
- Updates
  - SISC Update
- Patient Matching discussion



## SISC Update

Mary Woinarowicz



## Patient Matching

Nathan Bunker



## Patient Matching

- In-scope for discussion today:
  - Patient matching/deduplication/searching
  - How IIS matching/searching functions impact HL7 functions
  - How IIS architecture and matching processes impact EHRs
- Out-of-scope for today:
  - Vaccination deduplication
  - Specific technical strategies for IIS to reduce duplicates



## Patient Matching - Introduction

- Introduction to basic problem
- Common IIS Solutions
- How it impacts HL7 interfaces
- Questions for the group



- Creating a consolidated immunization history is the primary mission for IIS
- First step is to properly identify and match patient records
- First question from IIS: How many duplicates are in our IIS?
  - If we knew this answer then our problem would be solved
- Why can't we identify duplicates?
  - IIS do not collect DNA, fingerprints, photos, etc.
  - IIS do not have the staff to independently verify
  - All IIS have is a small number of fields sent by submitters



- Two mistakes possible:
  - Correct match is not identified
  - Incorrect match is identified
- Example

Name	Born	Mother	Phone	Dose #1	Dose #2
Avalon Harden	04/16/2018	Mary Harden	567-1039	Hep B 4/16	Pentacel 6/20
Avelin Harden	04/16/2018	Mary Harden	567-1039	Hep B 4/16	Pentacel 6/20



- There are three possible matching outcomes:
  - Confident that the two records do NOT match
  - Confident that the two record do match
  - Unable to determine if records are a match or not



- HL7 query searches add another layer of complexity
  - Conceptually at a different level than patient matching within the IIS
  - Assumes that all patients have a single consolidated record in IIS
  - Requires of advanced searching capability from IIS
  - May different matching process for queries than for updates
- HL7 does not allow sender to define matching criteria
  - Sender should send all data known about the patient
  - IIS may use MRN's, IIS lds, name, dob, and other information in the query to find a match
  - More information should not hurt results



- Query limits
  - IIS can set limit on the number of potential matches returned
  - EHR can also set limit and ask IIS to enforce
  - IIS enforces the lower of the EHR or IIS limit
- Policy impacts
  - Some IIS do not return potential matches
  - Some of these will indicate that this happened
  - Others will indicate that no match was found



- Most IIS originally developed as web applications
  - Most data was entered manually by clinicians
  - Clinician would be required to search for match first
  - Some data might be loaded by batch processes
    - Paper data entry
    - Flat file submission
- Batch processes became increasingly automated
  - Identify data quality problems, screen out bad data
  - Merge good data into IIS
  - Review records that couldn't be automatically merged in



- HL7 is now the most common method to submit data
  - HL7 support often built on top of old batch import processes
  - Design decisions made 20 years ago may impact current processes
- Updates and queries are core IIS functions
  - Improvements and changes here are difficult and complex



- Many IIS follow the add, merge, or queue for review model
  - Patient record is not updated with new data until a definitive decision is made on whether the patient is new or not
  - If the match process is not sure then the record is not added until it can be reviewed by IIS staff
  - Review might take from days to weeks to complete
- Other IIS add or update even if manual review is needed
  - Record submitted is immediately available to be queried back
  - Manual review may later identify records that can be merged
  - Ideal model for real-time HL7 interfaces



- Query interfaces add another layer of matching
  - IIS may leverage same matching process for updates
  - IIS may use a different matching process
    - Use sets of one or more queries on key fields
    - Submit sets of field to a probabilistic engine
  - IIS may optimize searches by ID
    - This type of search can be optimized, so it operates quickly
    - Often will still very other fields match as well to prevent fishing
  - IIS may take these potential matches and filter and sort by likelihood of match
- Some IIS never return multiple possible matches



- Search is specified by local implementations
  - Simply query
  - Probabilistic algorithms
- Submitters should send all they know about patient
- Submitter can not control query protocol used

Query Statement ID (Query ID=Z44):	Z44
Туре:	Query
Query Name:	Request Evaluated History and Forecast
Query Trigger (= MSH-9):	QBP*Q11*QBP_Q11
Query Mode:	Both
Response Trigger (= MSH-9):	RSP*K11*RSP_K11
Query Characteristics:	The query parameters may include demographic and address data. No sorting is expected.
	This profile does not specify the logic used when searching for matching clients/patients. The query parameter contents may be used for simple query or as input for probabilistic search algorithms. The search methodology should be specified by local implementations.
Purpose:	The purpose is to request an evaluated immunization history and forecast for one client.
Response Characteristics:	<ul> <li>In the case where no candidates are found, the acknowledgement response will indicate that no candidates were found.</li> </ul>
	<ul> <li>In the case where exactly one high-confidence candidate is found, an evaluated immunization history and forecast will be returned.</li> </ul>
	<ul> <li>In the case where one or more clients are a lower-confidence match for the criteria sent, the acknowledgement response will indicate no matches and no records will be returned.</li> </ul>
	<ul> <li>In the case where receiving system can't process the query, the receiving system will indicate an error in an acknowledgement.</li> </ul>
Resed on Segment Dattern	NA



- Expect IIS to return a consolidated view of the patient demographic record
  - May include data from other providers
  - May include data submitted
- MIROW: Consolidating Demographic Records and Vaccination Event Records
  - http://repository.immregistries.org/resource/consolidatingdemographic-records-and-vaccination-event-records/



- IIS may not return all patient identifying fields
  - Some fields may never sent back
    - Mother's maiden name
    - Race, Ethnicity
  - Might not include data sent by other submitters
    - Address, phone
  - Patient information should only be used for identifying matches
    - IIS do collect patient demographics for matching purposes
    - But do not necessarily see themselves as the point place for distributing patient level updates and changes



- Some IIS may accept a record positively but the record is not available when queried
  - It is queued for processing (minutes)
  - It may need to be reviewed before merging (days)
- EHR includes too much information on patient and does not get back match expected
  - IIS performs more strict searches when data is included
- EHR is unable to get back exact match
  - Can get list of possible matches but unable to construct query to get back single match



## Questions about Matching Algorithm(s)

- For incoming data (VXU's) what sort of patient matching algorithm(s) do you use?
  - For example, exact matches for certain demographic elements (e.g., first name, last name, DOB, gender)?
  - A weighted matching algorithm considering a number of different demographic elements?
- When responding to queries (QBP) what matching/searching strategies do you use?



#### Questions about MRN

• What role does the submitter MRN play in patient matching?

- Is there any sort of mechanism where a previously-stored MRN may "fast track" a subsequent patient match?
- How does the IIS differentiate between MRN values that may be the same for different patients but assigned by different provider organizations?



#### Questions about IIS ID

- What role does the IIS ID (a.k.a. "SR ID," or in some jurisdictions, "LR ID") play in patient matching?
  - Are there scenarios when its presence in a QBP would not guarantee an exact match RSP?
    - I.e., might there be a possibility of a "death loop" of Z31 multi-match RSPs if the queried demographics don't sufficiently match the patient in the IIS even when the SR ID is present in the QBP?



#### Questions about Multiple Matches

- What does your IIS return when a QBP matches a single patient in the IIS but not with the highest confidence?
  - A Z32 / Z42 RSP containing the immunization history for that patient?
  - A Z31 RSP containing just one potential matching patient?
  - A Z33 "Too Many" RSP? A Z33 "Not Found" RSP?
- What is your jurisdiction's expectation for how this scenario will be handled in an EHR?



Oct 2018

## Patient Matching in HL7

A high level overview of patient matching in HL7

**Kevin Snow** Envision Technology Partners, Inc.

# Challenges

- Garbage in and garbage out
- SSN is available less and less, and some systems intentionally do not support it
- Lack of a national identifier
- Lack of transparency and consistency in how matching algorithms perform
- Hard to measure accuracy when even humans are prone to mistakes in matching
- Attributes and features for matching can change over time
- First Name: John, Jon, Jonathan, Johnathon, Johnahton, Johnny, Johnnie
- Last Name: Myron, Merrion, Maureen
- Date of Birth: 01/08/2018, 08/01/2018, 01/01/2018
- Gender: M, F, U
- Good luck!



The above slide by Adam Culbertson, M.S., M.S. Illustrates visually the patient match possibilities. <u>https://www.himssconference.org/sites/himssconference/files/pdf/54\_0.pdf</u>

#### Different approaches to matching

- Deterministic, Probabilistic, Machine Learning, Hybrid, etc. <u>https://www.cdc.gov/vaccines/programs/iis/interop-proj/downloads/de-duplication.pdf</u>
- Open Source Algorithms: FEBRL, FRIL, CHOICE MAKER
- Use of Master Patient/Client Index
- Next generation "MPI replacements" (Verato, etc.)

Question #1: For incoming VXU messages what sort of patient matching algorithm(s) do you use? For example, exact matches for certain demographic elements (e.g., first name, last name, DOB, gender)? A weighted matching algorithm considering a number of different demographic elements?

- Algorithm Type: Deterministic and probabilistic hybrid.
- Blocking strategy ("getting a group of people"): Nearly 10 blocking strategies are applied
- Probabilistic Strategy: 30 features are used
  - Using the terms features instead of fields because Patient Last Name could be
    - Exact Patient Last Name
    - Phonetic version of Patient Last Name
    - Version of Patient Last Name with-out special characters
    - Version of Patient Last Name with Hyphenated name switched
    - Etc.

Question #2: Does the answer to question # 1 differ for VXUs compared to QBPs?

- Yes, query and update apply slightly different blocking techniques and weight requirements. At a high level query can be loosened up to be used as more of a search; update can be tightened to be more of a match.
- Taking this one step further, query messages might be configured to search slightly differently based off of the calling facility

#### Possible Query configurations

• Below is to illustrate why you might want query match to work differently for different locations

Query Options	Query Options	Query Options	Query Options
<ul> <li>Disable Queries (VXQ/QBP)</li> <li>Maximum Number of Query Matches *</li> </ul>	Disable Queries (VXQ/QBP) Maximum Number of Query Matches * 5	Disable Queries (VXQ/QBP) Maximum Number of Query Matches * 1	Disable Queries (VXQ/QBP) Maximum Number of Query Matches * 1
<ul> <li>Use Stricter Patient Name Matching For Query Messages</li> <li>Require Matching Address For Query Messages</li> </ul>	<ul> <li>Use Stricter Patient Name Matching For Query Messages</li> <li>Require Matching Address For Query Messages</li> </ul>	<ul> <li>Use Stricter Patient Name Matching For Query Messages</li> <li>Require Matching Address For Query Messages</li> </ul>	<ul> <li>Use Stricter Patient Name Matching For Query Messages</li> <li>Require Matching Address For Query Messages</li> </ul>
Do not have permission to search.	Configure for an interactive UI where a user could select between multiple possible matches.	More appropriate for an automated interface. You can have one and only one match and the name has to match (fuzzy logic will not be applied)	Used in a public portal scenario where address matching is key to whether you can gain access to the records.

• In a dream world, there would be time to build out highly configurable search profiles

Question #3: What role does patient MRN play in patient matching?

Is there any sort of mechanism where a previously-stored MRN may "fast track" a subsequent patient match? How does the IIS differentiate between MRN values that may be the same for different patients but assigned by different provider organizations?

- MRN can increase the relativity of a match
- MRN by itself does not yield a positive match, it must be combined with things like name and/or DOB
  - Part of the reason for this is MRN could change due to things like deduplication inside the EHR
- MRN is, however, more consistently and accurately used so likely some more weight could be added

Question #4: What role does the IIS ID (a.k.a. "SR ID," or in some jurisdictions, "LR ID") play in patient matching? Are there scenarios when its presence in a QBP would not guarantee an exact match RSP? I.e., might there be a possibility of a "death loop" of Z31 multi-match RSPs if the queried demographics don't sufficiently match the patient in the IIS even when the SR ID is present in the QBP?

- The IIS ID is used more like a primary key but still requires a matching first name, last name or DOB to be considered a positive match in a VXU
- If the IIS ID matched an existing patient in the system but the name and DOB both did not match then you would not successfully match the patient

Question #5: What does your IIS return when a QBP matches a single patient in the IIS but not with the highest confidence? A Z32 / Z42 RSP containing the immunization history for that patient? A Z31 RSP containing just one potential matching patient? A Z33 "Too Many" RSP? A Z33 "Not Found" RSP? What is your jurisdiction's expectation for how this scenario will be handled in an EHR?

- Yields a single patient match with immunization history as long the minimum required relativity was met
  - If multiple possible matches are found, a TM (Too Many) would be returned and the query would need to be further restricted
  - See previous slide about being able to customize query settings and number of returned results
    - Strict: get one and only one match and only use strict name matching which would be more tailored to an automated interface
    - Loose: multiple results which would be more tailored to an interactive UI experience

#### What happens with 20 identical VXUs in a split second?

- When this occurs there is an increased risk of creating duplicates of the patient
- If you duplicate copies created at the beginning, there is increased risk of creating additional duplicates as you have multiple matches with the same relativity
- This is one reason why we discourage asynchronous sending, or at a minimum not asynchronously sending all the update messages for the same patient in a split second

Possible solutions	Downside
Everything goes into a manual review table.	Data ages before it's available.
Populate a queue that forces serial execution.	Under high volume this delays when the data will be available for query.
Locking of patient table to only allow one record to go in at once.	Slows down everything that is trying to access the patients table.
For a weighted patient matching approach, if multiple copies score the same relativity then pick earliest record.	
Perform a duplicate check at insert: if multiple patients have the same hash, you pick the earliest record for update and treat it as a single match.	It's possible incomplete patient records have the same hash so you'll likely want to include additional data in the hash which would mean this is more just protection against a duplicate.

## Next Meeting

October, 11th 2:00 pm ET / 11:00 am PT



#### More Information

- Web Links
  - Subscribe to immunization group <u>http://www.hl7.org/participate/UserGroups.cfm?UserGroup=Immunization</u>
  - Public User Group Wiki
     <u>http://www.hl7.org/special/committees/iug/index.cfm</u>
  - Private User Group Wiki
     <u>http://iugwiki.hl7.org/</u>
  - HL7 Press Release <u>http://www.hl7.org/documentcenter/public\_temp\_F760602A-1C23-BA17-</u> <u>0C0D326E635471F9/pressreleases/HL7\_PRESS\_20140402.pdf</u>
  - AIRA Press Release
     <u>http://www.immregistries.org/events/2014/04/10/hl7-immunization-user-group</u>



#### Contact Information

If you have any questions or comments:

- Kim Salisbury-Keith <u>Kim.SalisburyKeith@health.ri.gov</u>
- Nathan Bunker <u>nbunker@immregistries.org</u>
- Kevin Snow <u>ksnow@envisiontechnology.com</u>
- Danny Wise <u>Danny.Wise@allscripts.com</u>

Thank you!

