Managing the Integrity of Patient Identity in Health Information Exchange (2009)

Save to myBoK

Over the past decade, multiple studies have documented the value of health information exchange (HIE). eHealth Initiative's recent "Fifth Annual Survey of Health Information Exchange at the State and Local Levels" found that 69 percent of fully operational exchange efforts reported reductions in healthcare costs.[1]

Respondents indicated that HIE decreased dollars spent on redundant tests; reduced the number of patient admissions to hospitals for medication errors, allergies, or interactions; decreased the cost of care for chronically ill patients; or reduced staff time spent on administration. These results support recent findings from a RAND report stating that the potential benefits of a connected, interoperable healthcare system could save an estimated $80 billion per year.[2]

However, in order to successfully exchange health information and reap the benefits of HIE, organizations must maintain accurate patient identification information. Patient identification integrity is a complex concept, and one that is not well understood throughout the healthcare industry. Many policy makers and industry leaders do not fully comprehend the negative effects of inaccurate patient identification information for even basic health information interchange.

Seemingly mundane errors such as transpositions in a patient's birth year, misspellings or culturally acceptable spellings of a patient's last name, and nicknames or default Social Security numbers (SSNs) can play havoc with successfully linking electronic records across clinical and administrative systems.

Even though most provider settings use a medical record number (MRN) as a unique identifier to connect records across and within their electronic systems, many of the interfaces across these multiple systems (commonly 30 to 50 within a single hospital) use some patient demographic data to validate the interfaced transaction.

Organizations that exchange health information and do not share a common unique identifier are completely dependent upon the accuracy and completeness of the key demographic data available in both records for successful electronic linking.

Healthcare standards for patient identity integrity have been slow to emerge. Existing standards address data format and position within an electronic transaction; however, data content accuracy has yet to be addressed.

Historically, little emphasis has been placed on the role patient identification systems play in the quality and safety of healthcare delivery. Local patient identification errors have been contained and managed within the healthcare organizations that create them.

Today's health reform initiatives push this backroom function to the very front of the health information exchange effort. Sending the wrong health information to the point of care can create critical patient care issues and risk privacy breaches, degrading consumer trust. The negative effects of local patient identification errors will expand as technology advances and the national health information network continues to expand.

To this end, this practice brief outlines how organizations can manage patient identification systems from front-end data capture to back-end quality control as an ongoing process and carry local quality operations into health information exchange efforts. It urges industry stakeholders to recognize that now is a critical time to address accuracy in patient identification systems.

Unless the healthcare industry takes the necessary measures to ensure complete and accurate data at the provider and HIE levels, the national strategic efforts under way to improve quality and safety will be more difficult to accomplish.

At the time of publication, the Markle Foundation proposed that meaningful use of health IT—a requirement for receiving health IT incentives under the American Recovery and Reinvestment Act—focus on "improving medication management and

the coordination of care."[3] Among the seven principles Markle outlined were factors that can only be accomplished through unimpaired and accurate identification of a patient or individual's healthcare information.

Support for better decision making, effective care processes that improve health outcomes and reduce cost outgrowth, and consumer benefit from health IT through improved access to patient health information are made possible when unbroken and dependable patient identification systems are operating.

Patient Identification Processes, Procedures

Ongoing, focused management and oversight of healthcare patient identification is critical to both internal operations and regional and national HIE efforts. Well-integrated front-end and back-end workflow processes offer optimal control of this function, which is set in motion with both a patient's initial and recurrent presentation to a healthcare provider for services.

Cradle-to-grave lifetime records that are accurate, complete, authenticated, and accessible by authorized providers can only occur if these processes are thorough and timely. As with the genesis of revenue cycle managers in healthcare organizations, data integrity managers of the enterprise's master patient index (MPI)—who oversee with authority the quality of patient identity data for all electronic and paper patient (and person) records regardless of the data source—are critical to achieving true health record interoperability.

Organizations must develop policies that ensure key demographic data are accurate and used to link records within and across electronic health record systems. These policies must address the accuracy of information at the initial point of capture using front-end verification, including timely correction of duplicate records and quality monitoring using a duplicate creation rate.

However, care should be taken that the emphasis on limiting the creation of duplicate MRNs does not lead to the greater evil of registering a patient to someone else's MRN and merging the identities of two separate individual's records. Record-linking algorithm effectiveness should be validated prior to linking records within an organization or releasing records to an HIE.

Organizations should outline duplicate record validity procedures and ensure they are followed. They must provide staff training at all levels to reinforce the importance of successful health information exchange. A well-managed patient identity integrity program will include an ongoing performance improvement process that assesses error rates and ensures progressive improvement. HIM professionals are well positioned to lead this effort.

Technology Efficiencies

Health data exchange increasingly occurs across disparate provider organizations using networks and Internet-based technologies. Interoperability requires the electronic transmission of data across organizations and assumes the data exchanged are accurate, comprehensive, current, consistent, relevant, timely, granular, precise, accessible, and well-defined.[4]

However, electronically exchanged data rarely meet the standard for each of these quality attributes, compromising healthcare's ability to achieve truly successful interoperability. Errors in the migration of data from one system to another, aging data, and lack of data completeness in each provider's patient record severely limits optimal interoperability.

HIEs currently use a variety of data delivery methods, which determine how patient records are sought and matched. Data from one provider may be "pushed" to another; for example, all electronic prescriptions generated within a hospital are automatically sent to a specific pharmacy, or all transcribed documents are forwarded to the physicians listed in the Health Level Seven header. In these scenarios, the receiving provider knows the patients about which it is receiving messages and uses internal procedures to process the incoming transaction. "Pull" technology is used in scenarios where a record locator service (RLS) provides a centralized index of all patients collected from HIE participants and a cross-reference to each participating organization's MRN is stored in the RLS. A provider searches the RLS and finds and selects the patient for whom it is searching.

Electronic messages are then sent to each of the participating organizations that have stored records pertaining to the patient. Specific types of electronic clinical results are pulled back from each participating organization to the requesting provider, such as lab tests, text-based reports, medication histories, and problem lists.

A centralized data model uses both push and pull technologies. Each organization participating in the HIE sends (pushes) clinical results to a central database managed by the HIE. Providers search the MPI of this centralized database and pull the

corresponding information for the applicable patient across to their system.

Record Matching

A fundamental and critical success factor for the RLS and centralized models is how the indexes within these databases link records for the same patient from the disparate participating organizations.

Ideally, the RLS or MPI identifies an individual with records at multiple locations using a unique identifier for the HIE. Many healthcare provider organizations have multitudes of unique identifiers for a patient (e.g., MRN, account numbers for billing, order numbers, and requisition numbers). Further, an organization's own set of unique patient numbers will not be the same for one patient across multiple provider organizations.

Therefore, once an RLS obtains demographic data from its participating organizations, it must collapse all of these individual demographic records for one patient into a single record. In this model, the HIE links the different provider records for one patient into one record and assigns that patient a unique numeric identifier. This unique identifier is sent back to each participating organization that holds medical records for that patient. Subsequent updates to the HIE by that participating organization use this unique identifier.

For example, John Smith has been seen at a physician office, hospital, and lab. His HIE unique identifier, 15894876, is included in the identifying data set from each provider. His records from all providers can be retrieved by entering this unique identifier in the search screens. Even when the search is via the name or other key identifiers, when the requestor selects the correct "John Smith," records from all providers will be included as the records are linked via his HIE unique identifier.

Records from various providers are frequently linked by algorithms embedded in the HIE's software. The algorithms available to perform this linking function fall into three main groups: basic, intermediate, and advanced.[5]

Basic Algorithms

Basic algorithms are the simplest technique for matching records. Comparisons are made on selected data elements—usually the name, birth date, SSN, and sometimes the gender.

Exact match and deterministic algorithms are both basic matching tools. With exact matching, the data elements used to search must match exactly with those in the database in order to return a particular record. Deterministic matching is slightly more sophisticated in that in addition to exact matches, partial matches or matches on Soundex codes (or those from other phonetic encoding systems) may be used to return a record.

With an exact match algorithm, if the name "Smith" is entered in a search for a patient with the last name of Smithe, the Smithe record would not be located, as it does not match exactly. However, a deterministic match using a substring of the first three letters (partial name) of the last name would be returned.

If the algorithm has deterministic capability and uses Soundex codes, a person with the last name of "Johansen" would be returned for a search on "Johnson" because the Soundex codes for these names are an exact match.

"Wild-card" linking falls under the basic algorithm definition. With wild-card linking, the user enters a few letters of the value being searched and adds a character (frequently a common keyboard symbol) that instructs the program to return every record that matches the limited letters entered.

A wild-card search on "Smi*" will return both Smithe and Smith, in addition to any other name in the database beginning with the first three letters and that match any other data element entered to refine the search, such as date of birth or gender.

Intermediate Algorithms

Intermediate algorithms use more advanced techniques to compare records. Fuzzy logic and arbitrary or subjective scoring systems are added to exact match and deterministic tools. A field match weight is arbitrarily assigned to key patient identifying attributes, such as last name, first name, date of birth, and SSN.

For example, a match on the SSN may be assigned a score of 40 points, while a match of the last name scores 25. Any records presented to the searcher must reach a minimum scoring threshold to qualify for inclusion.

Fuzzy logic and rules-based algorithms also may be a component of intermediate algorithms. These tools include nickname tables, rules to address transposition of characters or names, digit rotations, and typographical errors within the MPI database.

Intermediate algorithms may include an automated frequency adjustment, which decreases the field match score. This adjustment will decrease the score assigned to a field match across two records if the actual field value (such as a common last name or a common date of birth like 01/01/2001) is computed to be present in a high volume of records in that data set.

In this instance, a search for Elizabeth Jones would return records for Betty Jones as well as Elizabeth Jones. A search for Richard David would return records for both Richard David and David Richard. A search for James Smith with a date of birth of 6/17/1978 would return Jim Smith with a date of birth of 6/17/1987 as a possible match.

Advanced Algorithms

Advanced algorithms contain the most sophisticated set of tools for matching records and rely on mathematical theory. The core intelligence within advanced algorithms includes bipartite graph theory, probabilistic theory, and mathematical and statistical models, which are applied to determine the likelihood of a match on specified data elements.

Probabilistic matching uses the frequency of a specific element with a probability score assigned to adjust the relative value of the match or mismatch for the specified elements. The weight assigned to each field is relative to the weights assigned to other fields, but only after thorough research across millions of records (as opposed to a simple frequency analysis with an arbitrary field weight adjustment).

Advanced algorithms also include machine learning and neural networks, which use forms of artificial intelligence that simulate human problem solving. These systems actually "learn" as the program processes more data and will automatically tune the field weights as required based upon the learning achieved during the processing of the entire data set of records.

For example, a search for:

Erin Marie Kotnica, date of birth 8/18/1965, female, SSN 555-45-8888, 110 E 3rd Street

would return a record with the values:

Aaron Marie Skotnica, date of birth 8/13/1956, female, SSN 555-44-8888, 110 E 3rd Street

Despite a variance in multiple key identifying data elements, the advanced algorithm using the technologies listed above could identify the appropriate record.

Whatever algorithm an organization uses to link records, the results should be verified by staff using record-matching validity procedures. When applying designated HIE system requirements, a percentage of records from different participating organizations will be able to be automatically linked if a sufficiently sophisticated algorithm is used. However, a statistically significant sample should always be reviewed to ensure only true overlap records are autolinked (see "MPI Definitions" at right).

Even with a sophisticated algorithm, the HIE will achieve significantly higher rates of record links if potential overlap records that have a record match weight lower than the autolink threshold are reviewed and manually linked. There will always be potential intrafacility duplicate pairs that must be sent back to that participating organization for staff to review, validate, and manually combine.

False positives and false negatives will always occur with any algorithmic or manual system identifying potential duplicates. A false negative will result when the algorithm or other duplicate identifying process does not identify a true duplicate and the duplicate remains in the database. False positives occur when two records are matched together because they are presumed to belong to one person, when in fact they belong to different people. They are easier to find if a review is completed of each potential duplicate identified by the system.

Common pitfalls include linking two closely related people with very similar names and dates of birth who live near each other (e.g., cousins who are named after the same individual who recently expired); two individuals living in a dense urban area with the same common name, date of birth, and address; or the example of twins having the same first name. Failure to catch such errors can result in overlaid medical records and subsequently negative health outcomes, serious privacy breaches, and legal ramifications.

MPI Definitions

HIEs and organizations going through the process of linking their records and cleaning up their MPIs may need to brush up on common MPI terms.

Duplicate: more than one entry or file for the same person in a single facility-level MPI. This causes one patient to have two different medical records within the same facility.

Overlap: more than one MPI entry or file for the same person in two or more facilities within an enterprise. For example, patient John Smith has medical record number 12345 at facility A and a medical record number 447788 at facility B within the same enterprise-wide system. When both MPI databases are loaded into an enterprise MPI, the database does not link the two records. Thus, Smith ends up with two different enterprise identifiers and providers cannot view all clinical information across the enterprise for that patient.

Overlay: one MPI entry or file for more than one person (i.e., two people erroneously sharing the same identifier). Overlaid records are frequently caused when patient access staff select another patient's record during a scheduling or registration event. Sometimes interfaces cause the error if the receiving system lacks a robust patient record-matching program and "overlays" another patient's record from that inbound interface transaction. On occasion, overlays are caused by an incorrect merge of two records that belong to two different people. [1]

Note

1. AHIMA MPI Task Force. "Building an Enterprise Master Person Index." Journal of AHIMA 75, no. 1 (Jan. 2004): 56A–D.

How to Measure Duplicate Record Rates

Participating organizations such as hospitals and other healthcare delivery systems within the HIE are responsible for maintaining the integrity of the patient-identifying data within their own systems. Organizations that fail to carry out this responsibility not only compromise care within their own four walls, but also contaminate the HIE database and cause administrative complications or compromise care at other participating organizations.

Different methods can be found within the healthcare industry to measure the duplicate rate at a given point in time or measure the ongoing duplicate-creation rate. Algorithms used to identify potential duplicate records are also widely different. When choosing algorithm software or vendor consultant services, organizations are advised to investigate and understand proposed measurement techniques and ensure a consistent approach is used for subsequent comparative performance measurements.

Below, a basic, best-practice, standardized formula is described to ensure sound unit counting when measuring duplicate rates in any healthcare organization.

Facility Duplicate Rate for Static Database

Healthcare organizations sometimes choose to analyze their entire MPI database for potential duplicate records at a given point in time (i.e., a "static" database). The organization extracts the MPI data and analyzes that static group of records.

A computation determines the percent of records that are potential duplicates at the time the data were extracted (within that one database at that one facility). After the duplicates have been evaluated, those that truly represent the same individual qualify to be included in the calculation.

The following formula is a standard industry method of computing the actual duplicate record rate in a single database:

Total no. of individual duplicate
   patient records x 100

$$\frac{\text{Total no. of patient records in the}}{\text{MPI database}}$$

The total number of individual duplicate records is the count of the "extra" or duplicate patient records. Therefore, if 50 patients each had two records, the number of duplicate records would be 50 (representing each of the "extra" or duplicate records). If 90 patients had two records and 10 patients had three records, this number would be 90 + (10*2) = 110 because 10 of these patients had 2 extra or duplicate records.

For example, a facility has 10,000 duplicate pairs in the database, involving 20,000 individual records. The database at the time of the analysis contained 500,000 individual records. The duplicate rate is computed by dividing 10,000 by 500,000 and multiplying the result by 100 to obtain the percent result. In this example the rate is 2 percent:

$$\frac{\text{10,000 duplicate patient records x 100}}{\text{500,000 total patient records in database}} = 2\% \text{ duplicate rate}$$

Facility Duplicate Creation Rate

Measuring the ongoing duplicate creation rates involves dividing the number of records involved in the duplicates (the numerator) by the number of registrations performed (the denominator) within the same time period. The total number of registrations should include any opportunity that users have to create a new record when performing scheduling, preregistration, or registration activities. If the scheduling system creates a permanent person (or patient) record within the MPI database when scheduling an appointment, this represents an opportunity to create a duplicate record and should therefore be included.

Accordingly, the definition of the denominator may vary from organization to organization and is dependent upon the configuration and functionality of their applications. The numerator would always be the number of "new" or "extra" (duplicate) patient records created by any means during that time period. If an inbound ADT or scheduling transaction creates a new patient record, and that new record creates a duplicate patient record in the database, it should be counted in the numerator.

Presuming that all these activities present the opportunity to create a new record in the database, the formula for determining the duplicate creation rate is:

$$\frac{\text{Total no. of individual duplicate patient records for a given time period x 100}}{\text{Total no. of registrations, preregistration, or scheduling events for the same time period}}$$

Depending on the registration or scheduling system used, the denominator might be computed by counting the number of new account numbers generated during this time period and adding to that the number of temporary and scheduled visits created during that time period that do not yet have a permanent account number.

For instance, a facility has 10,000 inpatient and outpatient preregistration and admissions per month. (The scheduling system is independent of the MPI database and does not create records.) In one month, 150 duplicate records were created. The duplicate creation rate is computed by dividing 150 by 10,000 and multiplying that result by 100, arriving at a creation rate of 1.5 percent. Even a low rate creates too many duplicate records, compromising patient safety.

To understand what this duplicate creation rate might look like for organizations of different sizes, consider the following examples:

- 10,000 visits/month x 0.5 percent = 50 duplicates created each month (an average of 1.7 per calendar day)

- 50,000 visits/month x 0.5 percent = 250 duplicates created each month (an average of 8/calendar day)
- 100,000 visits/month x 1.0 percent = 1,000 duplicates created each month (an average of 33 duplicates created each calendar day)

As illustrated, even a very low creation rate places an organization at risk for inadvertent patient care errors. It is critical that every organization maintains solid data quality practices to identify and correct duplicate records on a timely basis and ensure proper feedback to the departments and individuals creating such errors.

Overlap (Across-facility Duplicate) Static Rate

When electronic medical records are shared across a multifacility organization, each patient record should be connected at the enterprise level. An "overlap" comprises two patient records from two different facilities that use different MRN "pools" of numbers. The patient may have only one medical record number from each facility, but when aggregated into an enterprise database, the two MRN records from the two different facilities do not link. This represents an "overlap" or an "enterprise duplicate." These overlaps may be measured with a simple calculation:

$$\frac{\text{Total no. of individual overlap enterprise patient records x 100}}{\text{Total no. of unique patient records across two or more MPI databases (i.e., facilities)}}$$

For an explanation of counting the total number of individual overlap records, see the description under the formula in the section "Facility Duplicate Rate for a Static Database."

Accuracy and Content Issues

An HIE must establish common procedures to ensure that records are linked accurately. All system users must understand how and when to validate their selection of the correct patient record and how and when to determine that their search does not include the patient record sought.

Some HIEs have opted to keep their search criteria strict to minimize false positives in the absence of the correct record in the patient list. Such an error would ultimately lead to inadvertent disclosure of the wrong patient records during a search that does not identify the correct patient due to false positives.

Sometimes, however, this strategy creates a situation where an existing record does not appear in the record result list. In this situation, the record-matching search algorithm of the index must be robust enough to allow entry of several patient demographic data fields and be error tolerant, allowing for mistakes in the entry of any one of these data points, yet still be able to find the appropriate record.

Validation Methods

It is critical that all HIEs routinely monitor linked or merged records within their systems and regularly work the lists of potential overlap (enterprise duplicate) records identified by their record-matching algorithms. Without such data maintenance, large databases will become riddled with data integrity problems. Providers using the system will start seeing duplicate and overlap records, lose confidence in the system, and stop using it.

Front-end verification . The best opportunity to ensure accurate patient identity information is at the front-end data capture point, usually performed by a healthcare delivery scheduling or registration staff member.

A user searching for a record must validate that the record presented represents the patient in question. This is generally achieved by reviewing key identifiers such as the patient's last, first, and middle names; date of birth; gender; and SSN, if available. Address and telephone numbers are also frequently used to ensure correct choice.

This selection process implies that the user has access to this information directly from the patient or his or her representative so that it may be verified prior to selection of a particular record.

Back-end verification . HIEs must ensure that the records organizations send through the network are linked appropriately. The HIE's software algorithms should be examined to ensure they work correctly, minimizing or eliminating records that match inappropriately (false positives) as well as records that should match but fail to do so (false negatives).

HIM professionals who are acquainted with the record match validation methods should perform a manual review of the proposed linked records, with ongoing review of new matches after implementation. Validity procedures provide staff with guidance on how to consistently decide whether two different records belong to the same patient.

Review of potential duplicate record sets and autolinked records can help set thresholds for record match weights. This review will help determine at what point records represent the same individual consistently and when manual review should be undertaken prior to linking.

While false positives are usually relatively easy to identify (although common names and multiple birth records frequently pose challenges), identifying a false negative (two records that fail to match) is more challenging. Research-based studies of record-matching algorithms have not been completed.

One way to determine if records have linked appropriately is to review specific examples identified at multiple facilities that can be verified as linked in the actual false positive or false negative rates of various algorithms.

Methods to Improve Patient Record Validation

While it is the duty of every healthcare organization to protect the privacy of patient information within its control, this identifying information is vital in establishing appropriate links within an HIE. Organizations must balance the amount of information sent to an HIE to ensure that privacy requirements and patient care needs are addressed.

One method organizations may employ to improve the validity of links is to ensure that adequate identifying information accompanies the records sent to the HIE. Key identifiers are essential, including last name, first name, middle initial or name, date of birth, gender, address, and telephone number.

Historically healthcare providers have collected SSNs and stored them in MPI databases to improve patient identification. However, with the increasing problem of identity theft, many individuals and organizations are reluctant to use this identifier for fear of compromising an individual's identity.

Some HIEs have addressed this concern by including only the last four digits of the SSN in the identifiers accompanying a patient's record. Including these values significantly increases the accuracy of record matching. Other identifiers also may be provided that are used only in the background and are not visible to the user searching for a record. For example, guarantor, next-of-kin, and insurance policy information is helpful in determining what two records should be linked between disparate systems that comprise an HIE. These identifiers can be used in the background by the algorithm and not displayed to the individual searcher.

A second method organizations can use to improve validation of patient identification stored by the HIE is to ensure that appropriate and efficient policies and procedures are in place to address any patient identification errors created within local facility MPIs and that the processes used by staff who create records have been refined to decrease the creation of new duplicates. Addressing standard naming conventions and search methods and ensuring appropriate training helps achieve minimal duplicate creation rates.

When patient identification errors are identified internally, it is helpful to study patterns and trends to assist in root-cause determination and subsequent corrective activities.

Organizational Training

HIM professionals have the expertise and ability to manage MPI database integrity and play a critical role in directing enterprise-wide activities that affect the MPI and the subsequent exchange of health information. HIM professionals in MPI leadership roles provide data and feedback necessary to improve registration accuracy, contribute to software selection, and maintain accurate data transference among downstream databases.

Performance data and duplicate creation rates may be effectively shared with key stakeholders via performance improvement functions within the organization, bringing support for development of new procedures, training of registrars, MPI staff, and

others who work with the demographic database. Providing aggregate information through a performance improvement mechanism heightens overall awareness of MPI value to an enterprise and HIE.

Preventing MPI duplicates directly affects the ability of the patient registrar to correctly locate a returning patient in the demographic database and to accurately update the information therein. In the case of a new patient, the registrar initially must enter all information accurately. Comprehensive training in record search routines coupled with data quality mechanisms that provide feedback to registration and scheduling staff are crucial.

Aggregate data review and individual case review can provide feedback on registration accuracy performance. Performance data, such as facility duplicate rate and duplication creation rate, can be used to focus educational efforts on problems that occur most frequently. Results data also may be used to make software enhancements that boost accuracy through the elimination of registrar decisions or help screens that assist registrars in selecting appropriate choices.

Comprehensive training requires that written policies and procedures for patient search routines are readily available to patient registration staff. Newly hired registrars should be trained using test registration environments, with access to live environments only upon satisfactory demonstration of knowledge. Successful training includes a review of the search methodology specific to the enterprise's MPI software. Software with less sophistication may require a "less is best" philosophy, where the preferred search method may be by SSN alone or last name and first name when the SSN is unavailable.

In the most basic systems, searches should be conducted using only partial last name and partial first name due to a lack of error tolerance in that product's search routines. As product sophistication increases, the patient is most reliably located when more information is entered (e.g., full name, SSN, date of birth, etc.). Some search engines will look for all persons who have received or might receive services, which would include a previous patient, subscriber, or guarantor.

Other systems can be set with a specific search mode, such as "starts with" as the default and the patient can be found based on the first part of the last name. These systems may be changed to a phonetic search that works based on the pronunciation of the last name, or "full," which will search with all the letters entered.

Inclusion of standardized patient-naming conventions in training programs is an essential key to prevent duplicate creation by registration and scheduling staff. Policies should delineate any acceptable prefixes and their values. Appropriate use of spaces, hyphens, and apostrophes must be established. Provision for use or exclusion of titles such as *Rev., Dr.,* and the proper designation of prefixes or suffixes such as *Sr., Jr., II,* or *III* should be included in documentation guidance. Strong emphasis must be placed on the use of a patient's legal name, exclusion of nicknames, and proper provision for middle name (or, at minimum, middle initial).

Inevitably, even processes intended to ensure the correct identification of patients and prevent creation of duplicate records by registrars can allow the correct patient to go undetected. In those cases, reports or e-mail alerts may be generated for MPI staff providing demographic information of each record in the potential duplicate set.

HIM or MPI staff members should review the potential duplicate sets, determine if they are the same patient, and perform merges in financial and ancillary systems as appropriate. Proper training in understanding causes of overlaid records (see "MPI Definitions" on page 65) is important to avoid merges of records belonging to two different patients. The strictness of validity procedures to determine if two or more patients are the same person varies with the degree of probability that an organization is willing to accept. A reputable firm can be contracted to assist organizations in establishing those policies for patient consolidation.

Organizations vary in the information they use to validate if two records belong to the same person. Guarantor information, insurance information, medical history, clinical results, previous names, addresses, and telephone numbers are useful resources; however, policy should dictate which resources MPI staff use. Data integrity specialists should be given access to thoroughly investigate potential duplicates. Validity decisions should only be made when there are enough data to determine validity with confidence. Guidance from external companies specializing in MPI cleanup can assist in the development of these validity procedures.

Identifying patterns of duplication errors is important, as patterns provide the basis for ongoing education and corrective action. Information that can be tracked can include the type of error, such as duplicate, overlay of two different patients, and overlaps between facilities. These can be calculated for each facility, as well as enterprise-wide. Examining these cases in detail also will provide the identity of the registrar who created the duplicate and allow for appropriate feedback, as needed.

A facility also may find it helpful to drill down further and track errors with greater specificity (e.g., correct patient not identified, SSN transposition, or naming convention not followed). Appropriate education, policy clarifications, communication, and biometric devices may be employed to decrease duplicate creation errors, resulting in a more accurate database and patient identity integrity.

The criticality of patient identity integrity has graduated from historical internal impact to one of national importance as successful HIE initiatives and interoperability models progress. The healthcare industry must commit to the improvement and maintenance of content accuracy in patient identification databases from which information exchange effectiveness springs.

Patient identity integrity aligns neatly with the messages of multiple management gurus: you can't manage what you can't measure. This wisdom is central to healthcare reform, quality of care improvement, administrative efficiency, and reduced healthcare costs. A single error rate snapshot is only a start in the plan-do-study-act cycle that, when done well, can make a critical difference to the quality of care a patient receives.

Striving for six sigma quality (99.999999 percent accuracy) is as important in healthcare as it is in landing an airplane. Setting this goal for patient identity accuracy will be a significant step in achieving high healthcare quality standards throughtout the entire healthcare delivery system.

Notes

1. eHealth Initiative. "Fifth Annual Survey of Health Information Exchange at the State and Local Levels." September 2008. Available online at www.ehealthinitiative.org/assets/Documents/eHI-HIESurveyResultsFinalReport-2008.pdf [link no longer active].
2. Hillestad, Richard, et al. "Identity Crisis: An Examination of the Costs and Benefits of a Unique Patient Identifier for the U.S. Health Care System." 2008. RAND Corporation.
3. Markle Foundation. "Achieving the Health IT Objectives of the ARRA: A Framework for 'Meaningful Use' and 'Certified or Qualified EHR.'" April 2009. Available online at www.markle.org/downloadable_assets/20090430_meaningful_use.pdf.
4. Just, Beth, et al. "HIM Principles in Health Information Exchange." *Journal of AHIMA* 78, no. 8 (Sept. 2007): 69–74.
5. E-HIM Work Group on Patient Identification in RHIOs. "Surveying the RHIO Landscape." *Journal of AHIMA* 77, no. 1 (Jan. 2006): 64A–D.

References

AHIMA Data Quality Management Task Force. "Data Quality Management Model." June 1998. Available online in the FORE Library: HIM Body of Knowledge at www.ahima.org.

AHIMA e-HIM Work Group on EHR Data Content. "Data Standard Time: Data Content Standardization and the HIM Role." *Journal of AHIMA* 77, no. 2 (Feb. 2006): 26–32.

AHIMA e-HIM Work Group on Regional Health Information Organizations (RHIOs). "Using the SSN as a Patient Identifier." *Journal of AHIMA* 77, no. 3 (Mar. 2006): 56A–D.

AHIMA MPI Task Force. "Maintenance of a Master Patient (Person) Index—Single Site or Enterprise." October 1997. Available online in the FORE Library: HIM Body of Knowledge at www.ahima.org.

AHIMA MPI Task Force. "Merging Master Patient Indexes." September 1997. Available online in the FORE Library: HIM Body of Knowledge at www.ahima.org.

AHIMA MPI Task Force. "Master Patient (Person) Index: Recommended Core Data Elements." July/August 1997. Available online in the FORE Library: HIM Body of Knowledge at www.ahima.org.

Altendorf, Robin L. "Establishment of a Quality Program for the Master Patient Index." AHIMA's 79th National Convention and Exhibit Proceedings, October 2007.

Prepared by

2009 HIE Practice Council

Beth Haenke Just, MBA, RHIA

Diane P. Fabian, MBA, RHIA

Lenore L. Webb, RHIA

Beth M. Hjort, RHIA, CHPS

Acknowledgments

The information contained in this practice brief reflects the consensus opinion of the professionals who developed it. It has not been validated through scientific research.

Article citation:
AHIMA. "Managing the Integrity of Patient Identity in Health Information Exchange (2009)"
Journal of AHIMA 80, no.7 (July 2009): 62-69.

Driving the Power of Knowledge