Patient Matching within a Health Information Exchange

by Tim Godlove, PhD, and Adrian W. Ball, MSc, PMP

Abstract

The purpose of this article is to describe the patient matching problems resulting from the Nationwide Health Information Network's automated patient discovery specification and propose a more effective and secure approach for patient matching between health information organizations participating in a health information exchange. This proposed approach would allow the patient to match his or her identity between a health information organization's electronic health records (EHRs) at the same time the patient identifies which EHR data he or she consents to share between organizations. The patient's EHR username/password combination would be the credential used to establish and maintain health information exchange identity and consent data. The software developed to support this approach (e.g., an EHR health information exchange module) could also allow a patient to see what health information was shared when and with whom.

Keywords: electronic health record, health information exchange, health information organizations, data exchange, patient matching

Introduction

The electronic exchange of health information, commonly referred to as health information exchange (HIE), plays a vital and central role in delivering coordinated, accountable, patient-centered care that is both less expensive and of higher quality. HIE between two health information organizations (HIOs) has two key prerequisites. The first is confirming that the requesting and responding HIOs' patients are one and the same. The second is confirming that the patient has consented to the exchange of the health information between the HIOs. Many of today's HIEs, including the Virtual Lifetime Electronic Record (VLER) of the Department of Defense (DoD) and the Department of Veterans Affairs (VA), are implemented using the Nationwide Health Information Network (NwHIN) specifications. The NwHIN specifications describe an automated patient discovery process for confirming patient identity and a manual process for collecting patient consent. However, standardization issues with the data used by the NwHIN's automated patient discovery process commonly result in false non-matches and even false matches, both of which can cause major patient safety issues. Incorrectly mismatching or matching a patient as part of the HIE process may also have privacy and security implications. One way to resolve this issue to have patients confirm their HIO identity when they provide their consent to exchange data between HIOs.

As emerging new delivery systems move more aggressively toward sharing identifiable health information across disparate settings, concerns about historically suboptimal levels of accuracy in matching patients to their health information are exacerbated by poor data quality and incomplete data collection. The importance of the ability to link data is best put by Scheuren (1997): "Record linkage can aid a society in achieving advances in the well-being of its citizens."[1]

Patient Matching

Patient matching is a relatively new process used by private and public healthcare institutions to detect common denominators in electronic health records (EHRs) on the basis of personal traits such as demographics, geographies, and histories. It is used to match patients with medical records when the records come from a variety of sources. That way, the patient can be identified by the records, instead of the other way around. The data may be retrieved from patient records at different hospitals, doctors' offices, and governmental databases. While patient matching applications have benefits such as the ability to identify best practices and to synthesize treatment, they also raise serious concerns about issues such as false matches, false non-matches, privacy rights, and patient consent.

As noted previously, two fundamental prerequisites for HIE between HIOs are confirming that the requesting and responding HIOs' patients are one and the same and confirming that the patient has consented to exchange their data. This prerequisite data must be stored and maintained within the HIE application. Figure 1 illustrates HIE between three hypothetical HIOs (Acme HIO, A1 HIO, and HIO-R-Us). The HIE system has confirmed that the patients in question are one and the same (i.e., John Doe in the Acme HIO [Patient ID 123] is John Doe in the A1 HIO [Patient ID 789] and is also John Doe in the HIO-R-Us HIO [Patient ID 456]) and has captured this patient match in the HIOs' HIE metadata stores. The HIE system has also confirmed what type of EHR data the patient has consented to exchange (e.g., John Doe has consented to exchange his Acme HIO lab results and prescriptions with HIO-R-Us and A1 HIO, respectively) and has confirmed that the HIOs have captured this consent information in their respective HIE metadata stores. Figure 1 illustrates the results of automated patient matching based on patient attributes (i.e., patient name, gender, and DOB) contained within each HIO's EHR. This patient matching approach is described by the NwHIN patient discovery specification.

Issues of Concern

The patient matching process is as controversial as any medical issue that involves potential invasions of privacy. It is not the matching process itself that is at the core of ethical concerns, but the overall ability to access the data. What makes this issue particularly sensitive is that in addition to the potential privacy violations, the process has a high probability of error. False matches, in which matches that should not have been made are made, and false non-matches, in which matches that should have been made are not made, are both dangerous and daunting. Until errors can be reduced and confidentiality (allowing access only to meet medical treatment needs and not for marketing purposes) and voluntary consent become a part of the process, patient matching will remain a controversial and volatile issue in the healthcare industry and among the general public.

Patient matching is a challenge for the healthcare industry as a whole and for HIOs in particular because of the need to ensure accuracy across organizations. The Office of the National Coordinator for Health Information Technology (ONC) has provided funding for a number of health information technology (IT) programs, including the development of the NwHIN, which is "a set of standards, services, and policies that enable the secure exchange of health information over the Internet."[2]

Patient Discovery: The NwHIN Standard for Patient Matching

Patient Discovery is the NwHIN standard for HIE patient matching. Prior to exchanging patient-specific data, HIOs need to confirm that they are dealing with the same patient. The NwHIN patient discovery specification dictates how HIOs will locate and identify patient information that resides on another HIO on the NwHIN.

The North Carolina Healthcare Information and Communications Alliance describes the NwHIN patient discovery process as follows: "The initiating HIO enters all the demographic data and local identifiers that can be shared about a patient. The responding HIO matches the demographics and identifiers. If a single match is found that is considered highly reliable, it is returned to the initiator, along with its demographic details and identifiers. If no patient match is found then the responder sends an empty response to the initiator, indicating that this patient is not known at this HIO. If a highly reliable match cannot be identified, an ambiguous response is returned."[3] At this point, the NwHIN patient discovery specification requires additional information or the use of a manual process to finish the patient matching process. The NwHIN patient discovery specification "is designed to avoid false matches at all costs."[4]

The NwHIN patient discovery specification dictates that patient demographic traits be exchanged between HIOs and that each HIO is to respond with a match or an empty reply. When a match is found, the responder sends back the patient's demographic traits and patient ID. In a study of a NwHIN-based HIE, Bouhaddou et al. (2012) note that this process "enables the sender to validate the match using the demographics received" and "is considered a 'no risk' approach. However, the patient match success rate is far less than ideal."[5] Bouhaddou et al. note that as of May 9, 2012, only 23,611, or 53 percent, of veterans who consented to share their data were successfully matched by the VA's NwHIN-based HIE with the DoD and several private-sector HIOs and that "the failures were mainly due to lack of accurate, standardized, complete data. In other words, without a complete set of the primary identifiers (i.e., first name, last name, middle name/initial, gender, date of birth, and social security number), it is unlikely to match a patient. Other reasons include partial coverage of opted in Veterans population, requirement for a second authorization, and difference in patient matching algorithms."[6]

Bouhaddou et al. further note that "other insights related to Identity, Privacy, and Consent Management gained from [this NwHIN-based HIE] include" the following:[7]

1. "When onboarding a new [HIO] to NwHIN that has millions of patients, creating initial [patient matches] is challenging. [Patient discovery] as a broadcast-out model is difficult to scale for a future NwHIN that supports hundreds of HIEs."[8]
2. The NwHIN patient discovery "specifications do not provide clear guidance on how to keep the patient [matches] up-to-date when there are patient ID changes (e.g., marriage), merges, etc."[9]
3. "Inclusion of the full Social Security Number (SSN) in the demographic traits is necessary for any reasonable level of matching success. Some organizations or states do not exchange SSNs or only exchange the last four digits, which makes it unlikely a unique match can be achieved."[10]
4. "Lack of a 'common/standard' consent model is a barrier. Patients may have to sign multiple consents, one for each HIO, before their data are shared."[11]
5. "Specification would benefit from more clear definitions of the . . . attributes [of the data to be exchanged] and more complete . . . samples of the permitted values."[12]

The Risks Associated with Exchanging Health Data on the Basis of a False Match

The major risk involved in using automated deterministic or probabilistic patient matching algorithms like the NwHIN's patient discovery is the occurrence of false matches. According to the Bipartisan Policy Center, "Error rates, which average eight percent and can range up to 20 percent—can result in sub-optimal care and medical errors," and "incorrectly matching a patient to a health record may also have privacy and security implications, such as wrongful disclosure—in addition to treatment based on another patient's health information."[13]

The two main reasons for false matches are human input error (e.g., misspellings, number reversals) and lack of standardization. Human error is natural, but it can be reduced through improvement of the other area contributing to false matches, that is, standardization.[14] According to the Bipartisan Policy Center, "The lack of standardization in the data attributes or fields used for matching, the information contained in those fields, and methods used, results in increased error rates as well as significant burden and cost within the health care system."[15]

Data exchange capabilities are a critical issue in developing national standardization. Developing compatible software and hardware that interface with one another on a national level is a monumental undertaking in terms of both logistics and cooperation. According to Stead et al. (2005), "Since the United States does not have a national patient identifier, data interchange begins with a mechanism—such as a master patient index or its equivalent function—to determine which records relate to a single person. . . . A mechanism is needed to authorize access to data by a person at a site other than the place it originated. Such a service, therefore, is as much about governance, trust, and common terminology as it is about technology."[16]

Synthesizing regulations and standards is as big of a challenge as is synchronizing software and hardware. A major part of the problem stems from a lack of compliance with regulation. A 2010 Identity Force survey, involving more than 200 compliance experts across the United States, revealed that, despite increases in laws and regulations relating to compliance with security standards, the number of security violation incidents in the healthcare industry continues to rise.[17]

In 2013, the ONC released the findings of an extensive study on patient matching that involved "more than 50 large health systems and health IT software developers."[18] The study had two primary objectives: "to define common features that achieve high positive match rates across different systems; and to define the processes and best practices that are most effective to support high matching rates."[19] The recommendations that were derived from the findings of the study centered on standardization. In particular, the ONC recommended that the following information be standardized in every HIE transaction:

- Current and past addresses
- Date of birth
- Full name
- Gender
- Phone numbers[20]

Other recommendations made by the ONC in its report included the following:

- "Additional data attributes to improve patient matching should be studied."
- "An open source algorithm should be developed to test and build patient matching capabilities."
- "Certified electronic health record systems should be required to generate and provide reports that detail possible duplicate patient records."
- "EHR certification criteria should include the ability to capture patient identifying attributes."[21]

The report also recommended development of the following:

- "Formal best practices for patient matching and data governance."
- "Policies to encourage consumers to keep their health information accurate and up to date."
- Educational and training materials for verifying patient data attributes."[22]

Most hospitals and other healthcare facilities are in need of more streamlined methods of communication and data sharing with other facilities and government agencies. Progress is being made, but a strategically designed integration strategy in which systems and interactions are standardized is still needed. Standardization will increase cost efficiency and decrease the potential for errors. However, it will not do much to help with the second major area of concern regarding patient matching, which is privacy.

Privacy Issues Related to National Patient Identifier–based Patient Matching

The main objective of patient matching is to provide more synergistic, comprehensive treatment and care to patients, but a secondary goal is to exchange and sell patient information for marketing endeavors that lead to monetary gain. This secondary goal is one of the primary risks associated with patient matching because it destroys expectations of privacy and security. Most patients do not mind other doctors' being able to view their records for medical treatment purposes, but the idea of drug companies viewing the records, just so they can send advertisements, does not sit well with many people. According to Peel (2013), "'Patient matching' is a method of involuntary, hidden surveillance, much like the NSA's surveillance of phone records and metadata. It enables thousands of hidden third parties to collect and aggregate our personal health data from many places without our knowledge or consent."[23]

The increased use of EHRs has given rise to a host of concerns regarding patient privacy and the confidentiality of patient data. One proposed, yet controversial solution to the problems associated with electronic health data exchange is the Unique Patient Identifier (UPI). The Health Insurance Portability and Accountability Act of 1996 (HIPAA) supported the development of a UPI. However, in 1999, Congress took countermeasures to prevent the creation of UPIs because of widespread concerns about privacy and confidentiality. Congress passed a law that prohibited the US Department of Health and Human Services from allocating any of its budget toward UPI development unless Congress approved it beforehand. Fifteen years later, this restriction is still active.[24]

Public perceptions of UPI development vary, but almost everyone agrees on a couple of points. The first is that although UPIs would be helpful, they will not be a cure-all for the issues that plague patient matching. The second is that patients should have to sign a consent form allowing for an identifier to be used only with their express permission. To make UPI use mandatory without gaining consent from patients brings to light numerous potential ethical violations, and possibly even legal ones.[25]

This understanding has led to proposals for a voluntary private-sector option known as the Voluntary Universal Healthcare Identifier (VUHID). This voluntary approach to assigning patients their own unique ID numbers eclipses the majority of objections that people have vocalized regarding mandatory national patient identifiers. According to Paxton (2009), "VUHID is based on two standards developed by ASTM (originally the American Society of Testing and Materials) and ANSI (American National Standards Institute). The system's goal is to make unique health care identifiers available at nominal cost to individuals who want one. But more than that, VUHID promises to shield patient privacy by using two categories of the identifier: an open identifier for information a person wants to have known to all of his or her care providers, and multiple private identifiers for medical information a person wants to keep private."[26] This solution seems to offer an ideal compromise. However, there is still a fair amount of objection to the concept, perhaps based on the principle alone.

While a blanket solution does not seem to be available, the Bipartisan Policy Center (2012) recommends that the following measures should be taken to ensure the accuracy, efficiency, and ethicality of patient matching applications such as UPIs:[27]

1. "Standardize Matching Methods. . . . This includes standardizing data fields, definitions and validation methods designed to improve the accuracy and the quality of the information gathered from patients."[28]
2. "Standardize Policies. . . including those related to the establishment of acceptable benchmarks or rates of error in matching."[29]
3. "Share Lessons Learned and Best Practices. . . regarding technology, human resources, workflow and policy. . . . More transparency in disclosing accuracy rates will facilitate assessment of methods and also promote improvement."[30]
4. "Collectively Organize and Support the Adoption of Shared Services. Common principles, policies, standards, and methods for matching patient data will facilitate the sharing of services for matching across many organizations, promoting standardization, improving results and producing economies of scale."[31]

Consent is without a doubt one of the biggest issues related to patient matching applications, in terms of both practical issues and ethical concerns. In general, people do not favor the concept of strangers being able to look at their private medical data without consent or strangers using or distributing private medical data however they please, without the patient's knowledge or consent. However, the voluntary model has failed to gain momentum, primarily because interested parties are convinced that most people will staunchly refuse to provide their consent.

The Need to Collect Patient Consent as Part of the Matching Process

Consent, as part of the patient matching process, encompasses issues related to permission, access, and transparency. According to Peel (2013), "Today, the nation's sensitive health records are exchanged by hundreds of hidden users without meaningful informed consent. Health technology systems violate our federal rights to see who used our data and why. Despite the federal right to an Accounting of Disclosures (AODs)—the lists of who accessed our health data and why—technology systems violate this right to accountability and transparency."[32]

For people to accept the idea of patient matching and national patient identifiers, they need to know that their involvement in the process matters. They need to know that nothing can be done without their express permission. They need to feel empowered in the decision-making process regarding who is allowed to access their personal medical information and who is not. They need to know that they will have access to the matching data so that they can personally verify whether the data is accurate or is in error. Until these needs are fulfilled, the national acceptance of UPIs is going to remain highly unlikely.

Solution

In the absence of a national patient identifier, the following solution is proposed to avoid the patient matching problems resulting from the NwHIN's automated patient discovery specification. Have the patient match his or her identity across HIO EHRs at the same time the patient identifies which EHR data he or she consents to share between HIOs. Figure 2 illustrates HIE between three HIOs (Acme HIO, A1 HIO, and HIO-R-Us). The HIE application would collect a patient's Acme HIO, A1 HIO, and HIO-R-Us EHR login credentials (username and password) and then validate those credentials to establish a patient match between these HIOs. As part of this process, the patient would also identify what EHR information he or she consents to share as part of the HIE (e.g., John Doe has consented to share his Acme HIO EHR lab results and prescriptions with HIO-R-Us and A1 HIO, respectively). As in the model described in Figure 1, this patient match and consent data is captured in the HIE metadata stores. This patient matching and consent approach assumes that patients have login credentials for the HIO's EHRs to which they belong.

Conclusion

As noted previously, the patient matching process is a controversial problem that involves possible invasions of privacy and a high probability of error. The NwHIN's automated patient discovery specification has the potential for matching errors that may result in patient safety problems and privacy and security implications.

The alternative approach described in this article would allow the patient to match his or her identity between one HIO's EHR at the same time that the patient identifies which EHR data he or she consents to share between HIOs. The patient's EHR username/password combination would be the credential used to establish and maintain HIE identity and consent data. The software developed to support this approach (e.g., an EHR HIE module) could also allow a patient to see what health information was shared when and with whom.

Until such an approach can be implemented and confidentiality and voluntary consent become an integral part of the process, patient matching will continue to be a source of concern both within the healthcare industry and among the general public.

Tim Godlove, PhD, is the director of information technology, business requirements, and administrative service at the National Cemetery Administration in the Department of Veterans Affairs in Arlington, VA.

Adrian W. Ball, MSc, PMP, is the senior director of software engineering at TurningPoint Global Solutions in Rockville, MD.

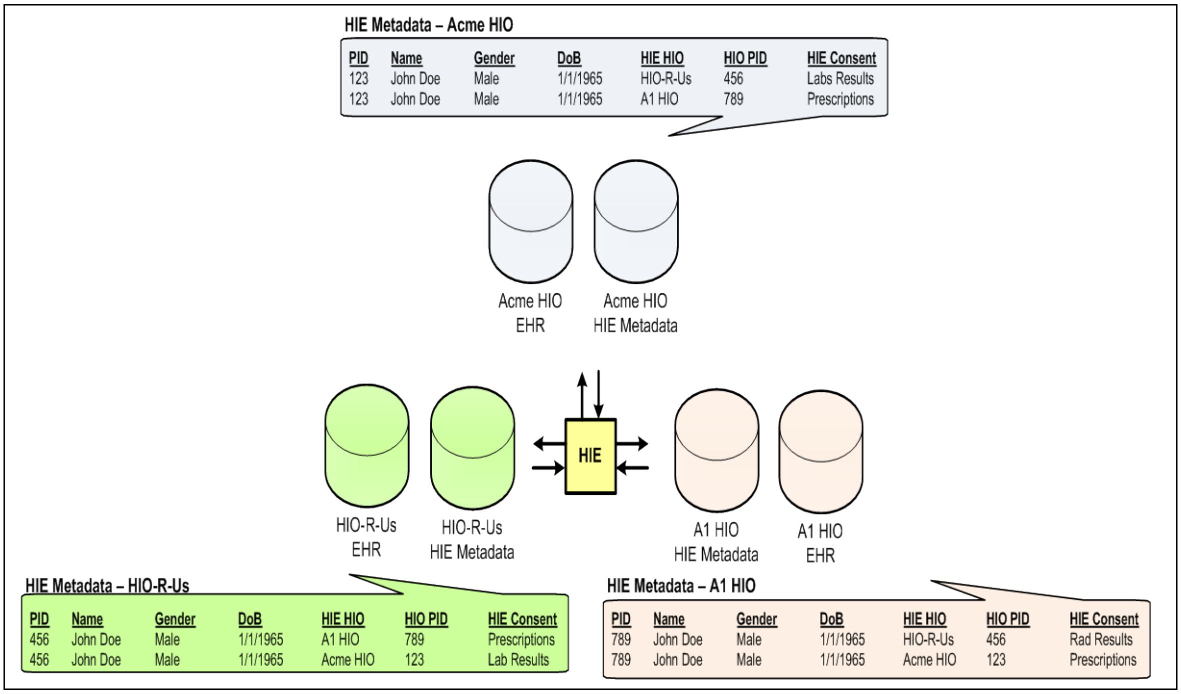
Notes

1. Scheuren, F. “Linking Health Records: Human Rights Concerns.” In *Record Linkage Techniques–1997: Proceedings of an International Workshop and Exposition*. Washington, DC: National Academy Press, 1999, 404–26.
2. HealthIT.gov. “Nationwide Health Information Network (NwHIN).” 2013. Available at http://www.healthit.gov/policy-researchers-implementers/nationwide-health-information-network-nwhin.
3. North Carolina Healthcare Information and Communications Alliance (NCHICA). “NwHIN Specification Preview.” Available at http://nchica.org/HIT_HIE/NHIN2/NHIN1209.htm.
4. Ibid.
5. Bouhaddou, O., J. Bennett, J. Teal, M. Pugh, M. Sands, F. Fontaine, M. Swall, S. Dhar, T. Mallia, B. Morgan, and T. Cromwell. “Toward a Virtual Lifetime Electronic Record: The Department of Veterans Affairs Experience with the Nationwide Health Information Network.” *AMIA Annual Symposium Proceedings* (2012): 51–60.
6. Ibid.
7. Ibid.
8. Ibid.
9. Ibid.
10. Ibid.
11. Ibid.
12. Ibid.
13. Bipartisan Policy Center. *Challenges and Strategies for Accurately Matching Patients to Their Health Data*. June 2012. Available at http://www.redwoodmednet.org/projects/events/20120719/Accurately-Matching-Patients-to-Their-Health-Data.pdf.
14. Li, X., and C. Shen. “Linkage of Patient Records from Disparate Sources.” *Statistical Methods in Medical Research* 22, no. 1 (2013): 31–38.
15. Bipartisan Policy Center. *Challenges and Strategies for Accurately Matching Patients to Their Health Data*.
16. Stead, W. W., B. J. Kelly, and R. M. Kolodner. “Achievable Steps toward Building a National Health Information Infrastructure in the United States.” *Journal of the American Medical Informatics Association* 12, no. 2 (2005): 113–20.
17. BusinessWire. “Delayed Compliance with New Regulations Has Increased Data Breaches and Medical Identity Theft in U.S. Hospitals.” April 20, 2010. Available at http://www.businesswire.com/news/home/20100420006272/en/Delayed-Compliance-Regulations-Increased-Data-Breaches-Medical#.VOywsfnF_jA.
18. iHealthBeat. “ONC Releases Findings from Patient Data Matching Study.” December 17, 2013. Available at http://www.ihealthbeat.org/articles/2013/12/17/onc-releases-findings-from-patient-data-matching-study.
19. Ibid.
20. Ibid.
21. Ibid.
22. Ibid.

23. Peel, D. C. "Patient Identification and Matching Initial Findings." Patient Privacy Rights. December 16, 2013. Available at http://patientprivacyrights.org/wp-content/uploads/2013/12/PPR-Patient-Matching-Testimony-for-12.16.13.pdf.
24. Bipartisan Policy Center. *Challenges and Strategies for Accurately Matching Patients to Their Health Data*.
25. Ibid.
26. Paxton, A. "National Patient ID: Could a Voluntary System Fill the Gap?" College of American Pathologists. November 2009. Available at http://www.cap.org/apps/portlets/contentViewer/show.do?printFriendly=true&contentReference=cap_today%2F1109%2F1109j_national_id.html.
27. Bipartisan Policy Center. *Challenges and Strategies for Accurately Matching Patients to Their Health Data*.
28. Ibid.
29. Ibid.
30. Ibid.
31. Ibid.
32. Peel, D. C. "Patient Identification and Matching Initial Findings."

Figure 1

HIE Patient Matching and Consent Metadata



HIE Metadata – Acme HIO

PID	Name	Gender	DoB	HIE HIO	HIO PID	HIE Consent
123	John Doe	Male	1/1/1965	HIO-R-Us	456	Labs Results
123	John Doe	Male	1/1/1965	A1 HIO	789	Prescriptions

HIE Metadata – HIO-R-Us

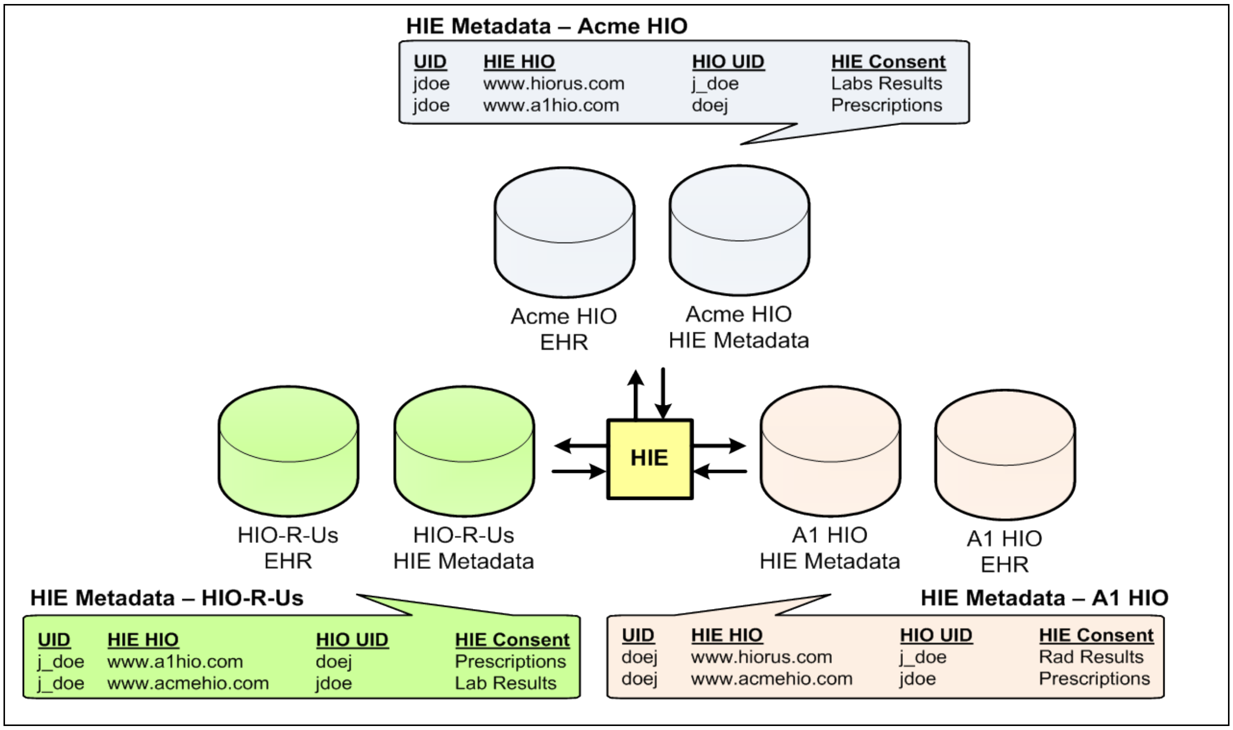
PID	Name	Gender	DoB	HIE HIO	HIO PID	HIE Consent
456	John Doe	Male	1/1/1965	A1 HIO	789	Prescriptions
456	John Doe	Male	1/1/1965	Acme HIO	123	Lab Results

HIE Metadata – A1 HIO

PID	Name	Gender	DoB	HIE HIO	HIO PID	HIE Consent
789	John Doe	Male	1/1/1965	HIO-R-Us	456	Rad Results
789	John Doe	Male	1/1/1965	Acme HIO	123	Prescriptions

Figure 2

Health Information Exchange (HIE) Patient Identification and Consent Metadata



HIE Metadata – Acme HIO

UID	HIE HIO	HIO UID	HIE Consent
jdoe	www.hiorus.com	j_doe	Labs Results
jdoe	www.a1hio.com	doej	Prescriptions

HIE Metadata – HIO-R-Us

UID	HIE HIO	HIO UID	HIE Consent
j_doe	www.a1hio.com	doej	Prescriptions
j_doe	www.acmehio.com	jdoe	Lab Results

HIE Metadata – A1 HIO

UID	HIE HIO	HIO UID	HIE Consent
doej	www.hiorus.com	j_doe	Rad Results
doej	www.acmehio.com	jdoe	Prescriptions