Implementing the Enterprise Master Patient Index

by Lisa Adragna

In implementing a cross-facility initiative, the importance of planning and understanding the implications for all facilities can't be overlooked. Here's how one integrated delivery network navigated the challenges of implementing a cross-facility enterprise master patient index.

Partners HealthCare System, Inc., established in March 1994, oversees the affiliation of Brigham and Women's Hospital, The Massachusetts General Hospital, and The North Shore Medical Center. Other member organizations include Dana-Farber Cancer Institute, Partners Community HealthCare, Inc., Spaulding Rehabilitation Hospital, and McLean Hospital.

As a newly formed integrated delivery network, Partners faced the considerable challenge of uniquely identifying patients within the enterprise and integrating clinical data over time using disparate entity-level information systems. This challenge became the primary goal of the enterprise master patient index (EMPI).

Project Scope and Related Goals

The EMPI was one of the first cross-facility initiatives Partners had undertaken since its inception. The ability to uniquely identify a patient at the corporate level was a strategic component of the Partners technology infrastructure and the first step to becoming a fully integrated delivery network. Although unique patient identification was the primary goal, secondary goals were just as important. These goals included:

- eliminating the need to renumber entity-level medical record numbering systems by assigning a unique Partners identifier
- providing a foundation for the clinical data repository
- developing a minimally invasive approach to current operations
- developing a rapid installation time frame

The EMPI project scope was defined as follows:

- *Providing a medical records crosswalk across Partners.* The EMPI would house each of the entity-level medical record numbers under the Partners corporate identifier, thus acting as an enabler for the clinical applications
- *Back-end real time (one-way) interface*. The decision to go with this back-end approach was based on the requirement that no modifications were to be made to the entity-level applications
- *Providing automated tools for identifying and resolving potential duplicates*. The back-end implementation approach meant that there would be no human intervention. Automated tools had to be provided to the end user in order to ensure data integrity

Project Structure

To implement a project of this magnitude and complexity, it was imperative to have the proper means to support it and to obtain buy-in from all Partners entities. In addition to the core project team, a formal project structure was established to provide continuity as well as operational and technical expertise.

- *The operational steering committee* was responsible for addressing cross-facility issues, providing operational expertise, approving the EMPI management model, and serving as the primary contact for the facility it represented

- *The technical steering committee* was responsible for providing technical expertise, evaluating vendor products and making recommendations, as well as reviewing and approving the EMPI technical model
- *The operational workgroup* was established to develop recommendations for the EMPI management approach, data set, and entity-level cleanup and conversion strategies. The workgroup members included representatives from all entity-level functional areas, including health information management, patient access, and information systems. The group's mission was to determine how the EMPI department would maintain the accuracy of the data stored at the corporate level. This task was not an easy one, given that the back-end approach had no human intervention. The group grappled with complicated issues, including:

    - Who could view/update the EMPI?
    - Who could view/update entity-level MPIs?
    - Who should resolve potential duplicates across entities?
    - Who should resolve potential duplicates within entities?
    - How should this corporate database be managed?
    - What are the confidentiality issues associated with viewing and updating entity-specific information at the Partners level?

The core project team, sponsored by the CIO, was responsible for planning and implementing the project.

Partners' EMPI Strategy

EMPI Data Set

In keeping with the project scope definition, it was decided that the EMPI would retain a limited demographic data set. This decision was made because the EMPI was designed to be the cornerstone for the clinical data repository, and because corporate management needed a six-month installation period.

Therefore, no visit or insurance information would be passed or maintained. It was impossible to map insurance and visit data across multiple entities within the aggressive six-month timetable. It was also decided that any incoming registration transaction would have to meet a minimum requirement data set. The Partners interface engine validates the incoming registration transactions for data completeness and accuracy. The check marks identify the minimum required data set needed to assign a unique EMPI number.

These data elements are currently stored in the Partners database:

- ☑ Last Name
- ☑ First Name
- ☑ Gender
- ☑ Date of Birth
- ☐ Social Security Number
- ☐ Address
- ☑ Entity Medical Record Number(s)
- ☐ Home Phone
- ☐ Work Phone
- ☐ Language
- ☐ Religion
- ☐ Marital Status
- ☐ Race
- ☐ Veteran Status
- ☐ Restricted Access Flag
- ☑ Facility ID(s)

Because several Partners entities used the same medical record numbering schemes, the facility ID from the originating facility was required. If an incoming registration transaction did not meet this minimum data set, then an EMPI number would not be generated for the record.

Technical Model

The EMPI relies on the passing of registration data by the entity system once a registration is completed at the site. Because Partners made the decision to be noninvasive to current operations, entity-level systems were not modified. The technical model primarily focuses on the data after it leaves the site. All transactions that are not HL7-compliant are converted and then

sent to the EMPI via the Partners Interface Engine. The EMPI application makes a determination as to whether the patient exists based on a sophisticated probabilistic matching algorithm. One of the following three options is possible:

- if the patient does exist, the transaction is processed and the EMPI is updated accordingly
- if the transaction potentially matches one or more patients in the database, a new EMPI number is generated for the patient and the record is flagged as a potential duplicate
- if the incoming registration transaction does not match any patients in the database, a new EMPI record is created and a unique number is assigned

Once the registration transaction is reconciled within the EMPI Oracle database, it is then sent back through the Partners interface engine to a replicated EMPI database. The Oracle database is replicated to more efficiently serve the Partners clinical applications, which are all internally developed applications. The replicated database is a mirror image of the Oracle database.

The EMPI application uses an Oracle database engine version 7.3.3 running the Windows NT operating system version 4.0 on a Compaq Proliant 5000 hardware platform.

EMPI Functionality

The initial phase of the implementation focused on the need to maintain the integrity of the data passed to the EMPI Oracle database. Following are the current capabilities of the application:

- Search for a given patient in multiple ways:
  - EMPI number
  - social security number
  - entity medical record number
  - patient name
- Add a new record
- Update an existing record
- Merge two EMPI records
- Unmerge two EMPI records
- Deactivate an EMPI record
- Add an entity medical record number to an existing EMPI record
- Flag a record as a potential duplicate
- Secure a given EMPI record within the database for patient confidentiality reasons
- Track all changes made to a given record through an audit trail function
- Produce desired reports both online and in hard-copy format

Vendor Selection Process

Industry Review

The Partners EMPI project began as a strategy initiative in January 1996 to determine best practices for implementing a unique patient identifier. Partners conducted an extensive industry review through the use of consultants, then the project team contacted several healthcare organizations and integrated delivery systems across the country to gain a better understanding of the current industry status. Through this extensive process, it was determined that the industry was still in a very immature stage. No integrated delivery system had successfully implemented an enterprise-wide master person index across multiple facilities using disparate hospital information systems.

Vendor Review

The project team then began to investigate possible vendors whose applications could support Partners' need for a unique patient identifier. Many vendors claimed they had an enterprise solution; however, the project team confirmed that most had few clients, and even fewer had operational sites. Nonetheless, 16 potential vendors were identified.

The project team prepared a request for information (RFI) based on the technical model that dictated the back-end approach, the EMPI data set, and the approved list of "must have" functionalities. The RFI was distributed to the 16 potential vendors. The RFI responses were reviewed and evaluated, and seven vendors were brought in to demonstrate their applications. Both the operational committee and the technical steering committee were invited to participate in the first round of vendor demonstrations. Vendor evaluation forms were prepared by the project team, and steering committee members were asked to submit their opinions in writing via the form.

After the first round of demonstrations was completed and references checked, three finalists were brought back to address their deficiencies and further demonstrate their products. After the second round of demonstrations was completed, the project team scored the evaluation forms and prepared the gap analysis. The forms and the gap analysis were reviewed with the technical steering committee, along with the prioritized vendor evaluation criteria.

The last option to receive careful consideration was the Partners internal build option. Partners had always reserved the right to internally develop the application if the supported vendor packages could not meet the specified requirements. But because internal developers were so consumed with the development of the Partners clinical applications, a vendor package was the preferred approach. Partners chose the vendor that best met its functional requirements, rapid installation time frame, and cost prerequisites.

EMPI Management Model

The operational workgroup believed that efficient management of the EMPI would require sophisticated coordination between the EMPI department and the entity-level medical records personnel. The workgroup determined that the management of the EMPI would require a dedicated corporate staff working in conjunction with each entity.

The Partners EMPI is currently managed by three full-time employees who are solely responsible for resolving cross-facility potential duplicates as well as updating and maintaining the EMPI database. Because the matching algorithms used by the EMPI application are more sensitive than those used by the entities, corporate staff often discover within-entity duplicates before the site does. In this instance, specific procedures were developed that describe the intracommunication that takes place between the enterprise and the entity on a daily basis.

The EMPI project team worked very closely with the EMPI department to help develop the procedures that currently govern the communication between the entities and the enterprise. For synchronization purposes, all within-entity duplicates continue to be handled by the entity and pass systematically to the enterprise on a real-time basis. The EMPI corporate staff has read-only access to the entity-level MPI information to help facilitate its research.

Matching

One of Partners' highest priorities was the inclusion of a sophisticated matching algorithm into the EMPI application. The back-end approach implied that no human intervention would be available at the time a registration transaction was received via the back end. It was this back-end requirement that prompted Partners to place such a high importance on the probabilistic matching algorithm.

When an interface transaction is received via the back end, the database is first searched by medical record number and facility ID. If a match is found, the record is updated accordingly. If the medical record number is not located in the database, then a patient name look-up is performed. The database searches in three distinct ways:

- patient's last name and first name
- patient's first name, date of birth, and gender
- patient's social security number

Once the list of candidates is identified using all three of the above searches, a weight is assigned to each. Based on the weights returned, one of the following actions is taken:

- if the patient search yields a definite no match, a new EMPI number is assigned and the new record is added to the database

- if the record is a potential duplicate of yet another record in the database, a new EMPI number is assigned, and the incoming transaction is flagged as a potential duplicate. The potential duplicate is then worked on manually by the EMPI administrative staff
- if the record is an exact match (on last name, first name, sex, date of birth, and/or social security number) with an existing record in the database, the medical record number is added to the existing EMPI record

MPI Cleanup

The primary reason for implementing the EMPI was to provide a foundation for the clinical data repository. The more potential duplicates each site had, the more fragmented the clinical view would become. The operational workgroup was concerned about the "cleanliness" of its own master patient indices. To get a better understanding of the depth of the cleanup effort as well as the degree of overlap across institutions, five of the seven Partners entities submitted an abstract of their MPI for analysis. Although the potential duplicate results came back under the national average, each entity's MPI had enough corrections to warrant both the cost and time required to conduct a full-scale cleanup effort.

Conversion

The cleanup decision would have a direct effect on the window of data converted. How many years of data should the entities clean up? The operational workgroup believed that the cleanup effort should coincide with the conversion window. Considering the number of records needing correction, coupled with the resources required to analyze and merge the records, the group agreed to convert back to January 1, 1995. With a live date of August 1997, this conversion would provide more than 2.5 years worth of data.

Implementation

The size of the individual entity MPI databases was key in determining the implementation sequence of the Partners entities. The operational workgroup felt that two sites should be identified for the initial implementation. Both the Brigham and Women's Hospital (BWH) and Massachusetts General Hospital (MGH) were selected based on the large size of their respective MPI databases. The Dana-Farber Cancer Institute was chosen as the next site to be implemented after the initial "go live" was completed. The remaining sites would be selected based on the patient overlap statistics and the institution's ability to resolve its own potential duplicate populations prior to conversion. The conversion process was slated to begin August 1, 1997, with an estimated completion date of August 10, 1997.

The EMPI has been live since August 11, 1997. The cross-facility potential duplicates created as part of the conversion process have all been resolved. During the last six months the project team has actively worked with end users to resolve system problems, refine policy and procedures, and begin rollout to other Partners entities. The following reflects the EMPI statistics to date:

- Approximately 1 million patients were converted
- Conversion of BWH and MGH data took approximately 10 days
- After the conversion process was complete, approximately 170,000 pairs were identified as potential duplicates
- A full 80 percent of the registration transactions received on a daily basis are updates to an existing EMPI record
- Between BWH and MGH, 1800 registration transactions are passed to the EMPI on a daily basis
- There are 1.2 million records in the EMPI database
- The sites have completed 3308 within-entity merges, which systematically passed to the EMPI
- The EMPI administrative staff has resolved over 200,000 potential duplicates, both systematically and manually
- The Dana-Farber Cancer Institute has been added to the system, and Spaulding Rehabilitation Hospital is currently in the process.

Lisa Adragna *is a corporate manager for Partners HealthCare System, Inc. Her primary focus is on planning and implementation of enterprise-wide applications. A version of this article originally appeared in a booklet distributed by the Massachusetts Health Data Consortium. Reprinted with permission.*

Article Citation:

Adragna, Lisa. "Implementing the Enterprise Master Patient Index." *Journal of AHIMA* 69, no. 9 (1998): 46-50.