Fundamentals for Building a Master Patient Index/Enterprise Master Patient Index (Updated)

This practice brief is made available for historical purposes only.

Editor's note: This update replaces the January 2004 practice brief "Building an Enterprise Master Person Index."

A master patient index (MPI) is an index of known patients within a single organization whose visits are linked together by a single identifier, typically the medical record number. MPI management activities typically pertain to a software application that identifies, coordinates, and lists database information.

Electronic MPIs have been used in healthcare since the 1980s, replacing the past process of manually completing index cards. Interest in electronic MPIs has markedly increased because of the consolidation of healthcare organizations and the implementation of electronic health records (EHRs). Industry efforts such as health information exchange (HIE) and the Nationwide Health Information Network (NHIN) also have provided a spotlight on a new term, *enterprise MPI (EMPI),* which combines MPIs of two or more organizations.

As the healthcare industry moves to a fully integrated, longitudinal EHR, the MPI and EMPI become vital databases. The information within an MPI is a key component in the accuracy of patient information, such as identification of allergies, medication lists, and prior visits.

This practice brief identifies fundamental components of building accurate MPIs and EMPIs and best practices for maintaining them. For the purpose of this practice brief, the term "MPI" encompasses both organizational- and enterprise-level activities of an electronically generated MPI.

Effect of EHRs

The healthcare industry is on a fast-moving train to implement EHRs. Accurate patient identification and its ability to affect all areas of the healthcare enterprise become the foundation for success. Accurate MPI databases will allow organizations and HIEs to improve patient care, reduce risks, improve operational efficiencies, support information exchange, and enhance the national health infrastructure.

Although MPIs traditionally may have been maintained on index cards, many organizations automated this labor-intensive paper index in the beginning stages of EHR development. The MPI is used during the registration application to ensure a patient is represented accurately and logically once in the database and ensures the same set of demographic and registration data are available throughout the organization.

With the adoption of EHRs, patients with more than one identifier will have isolated packets of information, such as treatment plans, allergies, and important historical medical care documentation, in the system. These isolated pieces of information create an actual loss of information because providers may not have a complete health record in front of them when making clinical care decisions.

Organizational-level MPI applications generally are categorized as vendor neutral or "best of breed" or a core vendor solution. The first classification implies that the patient information contained in the MPI can be integrated readily with any other EHR vendor application. Vendor-neutral applications can assist in patient identification across multiple EHR modules, such as laboratory, pharmacy, registration, and billing. A core vendor solution

generally is sold as an inherent or add-on module specific to a particular EHR system and may not readily integrate multiple disparate EHR systems across a large healthcare system.

Industry Effect

The American Recovery and Reinvestment Act of 2009 strengthened industry efforts regarding MPI applications. A key component to meeting stage 1 meaningful use criteria and, therefore, incentive payments is an organization's ability to exchange information for the purpose of care coordination. Such coordination requires tremendous efforts at the organizational and HIE levels to ensure care providers have access to the right information, for the right patient, at the right time.

Other industry initiatives include ongoing efforts of the NHIN and Connecting for Health's Common Framework. The NHIN is essentially a set of standards and policies that enables the secure transmission of health information over the Internet within an HIE environment. The NHIN will provide the foundation for health IT standards across organizations and communities and across the nation. This foundation will allow health information to follow patients as they move throughout the healthcare continuum. Currently, the NHIN is developing recommendations for standards for a broad audience.

Connecting for Health is a public-private collaborative that includes representatives from more than 100 organizations across the healthcare industry. Its purpose is to provide a catalyst for widespread industry changes needed to implement health IT while protecting a patient's privacy and security.

Connecting for Health's Common Framework approach was selected by the US Department of Health and Human Services as the initial prototype for the nationwide health information network. The Common Framework is composed of a set of technical and policy specifications designed to assist organizations and providers in creating data exchange and expected to require technical standards to support interoperability. One such infrastructure component is a record locator service (RLS). The RLS provides authorized participants of an HIE with the location of patient health information across the multiple participants (e.g., patient John Doe has health records at hospitals A and B and with physician A).

All of these current industry activities and initiatives will require accurate patient identification across multiple databases. The key to accurate and timely information exchange will be an accurate MPI.

MPI Deployment

MPIs are commonly deployed in either an active or passive mode by using existing Health Level Seven International messages, with additional data requirements being defined during the vendor selection process and implementation. An active deployment method implies that the MPI application is at the front end of the registration or scheduling process. Thus, patient identification is undertaken by using the MPI software, which requires integrating the MPI with legacy systems.

The user will identify the patient from a data repository, and at a select point in the identification pathway, the user drops to the facility-level registration or scheduling system. This process is generally transparent to the user.

A passive deployment method does not directly affect the registration or scheduling pathway. Rather, identification is undertaken behind the scenes or at the back end of the registration function. Generally, thresholds are established whereby a person is automatically linked or merged with existing data. If the threshold is not met, the registration data are held in a work queue for later resolution.

Many organizations and enterprises choose to launch MPI software applications in a passive mode initially and then migrate to active mode. Initial business goals, timelines, and budget will determine the deployment method for any given organization.

Both active and passive MPI methods should have the ability to identify persons at an enterprise and organizational level because initial deployment involves loading all databases to facilitate initial duplicate

identification. New patient identification policies and procedures must be formulated before launching an MPI, with careful consideration given to how the registrars will search for a patient if the MPI is deployed in an active mode, how to interpret scores and weights, and how to select the correct patient.

Data Elements

Because the MPI is essentially a database that maintains a unique identifier for each patient seen at the organizational or enterprise level, correctly identifying the data elements of the MPI is very important. Linking the correct clinical information from visit to visit and between enterprise activities requires that the core data elements be standardized and clearly defined. Once defined, these core data elements become the basis for matching algorithms, error reports, quality measurements, and staff training initiatives.

With no current healthcare industry standard for required data elements, organizations and HIEs often struggle in determining how much information to include in the MPI. At a minimum, MPI data elements should allow for the accurate matching of patients with a single identifier, facilitate a longitudinal health record, facilitate clinical system linkage, and improve access to patient health information. See appendix A for AHIMA-recommended data elements.

A review of the data elements associated with errors may reveal the need to collect new data elements or enforce the importance of accurately collecting existing data elements. For example, missing data from legacy system conversions or incomplete data collection during registration compromises the registrar's ability to select the correct patient. An important aspect of achieving and maintaining MPI data integrity is evaluating these procedural causes of duplicates and other issues affecting the MPI.

An organization should develop standard definitions of MPI data elements (data dictionary), standards for capturing and recording patient demographic data (naming conventions), and performance standards that hold staff accountable for accuracy. In addition, adherence to formal business processes and work flows will ensure data accuracy.

Algorithms for Identifying Duplicates

Algorithms are mathematical formulas used by an organization or enterprise that combine weighted data elements to determine the probability of a duplicate in order to identify potential duplicate MPI entries (see appendix B for a full list of terms and definitions). The matching algorithm is a critical component of any successful MPI solution.

The algorithm must be sophisticated, powerful, flexible, and accurate. Without a powerful algorithm to support accurate patient identification, the healthcare organization or enterprise will continue to create errors and will be forced to expend considerable time and money on maintenance efforts.

There are three types of matching algorithms available in the industry today: deterministic, rules based (sometimes known as ad hoc weighting), and probabilistic. Most organizations use deterministic, or exact match, algorithms, meaning they require exact matches on a combination of data elements such as name, birth date, sex, and Social Security number (SSN).

Deterministic algorithms are considered only 20 percent to 40 percent% accurate in patient identification and often result in a high volume of false matches. Therefore, one could expect less than half of the duplicates to be identified through the deterministic method. Deterministic algorithms are particularly weak in identifying individuals when there is transposition of numbers or letters, name changes, limited data, or large databases.

A more sophisticated technique for record matching uses rules-based algorithms. Rules-based algorithms are sometimes referred to as "fuzzy logic," or even mistakenly called "probabilistic." A rules-based algorithm allows an organization to assign weights, or significance values, to particular data elements and use these weights to

compare one record with another. This type of algorithm requires the facility to estimate weights in advance and then apply those weights to the data analysis process.

Because the organization determines the weights for data elements, the accuracy of rules-based algorithms varies widely. Using this method, only 50 percent and 80 percent of the potential duplicate record population may be identified. Usually, several repetitions of trial-and-error analysis on rules-based algorithms are required before an organization achieves acceptable results.

Probabilistic matching is considered to be the most sophisticated technique available, with an accuracy rate of 90 percent or higher. These algorithms are considered more advanced and more complex and able to support large data sets. Probabilistic algorithms are based on complex mathematical formulas that actually analyze the facility-specific MPI data to determine precisely matched weight probabilities for attribute values of various data elements.

For example, consider an MPI file where the name "Jones" appears more frequently than the name "Wheatley." A match on the name "Jones" has lower significance (less likely to be the same person) than a match on the name "Wheatley" (higher probability that a match represents the same person). Supporting digit transpositions and rotations, alternate name cross-referencing, distance editing, and enhanced phonetic searching can further enhance probabilistic algorithms. Therefore, many integrated healthcare delivery systems and HIEs choose probabilistic matching algorithms for enterprise-wide MPI activities.

Once determined, the algorithm is applied to a software application report that identifies potential duplicates. A threshold measure is used to interpret comparison scores based on the number of matches within required data element fields as defined in the algorithm.

For example, an organization has defined five data elements (patient name, date of birth, sex, SSN, and telephone number) as required matches for duplicate identification. The threshold for identification is defined as matching between three and five of the required elements. When scores are below the threshold (e.g., one data element matches), the records are assumed to represent different individuals, and the associated medical record numbers or enterprise identifiers are assumed to be accurate.

If the score is above the threshold (e.g., four of the five elements match), the record is assumed to be duplicative, and further research is required. If the score is between the two thresholds, the record is flagged and the information is placed in a work queue for review, resolution, and potential manual record review by the appropriate staff.

The table below shows an example of a deterministic algorithm. It shows how the exact matches regarding data elements date of birth, sex, SSN and telephone number are matched by the algorithm to produce a possible duplicate for John A. Smith.

Patient Name	Date of Birth	Sex	SSN	Telephone Number
John A. Smith	01-01-90	M	444-44-4444	555-555-5555
Johnnie A. Smith	01-01-90	M	444-44-4444	555-555-5555

Data Integrity

Data have integrity if they are complete, accurate, and consistent. A clean MPI contains only one record for each person and begins at registration. The diagram below illustrates how inaccurate information in the MPI has the potential to cascade to the various systems in an organization and can culminate in exporting erroneous information to an HIE:

? MPI
? Scheduling System
? Operating Room System

? Computerized Physician Order Entry
? Clinical Documentation System
? Transcription
? Radiology
? Pharmacy
? Laboratory
? Pathology
? HIM System
? Encoder
? Document Management
? Release of Information
? HIE

A review of the identified duplicates and overlays often reveals procedural problems that contribute to the creation of errors. Although HIM departments may be the hub of identifying, mitigating, and correcting MPI errors, that information may never be shared with the registration department. If the registration staff is not aware of the errors, how can they begin to proactively prevent the errors from occurring in the first place?

Registration process improvement activities can eventually reduce work for HIM departments. In addition, monitoring new duplicates is a critical process, and tracking reports should be created and implemented. Identifying and reporting MPI errors is important; however, tracking who made the error and why will decrease the number of duplicates.

The HIM, registration, and ancillary departments should establish communications to identify, report, and correct new duplicates, as well as methods for tracking, trending, and retraining individuals who created duplicate medical record numbers. Routine corrections of all identified duplicates must be a core organizational MPI maintenance function.

Calculating MPI Error Rate

Whether the MPI is at a local, enterprise, or HIE level, its primary purpose is to facilitate the link between clinical and administrative information between disparate systems. With so many patient care and industry initiatives at stake, the quality of MPI data can no longer be considered a back-end function.

Errors in MPI databases can lead to billing problems, unnecessary duplicate tests, and potential legal exposure. In addition, duplicates contribute to HIM operational workload and create inefficiencies as each new patient receives a new medical record number, file folder (in the paper world), and staff time in MPI maintenance activities.

Traditionally, MPI maintenance activities may not have received prioritization within the functional assignments of the HIM department for reasons such as chart assembly, analysis, coding, and abstracting. These other functions contributed to the revenue cycle of the organization and were given priority.

As EHRs were adopted and integrated, it became evident that the integrity of the MPI was in jeopardy because of the lack of merging and maintenance priorities. As disparate systems were implemented, an additional burden of correcting duplicates within multiple systems added increased administrative time in clinical departments such as the laboratory.

As the HIM department merges two duplicates together, the source system (laboratory) also must be corrected. This creates new challenges for organizations because merge functionality could be different in each system or module, which in turn creates data redundancy. Addressing ongoing errors within the MPI means an established quality measurement and maintenance program is crucial to the future of healthcare.

Organizations struggle to identify industry benchmarks and standards regarding MPI error percentages. Although

no organization expects to have duplicated patient information in the MPI, it is unrealistic to assume that duplicates will never occur. Even though errors may occur at the point of registration, the actual cause of error varies.

Duplicate information can occur as the result of insufficient patient information, clerical typos or transpositions, inability to capture important information (e.g., the patient is unconscious), and EHR system limitations in naming conventions. In addition, several processes also may contribute to duplicate creations.

The following processes can affect the quality of MPI data:

- Decentralized registration
- Converted data
- Lack of standards
- Lack of staff training
- Difficulties associated with registering laboratory specimens
- Accepting data from physician offices without verification procedures
- Lack of formal business processes and work flows

The duplicate error rate describes the quality of the MPI data. The error rate is calculated by dividing the total number of duplicate records with the total number of records multiplied by 100. An error rate is assigned to the MPI based on the file size and the number of duplicate records identified.

Industry experts estimate current organizational MPI error rates are between 7 percent and 10 percent and cost between $10 and $20 per duplicate to correct.[1] To put the duplicate MPI problem in proper perspective, if the organization has 300,000 patients in the MPI, there could be 30,000 duplicates, which, in turn, cost the organization $60,000.

Organizations can mitigate increased costs in MPI maintenance by developing MPI policies and procedures that clearly define review and maintenance activities.

Data Ownership

The issue of data ownership is a potentially difficult one that organizations must address early in their planning process, particularly if the corporation that is purchasing the MPI software license does not own all the participating facilities or sites. This challenge is further complicated by HIPAA because organizations must consider the relationship between covered entities, organized healthcare arrangements, business associates, and the obligations for disclosure of information to patients.

Participating for-profit entities of an integrated healthcare delivery system must consider the Gramm-Leach-Bliley Act limitations for sharing data.[2] Furthermore, as different facilities contribute data to the MPI, organizations should develop a comprehensive strategy to address demographic changes and duplicate resolution.

Participation in HIEs also will present challenges when patient-specific information changes. Specifically, organizations, integrated healthcare delivery systems, and HIEs should determine who will have the authority to change what level of data and how data changes will be communicated.

Enterprise and Corporate Identifiers

Inherent in the deployment of an EMPI is the assignment of the enterprise identifier. Although enterprise identifiers can be used for patient care, they are not commonly used by any downstream systems. Instead, they serve as a behind-the-scenes identifier to link and identify persons at a corporate level, with the existing identifiers such as medical record number or account number still providing identification at the local or facility level. However, with the push for more linking of clinical data to facilitate care across an integrated healthcare delivery system, organizations and technology may embrace the propagation of the corporate identifier.

Maintaining the MPI

The management of a high-quality, error-free MPI requires constant maintenance that includes oversight, evaluation, and correction of errors. The responsibility for MPI maintenance should be centralized under the direction of a qualified professional. Employees responsible for MPI maintenance must be carefully trained, have adequate tools and procedures, and be supervised to ensure their consistent compliance with established guidelines.

A comprehensive maintenance program should include:

- Ongoing processes to identify and address existing errors
- Advanced person search capabilities for minimizing the creation of new errors
- Mechanisms for efficiently detecting, reviewing, and resolving potential errors
- Ability to reliably link different medical record numbers and other identifiers for the same person to create an enterprise view of the person
- Consideration of the types of physical merges (files, film, etc.) and the interfaces and correction routines to other electronic systems that are populated or updated by the EMPI

The personnel performing duplication resolution activities require a foundation in the registration process to facilitate process improvement and ongoing communication between departments. This process includes an overview of the registration work flow, frequent obstacles encountered, and departmental expectations. In addition, training should include knowledge of downstream systems that are affected by duplicate identifiers and resolution activities in source systems.

Staff also should receive in-depth training, be supported by detailed policies and procedures, and be able to demonstrate competency in the areas of duplicate identification and resolution. Ongoing education should include feedback from a well-defined quality program. Updates to the training program should be performed periodically and should be based on a review of the initial training to incorporate system modifications, upgrades, and internal process improvements.

All training should be performed before beginning any duplicate resolution or maintenance activity.

Adequate staffing is needed to maintain and ensure the quality of the MPI. Staff members should have the authority to resolve duplicates, investigate demographic overlays, and link persons across the enterprise. See appendix C for a sample job description.

MPI maintenance policies and procedures also should outline:

- Whether to use the most recent information for each data element or criteria for determining which data elements will be stored as the patient's information
- Identification of duplicate data elements (e.g., prior name or alias)
- Communication of merges to ancillary staff for source system revision
- Regular review of error reports and trending of duplicate percentages

MPI management should be a key strategic activity for any healthcare organization, integrated healthcare delivery system, or HIE as the backbone of the EHR and the support for a longitudinal patient record. The integrity of the MPI is more important now than ever before. MPI maintenance and management requires a leader who possesses a variety of skills and experiences. There are vast opportunities in the healthcare industry today for professionals who can demonstrate knowledge of both information management and technology that supports MPI database integrity.

Appendices

Notes

1. Hess, Robbi. "Masters of Their Domain." *For the Record* 17, no. 17 (Aug. 15, 2005). Available online at www.fortherecordmag.com/archives/ftr_081505p30.shtml.

2. The Financial Services Modernization Act of 1999, also known as the Gramm-Leach-Bliley Act, requires financial institutions to provide customers with a notice of privacy policies and procedures and to satisfy various disclosure and consumer opt-out requirements.

References

AHIMA. "Reconciling and Managing EMPIs." *Journal of AHIMA* 81, no. 4 (Apr. 2010): 52-57.

Altendorf, Robin L. "Establishment of a Quality Program for the Master Patient Index." AHIMA's 79th National Convention and Exhibit Proceedings, October 2007. Available online in the AHIMA Body of Knowledge at www.ahima.org.

Brown, Eric, Jim Burke, Erica Foster, and Thea Campbell. "Accurate Patient Identification: The Foundation for EHRs." AHIMA's 78th National Convention and Exhibit Proceedings, October 2006. Available online in the AHIMA Body of Knowledge at www.ahima.org.

Connecting for Health Common Framework. "Correctly Matching Patients with Their Records." April 2006. Available online at www.connectingforhealth.org/commonframework/docs/P4_Correctly_Matching.pdf.

Connecting for Health Common Framework. "Record Locator Service: Technical Background from the Massachusetts Prototype Community." Available online at www.connectingforhealth.org/commonframework/docs/T6_RecordLocator.pdf.

Department of Health and Human Services. "Nationwide Health Information Network: Public Forums and Other Products." Available online at www.hhs.gov/healthit/healthnetwork/forums/.

Dimick, Chris. "Exposing Double Identity at Patient Registration." *Journal of AHIMA* 80, no. 11 (Nov-Dec 2009): Web extra. Available online in the AHIMA Body of Knowledge at www.ahima.org.

Greenwald, Constance Z., and Jason Broad. "Master Patient Index Overlaps (2 Patients, 1 Record): We Don't Know What We Don't Know." 2009 AHIMA Convention Proceedings, October 2009. Available online in the AHIMA Body of Knowledge at www.ahima.org.

Health Level Seven International. "V3 Messaging Standard." Available online at www.hl7.org/implement/standards/v3messages.cfm.

Hewitt, Joseph B., and Michele O'Connor. "Connecting Care through EMPIs." *Journal of AHIMA* 73, no. 10 (Oct. 2002): 32-38.

Hieb, Barry R. "The EMPI Magic Quadrant for 2001: A Maturing Market." Market Analysis,

Nunn, Sandra. "Risk-Based MPI Management." *Journal of AHIMA* 79, no. 6 (June2008): 48-49, 56.

Wheatley, Vicki. "Quality Impact of the Master Patient Index" *Journal of AHIMA* 79, no. 10 (Oct. 2008): 78-79.

Wheatley, Victoria. "Unique Identifiers: Preparing for HIPAA." 2002 AHIMA Convention Proceedings, September 2002. Available online in the AHIMA Body of Knowledge at www.ahima.org.

Prepared by

Lou Ann Wiedemann, MS, RHIA, FAHIMA, CPEHR

Acknowledgments

Cecilia Backman, MBA, RHIA, CPHQ
Jill Clark, MBA, RHIA
Jane DeSpiegelaere, MBA, RHIA, CCS
Angela K. Dinh, MHA, RHIA, CHPS
Lesley Kadlec, MA, RHIA
Carol Melvin, RHIA
Elaine O'Bleness, MBA, RHIA, CHP, CHPQ
Harry Rhodes, MBA, RHIA, CHPS, FAHIMA
Susan Richardson, MHSA, CPHQ, RHIA
Diana Warner, MS, RHIA, CHPS

Prepared by (original)
AHIMA MPI Task Force:
Lorraine Fernandes, RHIA, Chair
Mary Brandt, MBA, RHIA, CHE, CHP
Donna Fletcher, MPA, RHIA
Karen Grant, RHIA, CHP
Leanna Hatton, RHIT
Susan Postal, MBA, RHIA
Vicki Wheatley, MS, RHIA
Terry Winter, MEd, RHIA, CHE

Acknowledgments (original)
Victoria Barcena, RHIA
Jill Burrington-Brown, MS, RHIA

Appendix A: Recommended Core Data Elements for EMPIs

Data Element	Definition	Data Type*
Internal patient identification	Primary identifier used by the facility to identify the patient at admission (e.g., the medical record number)	Extended composite ID with check digit
Person name	Legal name of patient or person, including surname, given name, middle name or initial, suffixes (e.g., junior, IV), and prefixes (e.g., Father, doctor)	Extended person name

Date of birth	Year, month, and day of birth (e.g., YYYY/MM/DD)	Time stamp
Gender	For example, male, female, unknown, or undetermined	Coded value
Race	Race is a concept used to differentiate population groups largely on the basis of physical characteristics transmitted by descent. Races currently used by the federal government for statistical purposes are American Indian or Alaskan Native, Asian or Pacific Islander, Black, White, Unknown, Other, and Missing.	Coded value
Ethnicity	Ethnicity is a concept used to differentiate population groups on the basis of shared cultural characteristics or geographic origins. Ethnic designations currently used by the federal government for statistical purposes are Hispanic origin, not of Hispanic origin, and Unknown, Other, and Missing.	Coded value
Address	Address or location of patient's residence. Components include the street address, other designation (e.g., apartment number), city, state/province, zip or postal code, country, and type of address (e.g., permanent, mailing).	Extended address
Telephone number	Telephone number at which that patient can be contacted. This may be a home or business telephone number or the telephone number of a friend, neighbor, or relative.	String data
Alias/previous/maiden names	Any names by which the patient has been known other than the current legal name, including nicknames, maiden name, previous name that was legally changed, etc. All previous names available should be converted and retained.	Extended person name
Social Security Number	Personal identification number assigned by the US Social Security Administration	String data
Facility identification	The unique identification number of a facility where patients seek care. (The Centers for Medicare and Medicaid Services has developed a provider ID system for healthcare facilities.)	Person location
Universal patient identifier	Not yet established	N/A
Account/visit number	Number assigned by the facility billing or accounting office for all charges and payments for this encounter or visit	Extended composite ID with check digit
Admission/encounter/visit date	Date the patient actually arrived for care (e.g., YYYY/MM/DD/HH/SS)	Time stamp
Discharge or departure date	Date the patient actually left the facility or died (e.g., YYYY/MM/DD/HH/SS)	Time stamp
Encounter/service type	Categorization of the encounter, such as emergency, inpatient, outpatient, home care, or electronic (e.g., e-mail, Internet, telemedicine)	Coded value
Encounter/service location	Location in which the encounter, visit, or treatment occurred	Coded value
Encounter primary physician	Attending physician for the associated encounter, visit, or treatment, identified with the primary physician's National Provider Identifier	Coded value
Patient disposition	Patient's intended care setting after discharge. Examples include discharge	Coded

	home (not to home health service), acute care hospital, nursing facility, home to be under the care of a home health service, or other healthcare facility; left against medical advice; alive, other, or not stated; died; admitted to hospital; admitted for observation; transferred to skilled nursing facility, intermediate care facility, or other facility; or other disposition as dictated by type of MPI.	value

*Data types correspond to those described in Health Level Seven International's Application Protocol for Electronic Data Exchange in Healthcare Environments. Version 2.6. Version 2 Messaging Standard.

The table "Additional Recommended EMPI Data Elements" below lists additional data elements that can be stored in EMPIs.

Additional Recommended EMPI Data Elements		
Data Element	Definition	Data Type*
Mother's maiden name	The given, family, or last name of the patient's mother	String data
Marital status	Marriage status of the patient (e.g., never married, married, separated, widowed, divorced, or unknown). Organizations must determine whether marital status will be tied to a visit or just reflect current status and whether historical information for specific visits will be available.	Coded value
Place of birth	City, state, and country of the patient's birth	String data
County	County in which the patient lives	Coded value
Blood type/Rh	Patient's blood type or Rh factor	Coded value
Employer	Name of patient's employer	String value
Work telephone	Patient's work telephone number	String value
Advance directive and surrogate decision making	An indication that the patient has an advance directive on file. It describes an individual's current preferences about treatment should the person become incompetent or unable to communicate these preferences to medical personnel. Surrogate decision making is an alternative method for medical decision making on the individual's behalf. It is invoked in the absence of an advance directive when the individual is not competent to make an informed decision.	Boolean
Organ donor status	Whether the patient has consented to donate his or her organ(s) in the event of death	Boolean
Emergency contact name	Name of the person whom the patient wishes to be the primary contact if notification is necessary	String data
Emergency	Relationship to the patient of the person whom the patient wishes to be the primary	String

contact relationship	contact if notification is necessary	data
Emergency contact address	Address of the person whom the patient wishes to be the primary contact if notification is necessary	String data
Emergency contact telephone	Telephone number of the person whom the patient wishes to be the primary contact if notification is necessary	String data
Guarantor name	Name of the person responsible for the payment of the patient's bill	String data
Guarantor relationship	Relationship of the person responsible for the payment of the patient's bill	String data
Guarantor address	Address of the person responsible for the payment of the patient's bill	String data
Guarantor telephone	Telephone number of the person responsible for the payment of the patient's bill	String data
Payer information	Type of payer (e.g., commercial insurance, Medicare, self, etc.), including policy information such as payer name, policy number, etc.	String data, separate fields for each data element
Problem list	Master list of all of a patient's health problems or diagnoses	String data
Encounter primary physician contact address	Encounter primary physician business address	String data
Referring physician	Referring physician for the associated encounter, visit, or treatment, identified with the physician's National Provider Identifier	Coded value
Referring physician contact address	Referring physician's business address	String data
Receipt of notice of privacy practices	Whether the notice of privacy practices been given to the patient	Coded value

*Data types correspond to those described in Health Level Seven International's Application Protocol for Electronic Data Exchange in Healthcare Environments. Version 2.6. Version 2 Messaging Standard.[1]

Note

1. AHIMA. "Reconciling and Managing EMPIs." *Journal of AHIMA* 81, no. 4 (Apr. 2010): 52-57.

Appendix B: Glossary

Active Deployment: MPI software application deployment that uses a method for identifying patients at the front end of the registration or scheduling process.

Algorithms: MPI algorithms are mathematical formulas that combine weighted data elements to determine the probability of a duplicate. Used to identify potential duplicate MPI entries.

Corporate Identifier: unique link that identifies and links persons at a corporate or enterprise level.

Data Ownership: process of defining, within EMPIs, the owner of the individual data contained within the database.

Deterministic Algorithm: MPI algorithm that requires an exact match of combined data elements such as name, birth date, sex, and Social Security number.

Duplicate: more than one unique identifier (e.g., medical record number or person identifier) for the same person in the MPI. This causes one patient to have two, or more, different medical records within the same facility.

Enterprise Master Patient Index (EMPI): index referencing of all known patients in two or more facilities (e.g., integrated healthcare delivery system or HIE).

Error Rate: total number of duplicate patient identifiers within a database. Calculated by dividing the total number of duplicates by the total number of records multiplied by 100.

Integrated Healthcare Delivery System: coordinated healthcare delivery system that includes physicians and hospitals. Integrated healthcare delivery systems provide a broad range of health services, such as clinic visits, ambulatory surgery, urgent care clinics, inpatient hospitalization, rehabilitative services, and behavioral health services.

Maintenance: routine actions an organization undertakes to review and correct MPI data errors.

Master Patient Index (MPI): index referencing of all known patients to a single organization linked together by a common identifier, usually the medical record number (used interchangeably with EMPI).

MPI Data Elements: key data points within patient demographic information that will be included in the MPI (e.g., patient name, sex, race, date of birth, medical record number, Social Security number).

MPI Error Rate: percentage assigned to the number of potential duplicates within a given MPI. Calculated by dividing the total number of duplicate records by the total number of records multiplied by 100.

Overlap: more than one unique identifier for the same person across two or more facilities in the enterprise (e.g., patient John Smith has medical record number 12345 at facility A and medical record number 447788 at facility B).

Overlay: one EMPI record for more than one person (e.g., two people sharing the same identifier).

Passive Deployment: MPI software application deployment method--this is behind the scenes or at the back end of the registration function.

Probabilistic Algorithm: MPI algorithm based on complex mathematical formulas that analyzes facility-specific MPI data to determine precisely matched weight probabilities for attribute values of various data elements. Used

most often in large databases such as HIEs.

Record Locator Service (RLS): information service that locates patient health records across the multiple systems that subscribe or participate in HIE. A key piece of infrastructure within the health information environment in which authorized care providers submit a request for information to an HIE. The RLS responds to the request by providing information to the requester regarding pertinent records (e.g., health records for John Doe can be found at hospitals A and B and with physician A).

Resolution: act of identifying, assigning, correcting, and merging duplicates within the MPI (also called "merging").

Rules-Based Algorithm: MPI algorithm that assigns weights, or significant values, to particular data elements and later uses these weights in the comparison of one record to another.

Appendix C: Sample Job Description

Department Name:	Cost Center:
Status/Pay Grade:	Effective Date:
Job Title: MPI Coordinator	Reports to:
Hire Date:	Annual Review Date:

Position Summary/General Purpose: The master patient index (MPI) is a critical business function and resource because it links a patient to demographic, clinical, and financial information. For those purposes, the MPI must contain accurate, timely, and complete data that include a single identifier for each patient registered. The MPI coordinator will be responsible for reviewing, analyzing, and maintaining data integrity within the MPI. In addition, this position is responsible for ensuring that the information exchanged between (insert organization name) and health information exchange (HIE) databases is accurate and that corrections are communicated in a timely manner.

Position Requirements/Qualifications:

Licensure/Certification/Registration: RHIA, RHIT, CHDA certification preferred.

Education: Associate or bachelor's degree in health information management preferred.

Experience: Prior healthcare experiencepreferred. Previous health data analyst experience preferred.

Degree of Supervision: Employee must be able to work independently, effectively, and efficiently on his or her own. This position has no direct reports.

Physical Requirements: Employee must be able to perform the essential functions of this job. This position includes a requirement to walk, climb stairs, balance, and sit. Lifting is limited to health information charts and is not expected to exceed 25 pounds. Vision requirements include the ability to read paper and electronic health records.

Responsibilities:

- Monitor duplicate analysis reports daily to identify potential overlaps and overlays
- Review, prioritize, and facilitate merging of facility and HIE duplicates

- Report monthly MPI data integrity efforts to the appropriate committee
- Provide ongoing training to registration staff on MPI data entry and the importance of MPI data integrity
- Provide training to HIM personnel assigned to resolution and merging activities
- Provide facility-wide training on data-integrity efforts
- Review organizational policies and procedures regarding merging and provide input on key indicators such as which information to retain when merging duplicates
- Coordinate and communicate MPI activities and merges with other departments (e.g., laboratory) for source system resolution
- Retrieve, analyze, and make appropriate MPI changes in both electronic and paper records within the appropriate system

Knowledge/Skills

- Proficient in Microsoft Word, Excel, Access, and PowerPoint
- Ability to present complex information in an understandable and compelling manner
- Knowledge of coding classification systems
- Ability to analyze information to determine appropriate resolutions and merges
- Excellent written and verbal communication skills
- Strong customer service skills

Article citation:
AHIMA. "Fundamentals for Building a Master Patient Index/Enterprise Master Patient Index (Updated)." *Journal of AHIMA* (Updated September 2010).
