

HIT Policy Committee
Privacy & Security Tiger Team
Patient Linking Hearing
Department of Veterans Affairs – Veterans Health Administration
Office of Health Information – Data Quality submission
9 December 2010

1. What level of accuracy should be established for patient matching (i.e., matching patients to their data)?
2. What standards, if any, might need to be established to assist with patient matching?
3. Are there best practices that should be recommended to assist with patient matching?

THEMES/QUESTIONS FOR PANELS

Universal Questions for Testifiers:

1. What are your standards for identifying individuals?
Within the Veterans Health Administration (VHA) we use the OMG PIDS (Object Management Group Person Identification Service) standards for the exchange of patient identity/demographic information, and our unique identifier within VHA is based on the ASTM e1714 standard for universal health identifiers. We also utilize standards, where they exist, for data traits, such as the US Postal Service standards for any address traits used. Our base set of identification (ID) traits for identifying patients consists of full name (first, full middle name, and last), date of birth, full social security number (SSN), gender and additionally, if available we use mother's maiden name, place of birth, city and state, and phone and address components. For these traits, we have established standards. As far as formal standards on the matching process itself, such as what traits are used, matching thresholds, format of traits, etc., none currently exist that are used universally.
2. How can you be sure that you are accurately linking a patient with his/her data?
Before implementing our probabilistic matching algorithm, we did extensive analysis on our data, current matches, potential matches, and identified mismatch history. From this analysis, we have learned where the match and potential match thresholds should be set, along with what identity traits are required for matching. Presently, we use at least four identity traits for

matching, including full name, SSN, date of birth and gender. Slight permutations of the name data, such as nicknames, are permissible, but variations of other traits cause the potential match to be flagged for manual review. We also continue to monitor any reported issue, for opportunities to improve our business rules, matching thresholds, policies, and software applications to improve matching capabilities. In addition, we ensure that we provide enough identifying information to the users to be able to verify they have selected/retrieved the correct patient. We also see a significant degradation of matching with external partners when SSN is not provided. If a problem or potential problem in a patient record is discovered, we have a network of “super users” at each VA medical center that interface with the local staff and with our national Healthcare Identity Management (HC IdM) team to research and resolve any issues, utilizing a VHA-developed toolkit.

3. What problems are you having with patient-matching, internally and/or for information exchange? What is the source of those problems?
 - a. How do you handle patient matching problems (wrong/ambiguous match)? What should be done when there is no match although one was expected?

Within VHA, we have a team of data quality analysts working at the national level to identify and resolve any identity management issues and exceptions. They work with local contacts and technical teams to address issues, develop policies, and identify any software modifications that may be needed. The local contacts work with the national staff to manually resolve any issues identified to ensure correct matches are made and that the longitudinal health record is accurate.

For our experiences with data exchanges as broad as the Nationwide Health Information Network, we have done as much work on the “front end” as possible to minimize the occurrence of non-matches within the pilot phases. Defining a robust set of traits that are required for matching also helps to minimize non-matches and ensures that there is enough information to substantiate a match with the correct patient, i.e. strengthening the existing required identity traits by stipulating the provision of full name including middle name, requiring SSN, and including other non-changing traits such as mother’s maiden name and place of birth city and state. These are key traits necessary for authoritative matching in this effort.

Internally when a match is not made, although one is expected, policies and resources have been established to research and resolve any data

quality issues that may have caused the non-match. Having identified staff that are able to access information from each of the sharing partners and ensuring data quality errors are corrected, if possible, so that matches can be made. This piece has not been well-defined within the Nationwide Health Information Network and that aspect of the specification should be addressed specifically to outline expectations and resources, if needed. Privacy Laws from a health record perspective related to identity traits are also relevant here and must be followed. Internally within VA, the local and national identity management staff work together to resolve high threshold potential matches. We are unsure that this model is extensible given the level of effort, expense, Privacy Laws and concerns over security. We believe the specifications should be made strong enough to support matches and not require a significant level of adjudication on the part of the provider participant or impact sustainability.

- b. In your experience, what are the consequences of a wrong match? – to patient safety, privacy

Mismatches or wrong matches are very costly in terms of patient safety implications (not having complete information, incorrect information, etc.), resources necessary to correct them, and frustration for the patient. VA calls these wrong matches “catastrophic edits” if they are caused by changes being made to a record which alter the identity of the record or if an incorrect record is selected and information is entered for the wrong patient. These errors are given the highest priority for immediate action by the national identity management group, since the negative implications to patient care are paramount.

In data exchanges with external patients, the opportunity for wrong matches should be minimized by stringent matching thresholds and by not allowing the selection of a patient record from a list of potential matches. If a definitive match is not made with the traits that are provided, then no opportunity for selection from a list of potential patients should be allowed, to minimize potential for mismatches. We agree this has to be system-to-system based matching only.

4. What level of accuracy do you establish for patient matching?

Within VA, our threshold has been set at a level to ensure that erroneous matches are not made, if at all possible, to minimize patient safety risk. This, however, may result in a higher number of “potential matches” or false negatives that need to be manually reviewed. In general, we identify

an average of four catastrophic edits to patient records per month, out of more than 100,000 edits to identity traits for patients.

5. What lessons learned do you have from solving this problem?
 - a. *Efforts to gather quality data at the point of entry are essential and worthwhile; we continue to refine our data collection to improve data quality in patient matches.*
 - b. *Thresholds for defining matches, potential matches, and non-matches should be carefully established and continuously reviewed to ensure that they are effective.*
 - c. *Deterministic matching and a limited set of traits for matching submitted by any organization will continue to keep non-matches higher than expected.*
 - d. *Means to detect erroneous/catastrophic edits must be defined and structures put into place to support such efforts.*
 - e. *An effective infrastructure of tools and staff must be in place to identify and reconcile data quality issues within each organization.*

6. What are the cost implications of various solutions?
 - a. *Implications of patient safety errors are costly for patients and organizations and must be avoided whenever possible, which dictates a higher threshold for matching*
 - b. *Additional staff for manual review of near misses for analysis and research purposes could provide additional insights, but will be limited by the effects of deterministic matching and the limited set of identity traits provided by an organization*

7. What should ONC do to address patient matching problems in information exchange?
 - a. *Establish standards for identity traits and matching criteria and thresholds*
 - b. *Consider offering a “voluntary” universal health identifier, to reduce the need for repeated trait-based searches and matching*

Questions for Panel 2

- What are the solutions?

There are several actions/solutions that can and should be implemented to improve the accuracy and successful matching of patient

records within health care. The first is the use of a standardized, robust set of identity traits to be used for consistent matching, which must be mandated for all sharing partners. The use of full name (including middle name), full SSN, gender and full date of birth are essential in ensuring that successful, accurate matches are made. Additional traits that should be used to augment the minimum set could include Mother's Maiden Name, place of birth, city and state, phone number and address components. Using existing standards for those traits for which they exist also must be mandated, which relates to address traits, SSN, date of birth and place of birth fields.

The use of a sophisticated, probabilistic matching algorithm that is capable of utilizing this robust set of traits and adjusting to missing or incomplete traits, is also recommended. Careful consideration of match, potential match, and non-match thresholds will ensure that false positives, false negatives, and mismatches do not lead to data quality issues, patient safety concerns, or issues with a breach of data.

In addition, we recommend that each organization consider establishing a group of data quality analysts charged with maintaining the integrity of the patient identity data, which includes establishing and propagating policies, best practices, and directives for the entry and updating of identity traits is essential.

- What is the status of those solutions for healthcare?

In our experience within VA, some of these pieces are well-developed, while others are still under development. Standards exist within organizations for most identity traits, but across sharing partners many variations exist. The development and adoption of recognized standards for health care identity through the existing standards such as the ASTM E1714 Universal Healthcare Identifier, adoption of the HL7 Version 2.5 or higher for the contents, and format of the demographic traits should be required.

- What are the gaps?
 - a. Standards across organizations related to standardized traits for matching*
 - b. Variations in the algorithms used across organizations*
 - c. Addressing gaps in a and b above are key in being able to extend this to the number of provider and patient participants. Without addressing those gaps, especially within the specification for matching, matching with the Nationwide Health Information*

Network will occur at a very low rate which will affect participation.