Applying Innovation to the Patient Identification Challenge

Save to myBoK

By Lorraine Fernandes, RHIA; Jim Burke; and Michele O'Connor, MPA, RHIA, FAHIMA

Healthcare transformation activities involving data analytics, population health, disease registries, and consumer engagement have raised the stakes for stronger and real-time, or near real-time, patient identification and record matching.

Yet many organizations have hundreds of thousands—or even millions—of electronic records that can't be used in transformational activities because they can't automatically be matched/linked to other records, thus compromising organizational goals for patient-centric care.

As numerous studies, publications, and articles have demonstrated over decades, unmatched records result from "thin data" (poor quality data) being collected at some organizations, lack of standardization of data fields, human errors, and numerous other causes. Latency of data can be a particularly vexing record matching issue since demographic data frequently changes and consumers seek care in multiple settings.

While patient safety remains the top priority for all healthcare providers, industry pressures like regulatory compliance, the shift to value-based payments, and extensive use of matched data in analytics demand that health information management (HIM) professionals bring leadership and innovation to this challenge. And let's not lose sight of the Office of the National Coordinator for Health IT's Interoperability Roadmap, which requires organizations nationwide reach less than a 0.5 percent duplicate medical record number error rate by 2020.[1]

This article explores innovations that can move the healthcare industry beyond the traditional human resource-heavy, back-end retrospective approach to accurate, automated patient identification and record matching. This includes new approaches, such as augmentation using data from outside healthcare and using neural network software to resolve ambiguous identity data. It can also involve assigning a unique proprietary identifier with the assistance of external data.

The examples in this article address metrics, cost benefits, and privacy and security concerns that correlate with bedrock information governance principles being applied by payers, providers, and health information exchanges.

Terri Godar, director of technical operations and eligibility at Advocate Health System, based in Chicago, IL, says that healthcare organizations can't solve the patient identity challenges within their four walls. "They need to innovate by infusing richer data that comes from external sources that may include credit bureaus and government programs. In addition, we have found that demographic data from payers, which is increasingly important in the priorities of today's vast healthcare ecosystem, is outdated and often incomplete," Godar says. "Higher data quality and matching can also be achieved by implementing stronger patient registration, patient access, and patient search processes."

Using Cloud-based Data Services to Gain Trust in Your Data

The mainstream acceptance of cloud computing has opened an avenue to incorporate secure external data services into critical business processes such as patient registration, data exchange, and patient identification.

Cloud-based data services enable the infusion of referential or authoritative data that may come from large public databases outside healthcare, such as credit bureaus, loan servicing organizations, or telecommunications. These non-healthcare databases and associated business processes capture and validate identity data, update it continuously with each transaction, and retain the history of the person's demographics.

A common challenge in healthcare is that a patient's demographics can change between encounters at different facilities. Demographic data recorded at Facility A in January, while accurate at that time, can differ from demographic data recorded at Facility B in June—also accurate at that time—if the patient's actual demographic information changed between

those two dates. If the data from Facility A and Facility B do not closely match, the two identities may not automatically be linked. Referential identity data can help resolve these issues.

Real-time automation of patient matching with external data also addresses a critical latency issue associated with manual stewardship efforts, which typically don't resolve the ambiguous linkages/tasks (those records not automatically linked by an algorithm) until days or months after a patient presents for care.

How Healthix HIE Handles Patient Identification

HEALTHIX IS THE largest public health information exchange (HIE) in the nation, serving a comprehensive range of organizations in the greater New York City area.

- Before applying external referential data to augment identity reconciliation, Healthix had received: 51.1 million medical record numbers (MRNs) from their provider organizations and resolved them to 25.4 million actual identifiers (persons).
- Over a four-month period, Healthix applied referential identity data to the whole database. In the four months, MRNs continued to increase to 54.1 million, but referential data resolved them to 21.9 million identities.
- The number of MRNs per identity increased by 22 percent, from 2.01 to 2.47.
- The 21.9 million unique individual records are now available to meet the key clinical and operational needs.

Health information exchanges (HIEs) must deliver high-value and integrated data to their stakeholders despite data challenges. Tom Check, CEO of Healthix, a public HIE based in New York, says his organization faced very significant numbers of potential matches that weren't strong enough to automatically match given the conservative threshold that had to be applied (see the sidebar above). "Healthix needed more comprehensive, actionable information about an individual to serve value-based payment needs, and to provide trusted information to its providers, care managers and care coordinators, and research," Check says.

A linked and comprehensive view of patient/consumer data is needed to support quality and financial reporting for programs such as HEDIS and CORE Measures, as well as care delivery and predictive risk modeling. "By augmenting Healthix data with external referential data we are creating complete and current data," Check says. "The whole record, not fragments we previously encountered, is essential to tracking and predicting risks so providers can intervene early, thus enhancing quality and reducing costs." Additionally, Check pointed out the following healthcare Big Data and automation imperative: "The volume of data and automation applied to analytics necessitates timely, trusted, and comprehensive views of the patient/consumer data."

Auto-Stewarding with Neural Networks

Neural networks technology (machine learning) has been around for over 50 years and used in biometric operations like fingerprint analysis and facial recognition. It has enjoyed a resurgence in the past five years due in part to the explosion of Big Data, proliferation of large-scale networks, and the maturing of practical application of this technology.

The shift from a linear data model to a neural data model has proven to be the key to major neural network improvements that can be applied to patient identification. In his article "A Resurgence of Neural Networks in Machine Learning," Google research scientist Dan Gillick explains, "in machine learning, 'training' refers to the process of automatically choosing the weights given examples of the input where you know what the output should be."[2] Training the neural network through sufficient sample data is the key to producing the desired results, but in the context of healthcare this is quite practical.

As Graham Jones, MD, business development director at Kestrel Consulting Services, noted in his recent blog post, "the production platform that implements the neural network, once trained, is often very simple and cost effective."[3]

In the case of St. Joseph Health (SJH), based in Orange County, CA, they tested the use of neural network technology to resolve high volumes of ambiguous linkages or tasks. The data in these ambiguous linkages, which heretofore represented fragmented records, is critical to serving the needs of population health, consumer engagement, and numerous types of analytics. "Our data-driven initiative that began in 2012 prioritizes freeing data, sharing data, and using data to meet health systems' strategic initiatives," says Theo Siagian, director of HIE and interoperability at SJH. "Having a significant volume of

ambiguous linkages impeded meeting these priorities. In a proof of concept project, the supervised-learning technology (neural network) is constantly improved by mimicking the behaviors of our best data stewards, and freeing these individuals to address the more complex ambiguous linkages."

Further, Kathy Fitzgerald, RHIA, director of HIM and privacy officer at SJH, stated that "resolving ambiguous data sets must be timely, as delays impact data use."

"Applying machine learning through neural networks 'training' helps us meet our goals in this era of cost containment and data analytics," Fitzgerald says. "We can resolve in a few weeks what would have previously taken a team of people months or years, and save an estimated 80 percent of the expense."

SJH engaged a multi-stakeholder group of business and technology professionals to evaluate technologies, consider governance implications, and quantify benefits (see the sidebar below), as well as consider issues such as privacy and security. Daily and monthly dashboards and statistics are used to guide SJH patient identity governance activities, and integrate these foundational activities with the enterprise governance program.

Patient Identity Governance at St. Joseph Health

A robust proof of concept project was undertaken, applying governance principles, at St. Joseph Health based in Orange County, CA, in order to improve patient data matching. Key considerations for their cost analysis included:

- Traditional cost for data stewards ($21 to $25 per hour)
- Volume of tasks data stewards can typically resolve (50 to 100 per day)
- Estimated cost per task resolved by human intervention ($3.11 based upon rate of $25 per hour)
- Auto-stewardship estimated costs of $0.40 to $0.75 per task
- Financial savings of up to 80 percent per task can be achieved using automated tools, exclusive of management and overhead

Beyond the readily quantifiable savings, the value of having accurate, linked data to support consumer engagement and population health activities is immeasurable at this juncture.

Innovating with a Proprietary, Unique Patient Identifier

Provider organizations and payers are continuing to innovate in unique patient identifications. This may take the form of real-time or batch queries for a unique identifier that is held by a third party, frequently a healthcare business unit of a credit bureau. These unique identifiers are not universal, but rather proprietary to the vendor and their customers.

Michael Skvarenina, CIO at Holy Name Medical Center in New Jersey, says that provider organizations already do considerable business with credit bureaus for eligibility and credit inquiries, so extending that business model for patient identification is logical. "And, these organizations have already addressed with the credit bureau the common concerns of privacy, security, and trust," Skvarenina says. "A unique identifier created and maintained by a credit bureau can yield tremendous benefit particularly since their data is constantly updated, thus providing most current demographics for population health and prevention services, as well as data warehouses."

Innovation and Governance: A Natural Partnership

Innovation in patient identity should go hand in hand with existing or new data governance activities.

Innovation in patient identity should also naturally involve multi-stakeholder data governance councils or workgroups who may first consider a variety of strategic information governance perspectives, as sampled below, in addition to how innovation will be applied. This group should ask:

- Who owns the responsibility for the integrity of each identifier, particularly if a new identifier is created?

- If existing data is augmented with new external data, how is the new data integrated, and lifecycle of the new data managed?
- What are the acceptable uses for the identifiers in consideration of HIPAA personally identifiable information (PII), HIPAA privacy, and HIPAA compliance?
- How can/should an organization incorporate the patient identity technology with the long-standing human data steward work to ensure an enterprise compliance and governance perspective?

As organizations evaluate innovation for patient identity, it's important to elevate the discussion to a data or information governance realm.

Organization and healthcare professionals are understandably cautious in applying innovation to this long-standing problem, as the consequences of mismatching records can be profound. But this caution is not a reason to do nothing and wait for a silver bullet that some believe will come with a national healthcare identifier. Rather, when considering innovation, be prudent in assessing the potential risk, impact, and benefits. Consider a few basic steps when exploring the application of innovation:

- Innovation discussions should not only be a technology discussion, they should involve people, process, and technology and support data governance.
- Get business people involved if they aren't already. The multi-stakeholder group must include patient access/registration, health information management, and the technology team, and could include data users such as care coordination, analytics, and others undertaking data-oriented transformative activities.
- Consider whether you should or can apply innovation to data creation (patient access or registration) processes, along with data governance through standardization of procedures, processes, and data fields.
- Build a sample database to support the proof of concept/proof of technology (POC/POT), but don't think you can adequately test with 10,000 or 100,000 records. As the neural network discussion illustrates, using more data to test will produce stronger results and greater efficiency.
- Don't forget about the impact to downstream systems if auto-stewardship creates a high volume of resolved tasks.
- Incorporate identity data goals into an organizational data governance program.

As the world becomes more digitized, the authors hope that the innovations and guidance shared in this article inspire HIM professionals to apply innovation to the challenge of achieving accurate patient identification. All components of the vast healthcare ecosystem will benefit as a result of these efforts, including each of us as healthcare consumers.

Notes

[1] Office of the National Coordinator for Health IT. "Connecting Health and Care for the Nation: A Shared Nationwide Interoperability Roadmap."

[2] Gillick, Dan. "A Resurgence of Neural Networks in Machine Learning." Berkeley School of Information blog post. November 21, 2013.

[3] Jones, Graham. "The Resurgence of Neural Networks." LinkedIn blog post. July 27, 2015.

References

Butler, Mary. "Congress Passes the 21st Century Cures Act, Impacting Health Record Privacy, Documentation and Exchange." *Journal of AHIMA* website. December 7, 2016.

College of Healthcare Information Management Executives. "CHIME National Patient ID Challenge."

Fernandes, Lorraine M. et al. "Losing the Match Game: Study Reveals Gaps in HIM's Patient Identity Integrity Practices." *Journal of AHIMA* 87, no. 10 (October 2016): 39-47.

Fernandes, Lorraine and Michele O'Connor. "Accurate Patient Identification—A Global Challenge." *Perspectives in Health Information Management*. 2015.

Hall, Susan. "How New York RHIO Tackles Patient Matching." *FierceHealthcare*. December 5, 2016.

Le, Quoc V., Navdeep Jaitly, and Geoffrey E. Hinton. "A Simple Way to Initialize Recurrent Networks of Rectified Linear Units." Cornell University Library. April 7, 2015.

Office of the National Coordinator for Health IT. "Patient Identification and Matching Final Report." February 7, 2014.

Weber, Gerald. "Achieving a Patient Unit Record Within Electronic Record Systems." March 1995.

Weber, Gerald I. and Max G. Arellano. "Issues in Identification and Linkage of Patient Records Across an Integrated Delivery System." *Journal of Healthcare Information Management* 12, no. 3 (Fall 1998).

Lorraine Fernandes (lorraine@fhci.biz) is principal with Fernandes Healthcare Insights, and president-elect of the International Federation of Health Information Management Associations (IFHIMA). Jim Burke (jburke@himformatics.com) is EMPI and HIE practice lead at Himformatics. Michele O'Connor (Michele.OConnor@Collibra.com) is services manager, North America at Collibra.

Article citation:
Fernandes, Lorraine M.; Burke, Jim; O'Connor, Michele. "Applying Innovation to the Patient Identification Challenge" *Journal of AHIMA* 88, no.8 (August 2017): 26-29.

Driving the Power of Knowledge