A Framework for Cross-Organizational Patient Identity Management

TABLE OF CONTENTS

Table of Figures

Table of Tables

November 10, 2015

Dear Reviewer,

We in the healthcare industry are in constant pursuit of "better."

Better data. Better clinical decisions. Better care coordination. Better patient experience. Better outcomes. There is no perfection possible for us, because we can always strive to do better.

To support this constantly moving goal line of "better," healthcare information technology is evolving rapidly. With the support of the federal government, there is wide-spread adoption of electronic healthcare records (EHRs) in progress. For many, the focus is now shifting to using EHRs in better ways.

For EHRs to deliver on the promise of better healthcare, we need to ensure patient data are sent and received easily among providers across disparate systems. These shared records must be accurate and useable. Patient matching is critical to the successful sharing of patient records, but the eHealth Exchange, the nation's largest health data sharing network, and many others have observed patient data matching is an ongoing obstacle to seamless information exchange between organizations.

Without a national patient ID system, we must focus on optimizing the current patient matching strategies. The Sequoia Project and the Care Connectivity Consortium have researched best practice patient matching principles for the last two years. In a joint case study with Intermountain Healthcare, we've found cross-organizational matching deficiencies and opportunities for improvement among even our most progressive institutions.

The proposed framework for patient identity management presented in the following pages includes actionable best practices and a maturity model roadmap for future growth and improvement in nationwide patient matching strategies.

Please help us "better" this framework.

We value the input from across the healthcare continuum and government, and we invite everyone to contribute to our proposed framework through our public comment process. Your thoughtful analysis and feedback are key to bettering this proposed framework, and ultimately, bettering our national patient matching capabilities.

Yours in good health,

Mariann Yeager   &   Michael Matthews
CEO, The Sequoia Project   CEO, MedVirginia

The Sequoia Project

The Sequoia Project is a non-profit 501c3 chartered to advance implementation of secure, interoperable nationwide health data sharing. The Sequoia Project supports health IT interoperability initiatives, most notably: eHealth Exchange, a rapidly growing community of exchange partners who share information under a common trust framework and a common set of rules; and Carequality, a public-private collaborative effort to build consensus among existing data sharing networks regarding technical specifications and best practices.

The Care Connectivity Consortium

The Sequoia Project has teamed with the Care Connectivity Consortium (CCC) to address critical challenges to effective health information exchange. As pioneers, they share the goal of making data sharing a reality on a national scale to support health care providers and patients with easy, fast access to information at the point of care. The CCC, representing Geisinger Health System, Group Health Cooperative, Intermountain Healthcare, Kaiser Permanente, Mayo Clinic and OCHIN, works to develop and incubate new capabilities for adoption by its initiatives and the broader health IT community. The CCC are piloting a set of operational shared services focused on patient identity management, record location, and consent.

Contributing Authors

- Lead Author: Eric Heflin, Chief Technology Officer/Chief Information Officer, The Sequoia Project
- Shan He, Medical Informatics, Intermountain Health Care /Care Connectivity Consortium
- Kevin Isbell, Executive Director of Health Information Exchange, Kaiser Permanente/Care Connectivity Consortium
- Andy Kling, Director of IT, Geisinger Health System/Care Connectivity Consortium
- Chris Ross, Chief Information Officer, Mayo Clinic/Care Connectivity Consortium

- Seth Selkow, Director of CCC Program and HIE Engagement, Kaiser Permanente/Care Connectivity Consortium
- Sid Thornton, Medical Informatics, Intermountain Health Care/Care Connectivity Consortium

CHAPTER 1: A FRAMEWORK FOR PATIENT IDENTITY MANAGEMENT

Introduction

The healthcare industry is making significant progress towards technical interoperability, but continues to fall short of the promise of ubiquitous interoperable health data sharing. Until we can consistently send and receive accurate and useful patient data nationwide, we miss the opportunity to fully realize the documented benefits of seamless, interoperable health data sharing including improvements in clinical decision making and patient safety, business process improvement, and support for value-based payment. Among the remaining challenges to successful nationwide exchange are patient matching and identity management.

The inability to consistently and accurately match patient data creates a number of problems for physicians and other health care providers. With respect to patient safety and satisfaction, providers may have an incomplete view of a patient's medical

> Inadequate patient matching is impeding our ability to provide physicians with accurate, timely, and useful information

history, care may not be well coordinated with other providers treating the patient, patient records may be overlaid, unnecessary testing or improper treatment may be ordered, and patient confidence may be eroded. In addition, patient privacy preferences may not be honored across organizations without accurate matching.

With insufficient matching practices, providers may experience a number of clinical workflow inefficiencies that are costly. Those include prolonged troubleshooting to find the correct patient record, a reversion to manual telephone and fax information exchange workflows, ordering duplicate tests, and failure to detect and honor patient privacy preferences.

Development of a Framework for Patient Identity Management

To address these critical patient matching and identity management issues, The Sequoia Project, in collaboration with the Care Connectivity Consortium (CCC), is developing minimal acceptable cross-organizational patient matching rules, suggested matching traits, a framework for methodical improvement, and a maturity model serving as a roadmap for future growth and improvement.

The intent of this first draft framework is to solicit feedback on both a set of minimal acceptable cross-organizational patient matching principles and an identity management maturity model that will provide a method to evaluate, measure, and improve patient matching across organizational boundaries. These work products have been created in response to real-world, known issues from years of experience supporting large-scale, nationwide data sharing initiatives.

As an introduction to this work, we present our observations and a representative case study conducted by the Care Connectivity Consortium and Intermountain Healthcare.

Patient Privacy

The CCC and The Sequoia Project believe patient privacy should be at the center of patient identity management strategies. Specifically, we want to help advance the ability of patients to protect the confidentiality and integrity of their data, and to help patients stay aware of and in control of their data. We believe this includes: (1) allowing for anonymous or pseudonymous patient identities; (2) correct identification of patients so that their privacy preferences can be determined and honored; and (3) enabling correct matching of patients to their records (whether anonymous or identifiable). We have taken these objectives into account in this framework.

The Patient Identity Blind Spot

Correct, optimal, and safe patient matching has been a high priority for healthcare organizations since computers were introduced into those enterprises. And even after this multi-decade focus, patient matching is far from perfect. Many believe it will never be successful in the absence of a national identifier. This is described in the RAND Corporation's study, *IDENTITY CRISIS: An Examination of the Costs and Benefits of a Unique Patient Identifier for the U.S. Health Care System*. As a result we are left with

optimizing a less-than-perfect, complex, mission-critical system – a system upon which patients' lives depend. Correct identification of complete and accurate patient data is both a national and international priority.

This issue manifests in two broad forms:

- Organizations have an inability to detect shared patients within or among their organizations, known as a false negative, or;
- Organizations erroneously match the records of different patients, known as a false positive.

This Framework addresses these and other related difficult issues.

Your organization, like most many others, may have a "blind spot" in terms of patient matching. You likely have acceptable patient matching rates within your enterprise (i.e., your hospital or integrated delivery network). Here, your staff can identify problems, measure, apply fixes, re-measure, and continue to improve until problems are at an acceptable level for patient matching, consent issues, linking, merging, unlinking, and complex unmerging activities.

But patient matching across organizations is a very different problem. The vast majority of key factors influencing the correctness of patient matching are now out of your direct control, including:

- Default or temporary values;
- Data quality;
- Data completeness;
- Data field consistency;
- Software (vendors, update lifecycle, configuration);
- Vocabulary adoption and versioning;
- Consent, security, sensitive data sharing, and other policies;
- Research Institutional Review Board stipulations;
- Vastly different data characteristics;
- Human and system workflows (latency, variations, definitions, etc.);
- Corporate cultures (accepting "friendly" patient nick names vs. meticulous accuracy);

Patient matching across organizations is vastly different than matching inside an organization

- Data exchange latency;
- Vastly different scope of data (specialty practice vs. large integrated delivery network);
- Organizational size, resource allocation, project timelines, commitment, skill levels;
- Diagnostic capabilities;
- Change management;
- Vendor engagement, version updating strategy, staffing;
- Internal enterprise software architecture (presence/absence of an enterprise-wide active master patient index (MPI), use of multiple MPIs, different tolerances in terms of matching accuracy, different patient matching rules and algorithms, services levels/response times, etc.); and/or
- Legal jurisdictions and requirements (minors, reproductive health, etc.).

Resolution of patient identity issues are more daunting when they cross organizational lines. Such issues often involve six or more organizations (the two health information organizations, their two vendors, and often an intermediary such as a health

> Resolution of patient identity issues are more daunting when they cross organizational lines where they often involve six or more organizations.

information organization and their vendor). In such an environment, even mundane items such as scheduling cross-organizational working sessions often introduce days and weeks of delay in resolving each issue due to lack of availability of key personnel. In essence, health data sharing introduces dependencies upon these independent organizations, and intertwines the workflows of the organizations, where no single organization has direct control over the other. This plays heavily into cross-organizational diagnostics, manual fallback procedures when automated patient matching does not work, manual intervention to correct patient records, and manual intervention to gather consent.

More subtly, it is also a significant issue when determining the "truth." How can we measure the actual, predicted, and targeted patient matching behavior across organizations? Often that entails creating a manually validated subset of mutual patients to become a benchmark to measure patient matching performance. This is a significant effort.

The result: patient matching practices across organizations are inconsistent and often subpar, with match rates as low as 10-30%. In the next chapter, we present a case study of how one organization increased the cross-organizational patient match from only 10% to over 95%, including specific steps, avoidable missteps, and recommendations intended for application to your organization.

CHAPTER 2: CASE STUDY

In this chapter, we present a collaborative study by the Care Connective Consortium (CCC) and The Sequoia Project evaluating traits and processes for successful patient matching across organizations. This study is based on a live production pilot using CCC Services between Intermountain Healthcare and local exchange partners including Utah Health Information Network (UHIN).

Intermountain Healthcare is a not-for-profit health system based in Salt Lake City, Utah, with 22 hospitals, a broad range of clinics and services, about 1,400 employed primary care and secondary care physicians at more than 185 clinics in the Intermountain Medical Group, and health insurance plans from SelectHealth.

This case study illustrated the ability to increase patient match rates from as low as 10% to more than 95%.
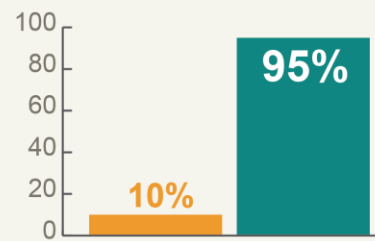
Intermountain's willingness to share their incredibly valuable knowledge gain so that the rest of the industry can build upon their work is laudable and provides an example for the industry in terms of "open sourcing" knowledge so we can all benefit from each other's experiences.

The Goal

Intermountain Healthcare was seeking to establish exchange of clinical information with two regional organizations as a preliminary step towards broader exchange. Intermountain had invested heavily in healthcare IT for many years, and frequently shares innovative ways to use IT with the global community. They are very highly regarded in the industry in terms of IT sophistication in the clinical domain. As such, the project was expected to achieve a reasonably high degree of success from the outset. Unfortunately, this proved not to be the case.

Case Study Executive Overview

Steps to Increase Patient Matching Rates

Performance Limit of demographic traits 90-95%

Unconstrained Demographics	Data cleaning, Normalization	Algorithmic refinement, Operational improvement	Pre-worked & reused correlations	Lessons Learned
	10-15%	60-70%	85-90%	95%+
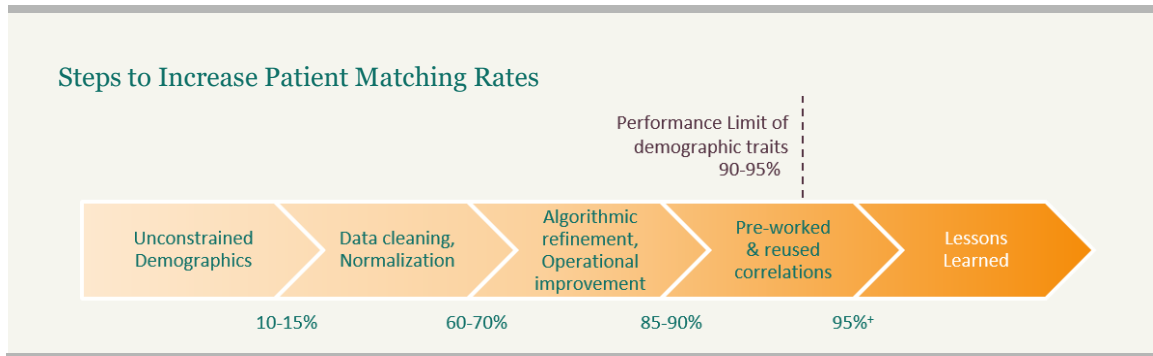
Demographics-based patient matching has inherent limitations in performance no matter how sophisticated the matching algorithm because demographic attributes by nature are not unique across individuals and because many demographics evolve over time as an individual traverses the healthcare community. Nonetheless, with proper data quality control and algorithmic adjustment, demographic-based patient matching can achieve mathematically promising matching rates around 90-95%. Unfortunately, diverse operational issues within healthcare data sharing networks often compromise inter-organizational patient record matching performance.

When the process of patient demographics collection is not governed among exchange partners, significant data quality issues can be introduced and the match rate can be as poor as 10-15%. Common data quality issues include missing information, typographical errors, misspellings, and transpositions. Simple process improvements such as data validity checking, normalization, and downstream data cleansing can increase patient matching rate to 60-70%.

Further improvements in matching rates among organizations accrue as healthcare data sharing network operational environments are refined to address challenges such as network timeouts, message encoding/decoding inconsistencies, synchronicity of patient consent, etc. Appropriate technical and workflow solutions can increase patient matching rates that approach the mathematical limit, which is typically about 90%.

In order to break through the inherent limitation of demographics-based patient matching, identity correlations of the fragile population should be proactively curated. Such correlations can be established by pre-working and subsequent reuse of identity correlations determined from human review and investigation of the problematic (fragile)

record pairs. Essentially, this cooperative approach allows patient matching based on reliable knowledge among disparate organizations and bypasses identity resolution based solely on demographic matching. As an additional benefit, the approach may compensate for minor lapses in operational rigor among the cooperating organizations. The matching rate can thus be increased to reach beyond 95%.

Exploration of definitive technologies based on immutable personal attributes, devices, or traits demonstrate the potential for perfect identity resolution but have not been broadly adopted across communities.

Next, we will review the step-by-step approach taken by Intermountain Healthcare as they share their process of optimizing patient matching with exchange partners.

Step 1: Small Sample Trial to Establish Baseline

Steps to Increase Patient Matching Rates

Unconstrained Demographics	Data cleaning, Normalization	Algorithmic refinement, Operational improvement	Pre-worked & reused correlations	Lessons Learned
10-15%	60-70%	85-90%	95%+	
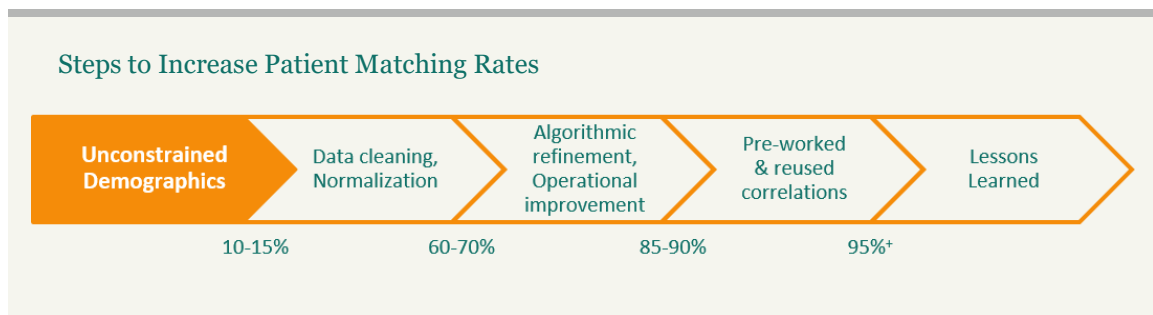
The first step in the project was to establish a baseline for patient matching success across organizational boundaries. Empirical testing was established with a small sample selection of 10,000 patients that were known to both Intermountain Health Care and one of its exchange partner. Given that all patients in this small sample were known to have been treated by both organizations, they expected that a large majority of the patients would be successfully matched. A "gold standard" (known correct) dataset composed of accurate matched patient pairs was established. This gold standard dataset was built by leveraging human-reviewed linked patient record pairs from pervious operational transactions. It included 340,000 pairs of linked patient demographic records. This important dataset established the benchmark upon which performance of various matching approaches could be accurately assessed.
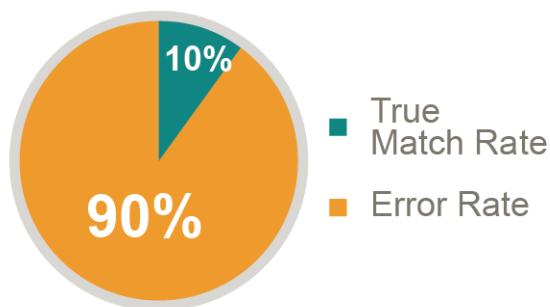
Figure 1: Initial Cross-Organizational Patient Match Error Rate

Surprisingly, the initial attempt resulted in only a 10% match rate. Even though this was a test that was not intended for production, the outcome of only achieving a 10% match rate was unexpected. **The sample data were fraught with data quality issues.**

Step 2: Trait Analysis
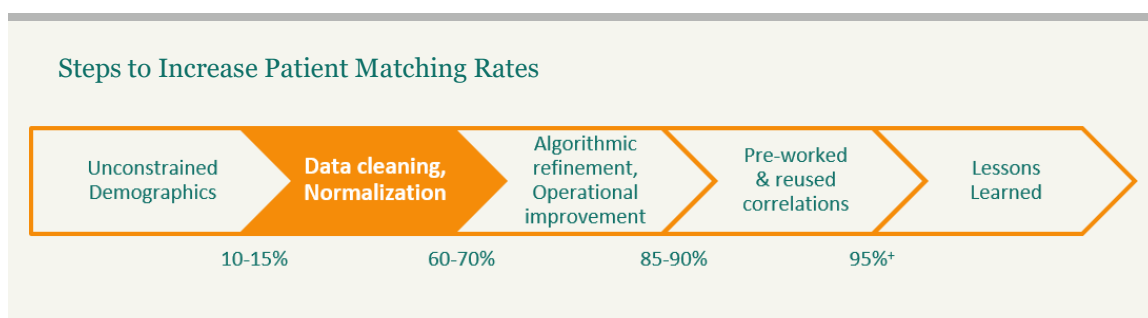
The next step for the organizations was to enhance the matching rate. They started by characterizing trait data to identify the identity attributes that contributed most to patient matching across these two organizations. This analysis was conducted using Intermountain's internal Enterprise Master Patient Index (EMPI) database that includes 6.6 million patient records. Several characteristics where analyzed to determine those traits most likely to be useful.

Completeness: At what rate is this trait captured and available?

Validity: Is this trait known to be correct? Patient demographics consisting of default or temporary values (e.g. "Baby Smith" for newborn's name) are complete but not valid.

Distinctiveness: Is the trait able to uniquely identify a person? For example, a trait such as sex (i.e. administrative gender) is not associated to a single

individual, whereas a trait such as an enterprise master patient index (EMPI) value is distinctive.

Comparability: Is the trait structured, coded (or numerical), or is it free text in string format? An address is an example of a relatively difficult to compare trait, whereas a social security number (SSN) can be easier to compare.

Stability: How much does the trait remain constant over a patient's lifetime? Although examples exist to the contrary, traits such as gender, birth date, and Social Security Number tend to be relatively consistent over time. Other traits, such as current address, tend to change relatively frequently.

The table on the next page shows the results of an analysis of potential traits and their suitability for use in patient matching.

Patient Attributes Analysis

Table 1: Patient Attributes Analysis

Attribute Name	Completeness	Validity	Distinctiveness	Comparability	Stability
EMPI	100%	--	100%	Very High	Very High
Last Name	99.85%	99.84%	5.1%	Medium	High
First Name	99.85%	99.33%	3.1%	Medium	High
Middle Name	60.54%	60.54%	2.6%	Medium	High
Suffix Name	0.08%	0.08%	0.08%	Medium	Medium
SSN	61.40%	60.92%	98.0%	High	High
Sex (Admin. Gender)	99.98%	99.98	0.00008%	High	High
Date of Birth	98.18%	97.38%	0.8%	High	Very High
Date of Death	3.36%	3.36%	3.4%	High	Very High
Street Address (1 or 2)	95.00%	94.61%	44.4%	Low	Low
City	94.84%	94.83%	0.8%	High	Low
State	94.81%	94.39%	0.8%	High	Low
Facility MRN	99.90%	99.90%	99.90%	High	Low
Postal Code	92.31%	92.0%	0.6%	High	Low
Primary Phone Number	90.68%	87.26%	51.6%	High	Medium
Work Phone Number	20.28%	19.79%	51.6%	High	Low
Ethnicity	25.25%	25.25%	0.0003%	High	Very High
Race	76.25%	76.25%	0.0001%	High	Very High

Items in Table 1 highlighted in yellow indicate desirable matching characteristics and are explored below. Items in red were identified as promising identifiers for future exploration.

The **EMPI**, which is the internal enterprise-wide unique patient identifier, value has many desirable characteristics. It is always internally available since the systems in this study required an EMPI value to be assigned as a prerequisite for all other clinical or data entry activities. It also should be valid, distinctive, very comparable, and very stable. However, the ability of an EMPI number to be used across organizations is problematic. Since the EMPI value is normally specific to one organization, it may not be accepted by the exchange partner. Alternatively, if a Master Patient Index (MPI) is **shared across** organizations, it can be a very valuable trait and is perhaps sufficient to establish high-confidence matching provided that demographic confirmation is also used to check the correctness of the link. Once this correctness has been confirmed, the linkage between patient identities is assumed.

However, the deployment of a cross-organizational MPI can be expensive and difficult from policy, legal, and technical perspectives. It becomes a new operational system that must be managed via a feed of demographic information from individual systems

> ... deployment of a cross-organizational MPI can be expensive and difficult from policy, legal, and technical perspectives.

to the shared MPI, or the shared MPI must be used to actively assign patient IDs in near-real-time as patients are initially entered into their respective systems. Additionally, in order to remain accurate, the EMPI must receive a constant feed of updates to the patient traits as they are corrected and/or change over time. The processes of achieving acceptable MPI accuracy levels, and diagnosing identified inaccuracies, can be a very time consuming, cross-organizational process. A cross-organizational shared MPI often receives a delayed feed, and in many cases, only receives incomplete information. The result is that a cross-organizational MPI can be of value, but it has significant limitations that do not exist when an MPI is used within a single organization. These limitations must be addressed in order to for a cross-organizational MPI to be successful.

Patients' first and last names also stand out in several regards. These traits are generally complete, valid, and stable. They are not, however, very distinct (5.1% and 3.1% respectively). In addition they are only moderately comparable largely due to spelling variations, inconsistent use of special characters, inconsistent use of middle name, the use of nicknames vs. formal given names, and software support of patient names with more than three components.

Patients' middle name is a very different trait in almost all regards (except for comparability and stability) than the patients' first and last names. The middle name is only present in the data set about 60% of the time and it is only valid about 60% of the time. It is also less distinctive, 2.6%, compared to first and last names.

The **patients' suffix name** can be quickly disregarded given its exceedingly low completeness, validity, and distinctiveness (0.08%).

The **SSN** looks like a much more promising trait, as it scores very highly for all criteria used for this assessment (including a 98% distinctiveness score). But the SSN is fraught with other challenges including fraud, medical and financial identity theft, sharing by multiple individuals, and more. These issues are compounded by the fact that some organizations require the use of SSNs for matching purposes, and even have made internal assumptions that the SSN will be provided. In contrast, other organizations have banned the use of SSNs, or only allow SSNs to be shared under limited conditions, such as only sharing the last 4 digits. The outcome of these SSN-related issues is that it remains a contentious trait. It holds promise, but is not in itself a solution. In addition, local policy makes it impossible, at the current time, to have cross-organizational patient matching that depends on this value being consistently supplied. The SSN trait remains a technical hurdle as well as an opportunity. One significant opportunity for improved use of SSNs will be explored later in this chapter.

Sex (more accurately referred to as administrative gender) is, as expected, largely complete, valid, comparable, and stable. However, also as expected, it is not distinctive (0.00008%).

Date of birth stands out in multiple ways, including completeness, validity, comparability, and very high stability. Its distinctiveness (0.8%) in this specific analysis, when combined with other traits, made it a useful trait. Overall, it was one of the most promising traits.

Street address looked promising from the perspective of completeness and validity. It also provided good (44.4%) distinctiveness. However, it was ranked low from the perspectives of comparability and stability. It may be of use for patient matching approaches that look at a patient's address history.

Postal code and **primary telephone number** also look promising when combined with other demographics. They are relatively complete and valid (above 80%) and they are easy to compare with minor normalization effort. Even though postal code by itself if not highly distinctive, the fact of it being numerical and collected by most organizations can make it an attractive trait, when combined with other traits, for patient matching.

Lessons learned from this analysis include:

1) More data do not necessarily mean better patient matching results. Depending on how sophisticated the matching algorithm is, traits with poor validity and comparability may cause a decrease in matching performance. Each organization should conduct a trait analysis on their internal patient population. The best combination of traits should be determined for each pair of exchange partners (or data sharing network).

2) Most patients (>90%) can be uniquely identified by a combination of common demographic data elements (e.g. name, date of birth, address, etc.) when available.

Step 3: Offline Algorithmic Performance Measurement and Refinement

Steps to Increase Patient Matching Rates

Unconstrained Demographics	Data cleaning, Normalization	Algorithmic refinement, Operational improvement	Pre-worked & reused correlations	Lessons Learned
10-15%	60-70%	85-90%	95%+	
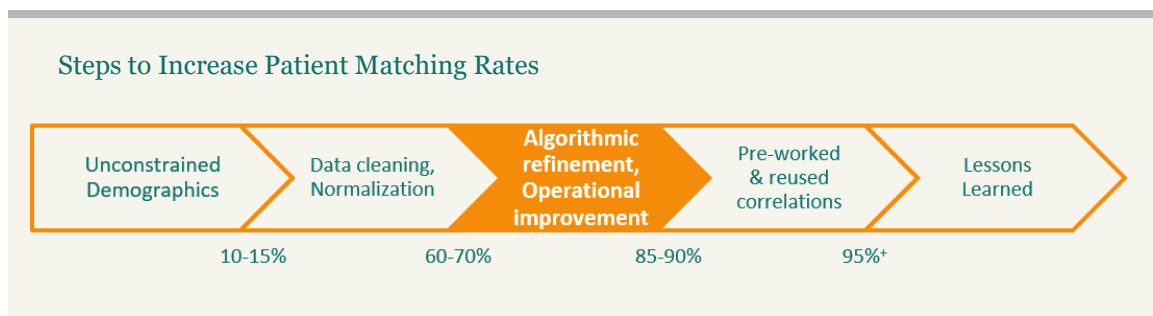
After identifying and analyzing potential traits to leverage in the development of more effective patient matching algorithms, the organizations next identified and implemented relatively easy, high-value improvements. **Missing data were gathered; inaccurate data were corrected.**

Default values (also known as temporary values) merit special consideration. In this context, we define a default value as a data item that is known to be fictitious due to lack of information. Default values are commonly employed when organizational policy or software limitations require certain fields to be supplied even if the correct value for that field cannot be ascertained at the time. A common example, is a hospital admission system which requires a patient name to be entered upon admission. If the patient's name is not known, staff are typically instructed to enter a value such as "Jane Doe." In this case, "Jane Doe" is the default value. Inside a given organization, default values are

often not harmful. But across organizational boundaries, default values can be harmful if not properly managed. Default values, when sent across organizational boundaries, can sometimes have the effect of "contaminating" the reciprocal organizations' MPI traits for that patient. Consequently, default values should never be exchanged across organizational boundaries.

For this case study, all known default values were inventoried and excluded from matching algorithms. Alternative name representations, such as nicknames or common misspellings, were accommodated during matching.

During this stage of the case study, it was observed that the patient identity data appears to become incorrect at a rate of 1% per month. It has been noted that similar data, such as mailing lists and provider directories, also have been found to age at this approximate rate.

> ... patient identity data appears to become incorrect at a rate of 1% per month.

During this stage, seven combinations of traits were assessed to determine their predicted success in terms of being complete and uniquely identifying patients. Table 2 was simplified by not breaking-out false positive, false negative, true positive, and true negative matching, selectivity and sensitivity but the below table is "directionally correct."

Analysis of Patient Trait Combinations

Table 2: Analysis of Patient Trait Combinations

Sequence	Combination of Traits	Completeness	Uniqueness
1	FN+LN+DoB	98.2%	95.7%
2	FN+LN+DoB+Sex	98.2%	95.9%
3	FN+LN+DoB+Sex+ZIP(first 5)	91.1%	99.2%
4	FN+LN+DoB+Sex+Phone	76.2%	99.5%
5	FN+LN+DoB+Sex+MN	59.9%	98.9%
6	FN+LN+DoB+Sex+MN(initial)	60.0%	97.7%
7	FN+LN+DoB+Sex+SSN(last 4)	61.9%	99.7%

The first combination explored (sequence 1) was that of **first name**, **last name**, and **date of birth**. This combination resulted in 98.2% completeness while relatively good in terms of completeness, was relatively poor in terms of uniquely identifying a given individual (95.7%).

The next combination reviewed (sequence 2) was **first name**, **last name**, **date of birth**, and **administrative gender**. This had no discernable impact to completeness (compared to sequence 1) with a near-trivial improvement on uniqueness from 95.7% to 95.9%. This was regarded as an insignificant improvement vs sequence 1.

In sequence 3, the prior sequence (sequence 2) combination was expanded to include **first name**, **last name**, **date of birth**, **administrative gender,** and the addition of the **five digit zip/postal code**. This combination of traits reduced completeness by a significant amount, down to 91.1%, but dramatically increased uniqueness to 99.2%. This was a significant finding and allowed the creation of an algorithmic rule stating, in essence, if these traits are all available for a given patient, then use this set of traits to match across organizations.

For sequence 4, the same set of traits was used as in sequence 3 with the substitution of a **telephone number** for the five digit zip code. This further reduced completeness to 76.2% reflecting the low availability of telephone numbers, but it also provided the best uniqueness at 99.5%. This resulted in another rule indicating that if these five traits are available for a given patient, they should be used. This became the highest precedence rule—if these traits are available, then this is the first rule to be applied. However, if these five traits are not all available, lower precedence rules will be utilized.

Sequence 5 used the same set of traits as was used in sequence 4, but with the substitution of a **middle name** for the telephone number. This reduced completeness to 59.9% and resulted in a relatively high uniqueness of 98.9%. A rule was created to use these traits, when available.

In sequence 6, a rule similar to sequence 5 was created but it only used the **first character of the middle name** instead of the full middle name. This resulted in 60% completion with a reduced uniqueness, as expected, to 97.7%. So this resulted in a rule that if no higher precedence rules was first applied, and if that patient's full middle name is available, then it is used to match. If the
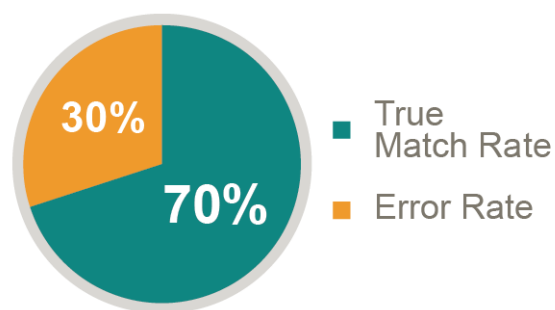

Figure 2: Initial Performance Analysis

full middle name is not available, but the patient's middle name first character is available, then it will be used to match.

Finally, sequence 7 used the **first name**, **last name**, **date of birth**, **administrative gender**, and **last 4 numbers of the SSN**. This resulted in low completeness of 61.9% but high uniqueness at 99.7%. This too became a rule, with careful placement between the other patient matching rules.

It should be noted this was an inward-facing analysis. That is, it was a review of these traits, for those patients, only at Intermountain Health Care. It still remains to be determined how these introspectively-derived rules would work across organizational boundaries for other exchange partners.

Expectations Based on Low Effort, High Yield Data Rework

After performing the above steps, a new analysis was conducted using the same small sample data set to determine the matching results. **The analysis showed a marked improvement with a true match rate of approximately 70%.** This represents a significant improvement over the prior results.

Next, testing of these new rules was expanded to include a larger sample data set of 340,000 patients.

When analyzed, Intermountain found that the patient matching success rate for this larger sample set was still unacceptably low. Note, this match rate is where Intermountain expected to start from in terms of cross-organizational patient matching.

Additional data quality interventional work was performed, focusing on:

- Data entry control
- Enforced data integrity checks
- Data transcription problems (from paper to EHRs)

A new analysis was conducted after making these internal data quality improvements. This analysis established the new **expected** values.

Figure 3: Expected Results After Data Rework and New Rules.

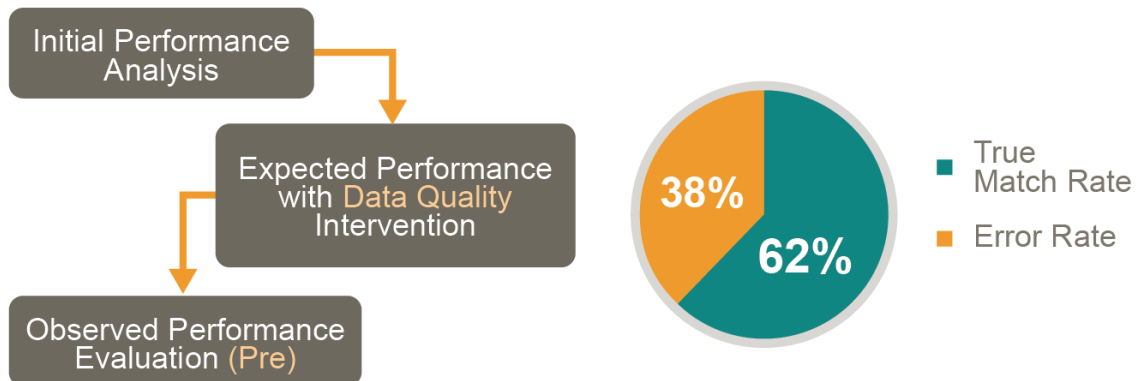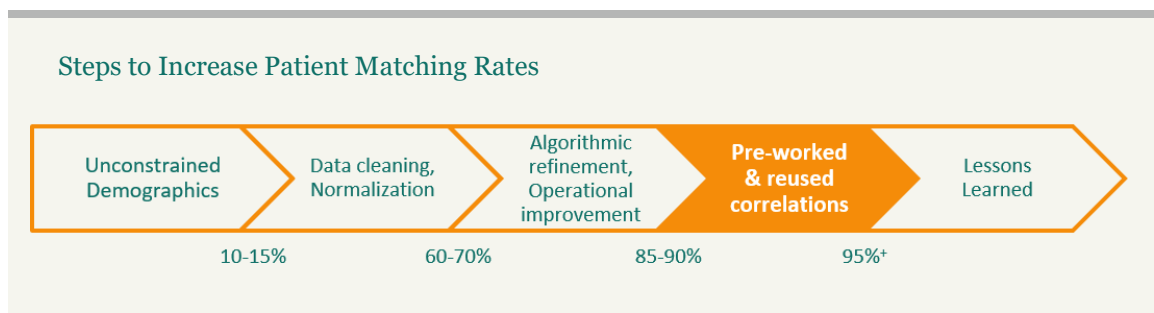Next, the actual performance with data intervention was determined.

Figure 4: Actual Observed Performance After Data Interventional Improvements

This resulted in a 28% worse performance than expected (a 10% error rate was expected, not a 38% error rate). Further investigation was conducted.

Step 4: Operational Performance Measurement and Improvement

Step 3 resulted in an observed correct match rate of 62%, which was far below the predicted 90% match rate. The disparity between the predicted match rate of 90%, compared to the actual results of 62% in Step 3, warranted further analysis to ascertain the causes of the failure rates.

The 38% error rate was broken down into more detailed contributing factors.

As shown in Figure 5: Detailed Analysis of 38% Error Rate, we uncovered five major factors: algorithmic, authorization, network, messaging, and a more intrinsic error category. Systematically, Intermountain Healthcare worked with their exchange partners to address each addressable issue.
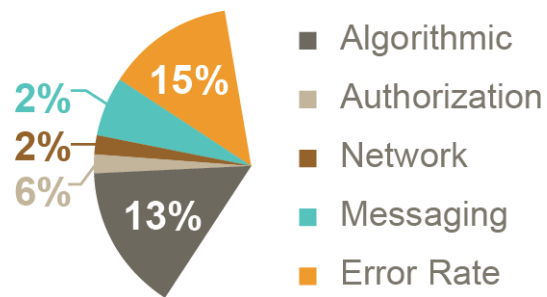
Figure 5: Detailed Analysis of 38% Error Rate

Algorithmic: Various algorithmic patient matching improvements were implemented at this phase including data normalization, selection of traits, blocking strategies, bucketing strategies, as well as additional improvements.

Authorization: Although outside the scope of this paper, the lack of an appropriate patient authorization resulted in what appeared to be a failure to match when the root cause was a permissions issue. Timing was an unexpected factor with respect to the patient authorization issue. For example, in some cases a patient would "opt-in" but there would be a delay before that status was reflected in all systems. *The CCC is prototyping an innovative approach for addressing this issue. Those interested in learning more on this topic are encouraged to reach out to the CCC.*

IT Networking Issues: Some IT technical issues were also identified, namely, network timeouts. This class of issues manifested themselves as <u>apparently</u> failed patient matches, however, the root issue was that the responding system for a given patient matching

... apparent patient matching failures were actually due to network timeout issues

request was not received before the initiating system gave up. An apt analogy is a telephone call. If the party being called doesn't answer after 15 rings, then the person placing the call may give up and disconnect. If the party being called would have answered reliably at 20 rings, then there was a failure that could have been avoided if the calling party waited longer. In a similar manner, Intermountain worked with University

of Utah (its exchange partner for this test) to configure both systems' network timeouts to higher values. As an aside: we are intentionally avoiding the issues of service levels and use case-driven response time requirements at this stage, but we will come back to this topic in the future. In summary, some apparent patient matching failures were actually due to network timeout issues.

Security Header Issues: In addition, a number of the failures were attributed to technical errors unrelated to patient traits. This specific case study leverages the eHealth Exchange (which is based on an open standard called IHE XUA), each message has an

> ... patient matching was effectively blocked by a technical error unrelated to patient traits

internal technical component called a security header. This component is part of the wrapper around each patient matching request message between exchange partners. Inside this security header is a very important set of data indicating the purpose of the request (such as for treatment, claims, patient authorized exchange, or emergency), the requesting person, the requesting person's role, the patient's authorization, and much more. These data are contained in a section of the message called the Security Assertion Markup Language (SAML) header. A number of issues were found to exist around the generation and/or consumption of the SAML header where, for example, the SAML header could not be properly understood by the receiver resulting in an error. The implication of this class of issues is that patient matching was effectively blocked by a technical error unrelated to patient traits. An analogy is when a physical package is not delivered because the recipient address on the package was illegible. This class of issues were resolved by the technical implementation teams.

Data Encoding: A number of additional issues were related to data formats. This was a key area of improvement for this case study. It was originally assumed that data would be consistently formatted internally and across organizational boundaries. Unfortunately, this was far from being true.   Most fields in fact had different representations that required normalization before they could be compared. Telephone numbers had to be tagged so that work telephones were not being compared with mobile telephones.   And telephone numbers had very little consistency in the use of special characters.   These issues were addressed by removing all special characters and normalizing them to an international standardized format. It was also discovered that in many cases first names were being combined with middle name or initials. This was also corrected. Another large group of problems was associated with the use of names with

non-alphabetic characters (e.g. O'Toole) or internal spaces (e.g. Van Der Camp). **Standards were developed and applied to normalize these names.**

As a result of these improvements, the true match rate improved to approximately 85%. That is, about 15% of the matching problems still remained. Obviously, this was a significant improvement with the best result to date for the study. Next, a number of best practices were utilized to make further incremental improvements.
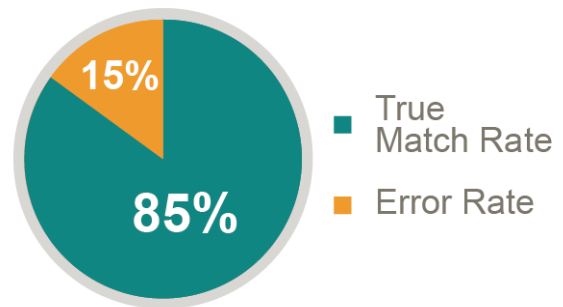
Figure 6: The True Match Rate

Step 5: Implementation of Best Practices and Lessons Learned
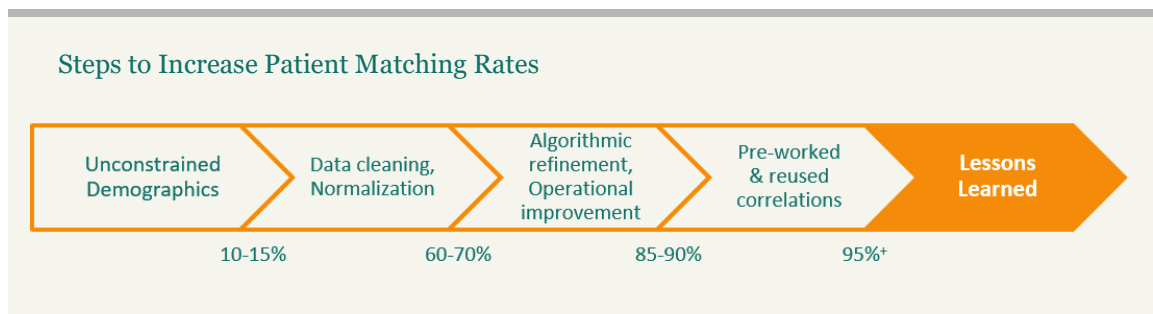
To address the remaining 15% error rate, a number of best practices were identified and implemented. These practices included:

- Working systematically with partners to reduce or eliminate prior issues;
- Applying results of prior work (see Lessons Learned);
- Agreeing with trading partners on data standardizations;
- Agreeing on consent synchronization;
- Expecting no less than a 90% match rate across organizations;
- Keeping investment reasonable by agreeing when goals are achieved;
- Focusing on scalable solutions;
- Pre-working fragile identities when possible (see Lessons Learned);
- Improving the human workflow; and
- Leveraging CCC Shared Services.

After employing these identified best practices, Intermountain expected a successful match rate of approximately 90% as shown in Figure 7: Expected Error Rate at this Phase. However, as seen in Figure 8: Actual Results, at this stage they achieved a 95% match rate, which was significantly better than expected.

Figure 7: Expected Error Rate at this Phase
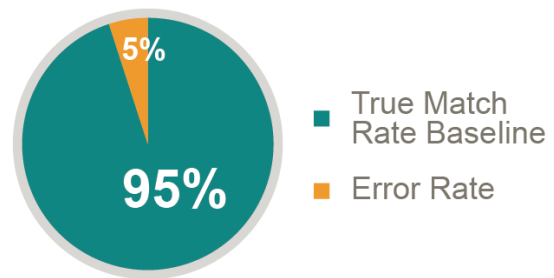
Figure 8: Actual Results

After a review of the reasons for the increased patient matching success rate of approximately 95%, the major factors contributing to this successful result were identified as:

- The use of the CCC Shared Services;
- Collaborating with external exchange partners regarding standardized data formats; and
- Addressing patient consent issues.

In the next section, overall case study lessons learned will be shared.

Lessons Learned

Fragile Identities: One interesting outcome of the case study was the emergence of a category of patients that repeatedly failed to match correctly (false positive match or a false negative match) even

> … a category of patients emerged that repeatedly failed to match correctly…

after multiple interventions. These patients' traits were manually edited to enable successful matching, but later the patients again failed to match. This process repeated several times (failure to match, rework, match, and a failure to match) during this study. As a result, these patients were put into a category called "fragile identities." Some of these patient identities were studied in detail and the root causes of their repeated failure to correctly match were identified. In most cases, the repeated matching problem was due to "thin" demographics, such as just a first and middle initial instead of full first and middle names, a missing address, or an address using non-standard abbreviations. There

were also a few extreme cases that patients have the same date of birth, address, last name, very similar first name (twins). A strategy was developed to (1) identify the characteristics of patients in this category, (2) query for and create a work list of patients falling into this category, and (3) implement manual remediation of those patients' identities to **proactively** attempt to resolve future patient matching problems. The case study demonstrated the effectiveness of the approach in a **reactive** scenario but it is expected that by providing a proactive remediation standard operating procedure, we can eliminate this entire class of patient matching issues.

Well Behaved Group: In sharp contrast to the "fragile identities" group, another group of patient identities emerged that exhibited the opposite behavior: their identities seemed to almost always match correctly (true positive or true negative match). This group was regarded as being robust with respect to patient matching. Members of this group were analyzed in an attempt to determine the reason(s) this group exhibited desirable matching behavior. The single most important factor distinguishing this group from the fragile group was the presence of complete demographics. These patients had a full, correctly spelled name including their middle name and any special characters. The patients had a complete current address and telephone number. These patients also had historical name information and historical address information. The traits of this group of patients are being used to inform the best practices for the other patient groups including the human workflow implications such as ensuring hospital admissions staff are trained and motivated to enter and use robust and complete patient traits.

Knowledge Reuse: Manual work on a patient's identity is expensive, slow, and error prone. It has a large negative impact on the speed that patients can begin treatment. However, it represents very valuable information gained. Once a patient's records have been manually analyzed and remediated, that information can and should be leveraged in the future to prevent repeated manual rework on the same patient record. For example, a correction made to a record should not be made in a read-only system rather it should be updated via an approved secure workflow to a master patient record. This allows future patient matching activities to leverage the improvements made from the manual process and ultimately makes the organization more efficient. In a similar manner, the links created between patient records also can and should be leveraged. Those linking decisions should be stored in such a way that they can be reused in the future.

Involve Patients in Identity Management: Patients themselves are valuable allies in helping maintain their identities. Two methods have been identified so far. A critiquing service involves patients, at the point of care, in helping to link, correct, unlink, and otherwise update the patient records. It is also envisioned that in the future a patient portal (or other self-service application) could perhaps help patients understand their identity completeness in a manner similar to a password strength test offered by many sites and applications.

Observations and Recommendations

- The biggest opportunity to immediately impact matching rates is standardized formats for demographic data among data sharing participants.
- Consistent name representation will be a challenge without probabilistic assistance because of data collection workflow issues that favor alternate representations (such as preferred name over legal name).
- Acceptable patient matching data integrity (99.99%) may require a supplemental identifier in addition to the required fields. This allows for probabilistic linking where alternative representations are allowed among the exchange participants and where established linkages are expected to be reusable for future exchange transactions.

Case Study in Review

The next generation of patient matching is still on the horizon. Healthcare is still in its infancy with respect to patient identity management between organizations (also more correctly known as record linkage). Many other domains have studied patient matching at an industry and academic level for many years. Several industries, specifically financial services and airline transportation, have legislative support for unambiguous matching of their customer records. However, legislative support for patient identity matching is not assumed or suggested in this draft framework. We can perhaps learn from the consumer credit reporting bureaus whom have been working on this problem for many decades. Today, credit bureaus use approximately 140 separate traits to match people to the correct database record. Not only are consumers matched, but their physical addresses are also matched in a similar way as patients. Additionally, credit bureaus link consumers to a web of other entities that provide credit to that consumer. The net result is not perfect, but it allows for a national or international scale solution that can inform our work within the health IT domain.

The CCC is also innovating and developing multiple next-generation patient matching approaches. Several of their key approaches include:

- A critiquing service to involve patients and providers in the patient matching workflow at the optimal time with a feedback loop to leverage such knowledge gained;
- Authoritative sources on a field-by-field basis;
- Identification and re-work of fragile identities;
- Staff incentives;
- Consent shared services;
- Patient matching/shared record locator services;
- Data quality analysis; and
- Empirical analysis across several organizational boundaries.

As Intermountain continues to strive for perfect patient matching within and across organizations, they anticipate that these improvements will allow for patient matching rates to exceed 99%. Proposed future improvements may include:

Figure 7: Future Match Rate

- Use of biometrics, specifically patient retinal scan;
- Proactive correlations; or
- Patient engagement in identity management.

CHAPTER 3: CROSS-ORGANIZATIONAL PATIENT MATCHING MATURITY MODEL

Introduction

As mentioned in Chapter 1, patient identity management has remained in the national spotlight as a key prerequisite to successful health information exchange. The purpose of the maturity model, described in this chapter, is to provide a method to evaluate, measure, and improve patient matching deployments across organizational boundaries. The proposed maturity model is designed to provide a simple framework aiding in the comprehension of this domain and to focus on process change. We believe that more precise definitions of the maturity model will give organizations the ability to adopt more advanced patient identity management in a methodical manner.

This framework is based in part upon the International Organization for Standardization (ISO) framework (which includes people, process, and technology) with the added dimension of governance.

Scope

Patient matching is often thought of in two very different domains: (1) patient identity management within an organization and (2) identity management across organizational boundaries. The scope of this paper is largely focused on patient matching across organizations. While there is overlap, and these areas will also be discussed, this paper does not focus on patient identity management inside organizations otherwise.

This chapter is intended to solicit feedback on a proposed maturity model of patient identity management and is therefore being submitted to health IT community in draft form for public comment.

Feedback Requested From

Feedback on Chapter 3 is requested from all organizations participating in health data sharing across organizational boundaries, vendors, subject matter experts on patient matching, record linkage and entity resolution, state and federal government, standards bodies, and the general public.

Timeline

Please provide feedback on this chapter by January 22, 2016. The Sequoia Project plans to publish the final version of this document in 2016.

Process

Please send all comments to feedback@sequoiaproject.org. Once comments have been received, The Sequoia Project will convene public working sessions with all interested parties to reconcile the comments. The resulting model will be made publicly available for customization and adoption.

Instructions for Reviewers

While reviewing the proposed maturity model please answer the following questions:
1. Are these maturity model levels consistent with various levels of patient matching results?
2. Will these maturity model levels be possible for my organization to implement? If so, when?
3. If an organization achieves the highest level defined in this paper, what else could it do to increase cross-organizational patient matching success? In other words, what is missing from these maturity levels?
4. If all of these levels cannot be adopted by my organization, why? Are the levels inapplicable or insufficient? Are they difficult or impossible to adopt?
5. Finally, are there components of each level that should be moved or removed?

Please use the comment submission form at the end of the chapter to provide your feedback on this maturity model. We anticipate that this model will be finalized over the next few months so please review carefully and comment now by sending your completed comment form to feedback@sequoiaproject.org.

Characteristics of Mature and Immature Organizations

It may be helpful to define some general characteristics of immature and mature organizations.

Immature organizations generally possess the following characteristics:

1. Process is improvised
2. Known processes are commonly ignored
3. The organization is in reactive mode
4. Schedules, staffing plans, and budgets are not fact-based
5. Quality is sacrificed
6. Quality is not objectively measured

Mature organizations typically possess the following:

1. Coordination, communication, and collaboration across silos
2. Work plans are generally realistic and accomplished for common project types
3. Process and practice are largely in agreement
4. Processes improve over time
5. Staff understand their responsibilities and there are no key gaps in staffing or skills
6. Management and staff are aligned

Summary of Levels

Carnegie Mellon University's Software Engineering Institute (SEI) is generally regarded as having created the first Information Technology maturity model. (http://www.sei.cmu.edu/cmmi/index.cfm). There are five levels defined for their model, which, according to the SEI, "Predictability, effectiveness, and control of an organization's software processes are believed to improve as the organization moves up these five levels. While not rigorous, the empirical evidence to date supports this belief." The SEI Capability Maturity Model's (CMM) Five Maturity Levels of Software Processes are:

1. Initial (chaotic, ad hoc, individual heroics) - the starting point for use of a new or undocumented repeat process.
2. Repeatable - the process is at least documented sufficiently such that repeating the same steps may be attempted.
3. Defined - the process is defined/confirmed as a standard business processes.
4. Managed - the process is quantitatively managed in accordance with agreed-upon metrics.
5. Optimizing - process management includes deliberate process optimization/improvement.

79 The Sequoia Project believes these levels are a valuable construct to help guide our
80 thinking about cross-organizational patient matching.
81

> Level 0: Indicating ad hoc processes and outcomes, and little to no management oversight or recognition;

> Level 1: Indicating adoption of basic defined processes with associated repeatable outcomes, and limited management involvement;

> Level 2: Indicating increasing maturation of processes, definitions of most key processes, data governance, algorithm use, active management involvement, and accumulation of quality metrics;

> Level 3: Indicating advanced use of existing technologies with associated management controls and senior management awareness, and use of quality metrics; and

> Level 4: Indicating innovation, ongoing optimization, and senior management active involvement.

82

83 At Level 0, people are disorganized, processes are not well understood or defined, and
84 the overall view of cross-organizational patient identity management is that of a chaotic
85 system because of the lack of replicable results. Success often depends on individual
86 heroic efforts. The organization is often in reactive mode instead of one of proactive
87 management.

88

89 At Level 1, there is a growing awareness of the critical role of cross-organizational patient
90 matching and a corresponding recognition of the need to apply basic management
91 controls by lower and mid-level management staff. At this level, some processes are
92 repeatable, but not all. Success is more predictable. The quantity of reactive mode
93 events declines. Level 1 organizations, and above, have implemented all applicable *Cross-*
94 *Organizational Patient Matching Minimal Acceptable Principles,* as described in chapter
95 4*.*

96

97 At Level 2, all key processes related to cross-organizational patient matching are
98 understood and documented. They may be enforced somewhat inconsistently. The
99 organization is normally not in a reactive mode, and unexpected events become relatively
100 rare where they were the norm in Levels 0 and 1.

101

102 At Level 3, organizations monitor, analyze, and systematically improve their ability to
103 manage patients across organizational boundaries. Most if not all processes are defined
104 and documented. The processes are somewhat rigid. However, at Level 3, the processes
105 are largely documenting the system behaviors "as is," as opposed to Level 4 where the

106	processes are innovative. At Level 3, the organization achieves consistency but it is not
107	optimal.
108	
109	At Level 4, innovation becomes a standard component of patient matching. Management
110	uses data accumulated to model and, as is deemed viable, implement sometimes
111	significant improvements. Key staff members are considered leaders in this domain and
112	contribute to the community.
113	
114	The time an organization has existed is not necessarily strongly correlated to an
115	organization's maturity. Each level will likely have to introduce innovation to advance to
116	the next level.

Level Characteristics

(line 117)

Workforce
(line 118)

- At Level 0, an organization has staff focused on **internal** patient identity and matching with no dedicated staff **outwardly** focused (i.e., on patient matching across organizational boundaries). In the unlikely event that some staff are focused on outward patent matching, then they do such because of a personal recognition of the need, instead of this recognition occurring at an organizational level. Or they are focused on cross-organizational patient matching due to a limited scope project. Workforce patient identity management formal processes, training, skills development, and career path are not recognized. Staff members are not trained. Job titles do not exist for patient identity management staff; cross-organizational patient identity management is often a responsibility that is added on to other positions.
- At Level 1, management has recognized the need for specific assignments for external patient matching, and have started formulating plans.
- At Level 2, staff are devoted, at least part-time, to cross-organizational patient matching.
- At Level 3, staff include formal responsibility for cross-organizational patient matching. Training is accepted as necessary and appropriate. Staff are involved with industry initiatives.
- At Level 4, staff involved in patient identity management are involved at more senior levels within the organization and are leading innovation with respect to this topic. See also, Standards Development characteristics.

Patient Involvement
(line 140)

- Patient involvement in their identity management does not exist at Level 0. At Level 0, organizational staff do not consider the patient to be a part of the identity management workflow other than to confirm their demographics when checking-in or registering.

145 • At Levels 1 and 2, the patient is starting to be recognized as a potential active
146 participant in their identity management.
147 • At Level 3 the patient is involved via manual workflows and processes, but no
148 system changes are made to accommodate such involvement.
149 • At Level 4, the patient is recognized as a key ally in optimal patient identity
150 management. In addition, at Level 4, the knowledge gained as patients become
151 involved in their own identity management is durable, shared across the
152 enterprise, and reused for subsequent cross-organizational patient identity
153 management.

154 *Use of Technology*
155 • Use of technology at all levels is assumed.
156 • However, the deployment of technology at Level 0 is
157 largely built around ad hoc processes and standards,
158 such as using custom data interfaces that are not
159 fault tolerant, robust, performant, or well
160 documented.
161 • At Level 3, the organization is using software of
162 commercial quality, either from a strong performing
163 third party vendor, or custom developed with the
164 same degree of robustness, performance, fault resilience, and internal
165 documentation.
166 • At Level 4, the organization has developed new technology, is continuously
167 testing their innovative technology, and is submitting refined version to
168 Standards Development Organizations (SDOs) to help advance the industry.

169 *Communication and Community Involvement*
170 • In terms of participation in communications within the health information
171 technology communities, organizations at Level 0 are largely isolated. This
172 results in them being unaware of standard approaches to common problems and
173 in deploying solutions to problems they believe are unique which, in fact, are
174 common.
175 • At Level 1, organizations become aware of standards and communities and begin
176 formulating plans to begin participating in broader communities.
177 • At Level 2, organizations are involved in appropriate health IT communities, such
178 as those curating relevant standards, state or regional exchanges, and state-wide
179 approaches.
180 • At Level 3, organizations are fully integrated into most relevant health IT
181 communities, such as SDO committees.
182 • At Level 4, organizations exhibit leadership in relevant communities such as by
183 co-chairing workgroups and testifying in front of state and national legislative

184		bodies and agencies. They participate in board, state, and federal advisory
185		committees, etc.
186	•	At Levels 3 and 4, key staff members also frequently share negative and positive
187		knowledge gained to help others understand patient matching problems and
188		solutions better so that they may leverage prior work.

189	*Workflows*	
190	•	At Level 0, workflows are based on speculated needs and are not driven by
191		confirmed, high-priority use cases.
192	•	Workflows at Levels 1 and 2 are largely driven by the desire to meet federal
193		regulatory requirements.
194	•	Level 3 workflows are driven by more advanced objectives such as full round trip
195		immunization query, administration, update, and reporting. Cross-organizational
196		partners are partially incorporated into workflows.
197	•	Level 4 workflows are driven by advancing the state of the art and tracking
198		adherence to the best demonstrated practices. Level 4 includes optimization of
199		workflow to incorporate partner organizations.

| 200 | ## External Matching Focus |

201	**Organizations that are at Level 0 often do not yet understand that rules of patient**
202	**identity management that work for them within their enterprise do not necessarily**
203	**work across organizational boundaries.** One example is an organization that makes no
204	distinction between patient matches across organizations from those that occur
205	internally. It assumes that patient demographic feeds, including merges, links, unmerges,
206	unlinks, and demographic updates are occurring both for internal patient matching and
207	for their partner organizations externally. This manifests itself as policies and procedures
208	for patient identity management that are not viable because they cannot enforce their
209	internal enterprise policies and procedures across organizational boundaries. A more
210	specific example of this is that the organization uses a master patient index (MPI)
211	configuration that is only effective at matching patients if the patient's SSN is provided.
212	Internally, they can enforce that policy; externally, they cannot.

| 213 | ## Testing |

214	•	Testing of identity management solutions at Level 0 is minimal, manual, ad hoc,
215		and does not consistently assure successful deployment of "passed" systems.
216	•	At Level 2 and above, testing is largely automated, based on significant real
217		world lessons learned, and is a good predictor of a successful deployment.
218	•	At Levels 3 and 4, testing programs assure, with a high degree of confidence,
219		successful deployments.

Use of Patient Matching Quality Metrics

- At Level 0, the value of patient matching quality metrics are not recognized.
- At Level 1, the organization has begun to recognize the value, and has started planning for the future capture of metrics.
- At Level 2, quality metrics are being captured.
- At Level 3, the metrics are being used to actively improve.
- At Level 4, the metrics are being further refined, and include feedback loops to the systems and organizations involved in patient identity management. Their external health IT trading partners join in metrics capture, use, and feedback.

Diagnostic Approaches

Diagnostic approaches vary with the level of maturity (ad hoc, some automation, full automation, innovative approaches). **At Level 0, all patient matching exceptions and errors require human intervention.** Processes are not well understood or documented. Errors are often not recognized in a timely manner. Manual work queues are not consistently staffed. Management has little to no visibility into exceptions (frequency, types, root causes, impact, remediation plan, etc.).

System Stability

- At Level 0, system stability and performance is brittle. The overall patient matching system is not well regarded by end-users, administrators, or management. It is considered error prone and is not trusted to be reliable or available. It often returns unexpected patient search results or errors. False positive and false negative matches occur frequently. Manual disambiguation of returned patient searches is frequent. User disillusionment and abandonment is common at this level. Clinical users, in particular, see their hopes of improved patient medical records availability dashed and stop using the system.
- Incremental improvements are made at Levels 1-2.
- At Level 3, end-users see a personal return on investment "ROI" of making the effort to query resulting in significant payoff thereby accelerating adoption of cross-organizational patient matching into their personal processes. Organizations incorporate use of cross-organizational patient records as a best demonstrated practice or standard of care.
- At Level 4, organizations have feedback loops with senior representatives of their staff to identify innovative new approaches. This results in the

256   identification of new patient identity management strategies that are novel and
257   valuable.

Management Oversight

259 • At Level 0, management oversight is virtually non-existent. Management may not
260   even have a firm grasp of the definition of cross-organizational patient matching and
261   records exchange. The business and clinical value of cross-organizational patient
262   matching is not recognized.
263 • At Level 1, management awareness has increased and basic management
264   controls are being defined.
265 • At Level 2, management is actively
266   involved in cross-organizational
267   patient matching. Initial
268   management controls have been
269   implemented, are being used, and
270   being improved. Metrics are being
271   captured but are not yet being fully used.
272 • At Level 3, management is leveraging metrics. Senior management is aware of
273   the importance of cross-organizational patient matching as being of strategic
274   importance as a prerequisite for other activities such as care summary
275   exchanges. Management ensure that key staff are trained and skills developed.
276   Workflow is reviewed and optimized at a system-wide level to ensure that
277   patient matching dependencies, such as proper staff incentives, are in place.
278 • At Level 4, management has empowered the organization to assume a
279   leadership role in the industry. Innovation projects are funded and staffed.
280   Innovations, once proven, are incorporated into production operations.
281   Knowledge is shared with SDOs and with the wider community via significant
282   industry involvement. Senior management includes at least one member that is
283   focused on cross-organizational identity management as a formal area of
284   responsibility.
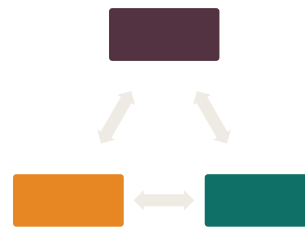
Use of Industry Standards

286   Use of industry standards and calibration/backfill of those standards is very
287   different across levels.
288 • At Level 0, the organization may use custom solutions that are based on
289   standards, or that are based on standards that are not well understood.
290   Organizations at this level may also have the naïve belief that the standards will
291   solve more problems than they actually do.

292   •   At Levels 2 and 3, use of standards has matured, organizations understand that
293   standards have limits but they leverage those capabilities fully, and provide for
294   backend system support, such as more advanced internal algorithms, to make
295   the best use of standards. Organizations also work with SDOs to fix errors and
296   vagueness in the standards.
297   •   At Level 4, organizations take the initiative to create new standards and to
298   suggest significant improvements to existing standards.

Establishment of Feedback Loops

300   •   At Level 0, feedback loops do not exist; nor does the
301   recognition exist that feedback loops are of value.
302   •   At Level 1, feedback loops are established with a few
303   primary data sources.
304   •   At Level 2, feedback loops are established with most data
305   sources and other key workflow participants.
306   •   At Level 3, feedback loops are established with all participants, human and
307   system, in the patient matching process inside an organization.
308   •   At Level 4, feedback loops are established with all participants, human and
309   system, in the patient matching process across organizational boundaries.

Fragile Identities

311 At Level 4, organizations recognize that some patient identities are "fragile" and tend to
312 consistently be false negatively matched or false positively matched. This can be due to
313 demographics that are not sufficiently rich or that are very similar to other people, or
314 errors such as an incorrect identifier. Level 4 organizations recognize this class of issues
315 and implement an associated process to systematically identify identities in this category,
316 assign appropriate staff to remediate these fragile identities, and then measure the
317 results to confirm the resolution. Note that certain identities are the opposite from fragile
318 identities, specifically those that rarely match incorrectly. Level 4 organizations analyze
319 these consistently matched identities, learn what characteristics make these identities
320 largely immune from mismatching, and then leverage this knowledge to help manage
321 their entire census.

Flow Down

323   •   At Levels 0 to 2, organizations do not have any special provisions in their various
324   legal agreements covering the organizations with which they contract to enforce
325   any degree of patient identity management practices. For example, if the
326   organization is an data sharing network with participation agreements with their

327		hospitals, the data sharing network may not require their hospitals to quality-
328		assure patient demographics.
329	•	At Level 3, organizations recognize that their patient matching processes will
330		only have limited utility unless they legally bind vendors and organizations
331		connected to their exchange partner to patient matching principles. Terms
332		covered in these binding agreements include data quality, data completeness,
333		key workflow components such as dealing with minimization of duplicate
334		records, minimal patient matching practices, exception handling, and service
335		levels.
336	•	At Level 4, organizations must obligate their internal patient matching data
337		sources, consumers, vendors, and systems to the same. Data sharing networks
338		legally bind their network members and HIEs legally bind their participants to
339		comply with patient matching principles.

Knowledge Sharing

341	•	At Level 0, knowledge gained is often lost since the organization is largely in
342		reactive mode and is often "fighting fires."
343	•	At Level 1, the organization has some recognition of the
344		value of capturing knowledge but there is no formal process
345		for capturing it.
346	•	At Level 2, knowledge about patient matching processes is
347		captured and shared internally to a limited extent.
348	•	At Level 3, the knowledge is shared with partners and is
349		starting to be shared with the broader health IT community.
350	•	At Level 4, the knowledge gained about a specific patient match is implemented in
351		automated systems that leverage the information broadly and durably. For
352		example, take an organization that has manually resolved an external patient
353		matching investigation with a partner and identified consent as being the root
354		cause. The organization has a method of incorporating that knowledge into their
355		patient matching data and processes, allowing this entire class of issues to be
356		permanently resolved.

Temporary (Default) Values

358	•	At Level 0, default or known temporary values are defined for the organization.
359	•	At Level 1, temporary values are inventoried and defined for the organization
360		and its key exchange partners. The inventory is not enforced through user
361		interface data capture, though, and occasionally temporary values are
362		discovered that were previously missed. Partner temporary values inventories
363		are also incomplete.

364	•	At Levels 2 and above, the inventory of known temporary values is accurate.
365		Technological enforcement of these values is in place including at the staff data
366		capture levels as well as at the automated data exchange levels. Staff processes
367		are also in place to enforce this list. As new partners begin the onboarding
368		process, temporary value inventories are exchanged as part of the formalized
369		process.

New Partner Connectivity

371	•	At Level 0, each new data exchange partner is connected using an ad hoc process.
372		Patient matching attributes are not documented or incompletely documented.
373		Partner data considerations, such as quality and availability, are not accounted for
374		during the planning and implementation process. This leads to many production-
375		level unfulfilled expectations and errors: This is a large component of the reason the
376		Level 0 organization is often seemingly fighting fires. They simply have not yet
377		gained enough experience to proactively manage, predict, and resolve common
378		issues with partner patient identity management.
379	•	At Level 1, key patient matching related considerations are documented and
380		considered during the implementation process with each new partner or for the
381		network(s) the organization participates within.
382	•	At Level 2, key patient matching processes with new data exchange partners are
383		included into manual testing processes. Patient identity management
384		considerations are well documented into complete, implementable
385		specifications. Standards are adhered to, when possible, and manual testing
386		confirms adherence.
387	•	At Level 3, patient identity management testing processes are automated and
388		patient identity management related on boarding processes are forward-looking
389		and request adherence to best practices.
390	•	At Level 4, heuristics and advanced processes allow for deeper insights into data
391		exchange partners to understand their patient identity systems, processes, and
392		workflows. This results in advanced configuration of partner systems'
393		integrations to optimize their success in patient matching.
394		

Data Quality

396	•	**At Level 0, data quality is an unknown.** Technical staff and management have
397		very limited awareness of the importance of data quality, or the status of their
398		data.
399	•	At Level 1, data quality has been identified as a key component of patient
400		matching across organizations. The organization may attempt to side-step their

401	own data quality issues by asking their exchange partners to make adjustments
402	to their data. No formal analysis of data quality exists, but there is a growing
403	awareness of the need for formal control of data quality.
404	• At Level 2, data quality has been assessed and is well understood. The
405	organization understands its data quality situation, and that of its primary
406	exchange partners.
407	• At Level 3 and 4, data quality is expanded in scope to include human workflow
408	considerations, staff training and formal responsibilities, and is being reported to
409	senior management, whom is measuring and tracking progress towards targeted
410	improvements.
411	• For more information see Chapter 1 and the Feedback Loops topic in this
412	chapter. The majority of topics discussed in the case study presented in Chapter
413	1 are ultimately focused on the central topic of data quality.

414 *Table 3: Overview of characteristics by level*

Characteristic	0	1	2	3	4
External Matching Focus	○	●	●	●	●
Testing	◕	◔	◑	●	●
Use of Patient Matching Quality Metrics	○	◔	◑	◕	●
Diagnostic Approach	○	◔	◑	◕	●
System Stability	○	◔	◑	◕	●
Management Oversight	○	◔	◑	◕	●
Use of Industry Standards	○	◔	◑	◑	●
Establishment of Feedback Loops	○	◔	◑	●	●
Fragile Identities	○	○	○	○	●
Flow Down	○	○	○	◑	●
Knowledge Sharing	○	◔	◑	◕	●
Temporary (Default) Values	◔	◑	●	●	●
New Partner Onboarding	○	◔	◑	◕	●
Data Quality	○	◔	◑	◕	●

415 ## Summary

416 This chapter is designed to facilitate a broad discussion on the need to define cross-
417 organizational patient identity management characteristics in terms of levels of maturity.
418 Once updated to reflect public comment received and dispositioned, The Sequoia Project
419 and the CCC are hopeful that this paper gives management a useful model for methodical
420 assessment and improvement in cross-organizational patient matching. Moreover, we
421 hope this leads to national-scale improvements in our ability to accurately exchange
422 patient information while honoring patient privacy preferences, and ultimately providing

423 better care and outcomes to those patients. We welcome your feedback and look
424 forward to collaborating with you on this critical topic.

Outstanding Issues for Reviewers

426 Should the following topics be included as defining characteristics for this maturity
427 model?
428 • IT internal compartmentalization, which we define as the administrative
429   overhead necessary to diagnose common patient matching problems without
430   involving multiple departments
431 • Security considerations/threat models/risks
432 • Patient consent and/or authorization and its impact on matching
433 • Manual "pick up the phone" types of intervention that are needed for false
434   negative matching and related expenses
435 • Workforce training and incentives aligned with patient matching objectives
436 • Patient matching metrics identified, periodically assessed, managed, and
437   targeted for methodical improvement
438 • Active staff engagement in work queue management
439 • Fraud considerations (such as intentional misrepresentation of patient identities)
440 • Different assumptions made for interactive disambiguation vs automated
441   systems

Comments Submission Form:
Cross-Organizational Patient Matching Maturity Model

444 *Table 4: Cross-Organizational Patient Matching Maturity Model public comment response form*

Comment Number	Referencing Document Line Number	Original Text	Suggested Revised Text	Comments
1				
2				
3				

CHAPTER 4: CROSS-ORGANIZATIONAL PATIENT MATCHING MINIMAL ACCEPTABLE PRINCIPLES

Introduction

The Sequoia Project, in collaboration with the Care Connectivity Consortium, has identified patient matching and identity management as a key national impediment to successful health data sharing across organizational boundaries. In this chapter we are providing and **soliciting feedback on a list of minimal acceptable cross-organizational patient matching principles**.

This list of principles has been created in response to real-world production experience supporting large-scale health data sharing endeavors. These principles, after comments are received and incorporated, are intended to establish minimally acceptable expectations that an organization must meet in order to exchange with organizations adopting Cross-Organizational Patient Identity Management Maturity Model Level 1, as described in chapter 3: *Cross-Organizational Patient Identity Management Maturity Model*. Some rules seem obvious; however, experience has shown that there is a lack of consistent application of these rules in production across the nation. This paper aims to begin illuminating, enforcing, and improving the consistency of the application of these rules. **Once published in final form, we believe this list of principles will serve to create a component of a "Level 1" adoption model that organizations can target, test against as appropriate, and declare conformance to.** Additional levels beyond level 1 are described in Chapter 3.

Unlike previous chapters in this paper, Chapter 4 is oriented towards a technical audience.

Feedback Requested From

Feedback on Chapter 4 is requested from all organizations participating in health data sharing across organizational boundaries, vendors, subject matter experts on patient matching, record linkage and entity resolution, state and federal government, standards bodies, and the general public.

Timeline

Please provide feedback on this document by January 22, 2016. We plan to publish the final version of this document in 2016.

Process

Please send all comments to feedback@sequoiaproject.org. Once comments have been received, The Sequoia Project will convene public working sessions with all interested parties to reconcile the comments. The resulting principles will be made publicly available for customization and adoption.

Instructions for Reviewers

While reviewing the proposed rules in this document please answer the following questions:

1. Will adopting these patient matching rules help my organization, or the data sharing networks I support, achieve better patient matching results?
2. Will these rules be possible for my software to implement, and if so, when?
3. Are these rules so important, that I would prefer not exchanging with an organization if they don't implement each of these rules?
4. If these rules cannot be adopted by my software, why, and how do we compensate for these potential patient matching deficits?
5. The proposed patient matching rules in this chapter, are focused on IHE profile based exchanges using XCPD and XCA. Should these rules be made more generic at the expense of loss of specificity?
6. Finally, are there any minimal acceptable practices missing from this list?

Please use the comment submission form at the end of this chapter to provide your feedback on these rules. We anticipate that these principles will be finalized over the next few months so please review them carefully and comment now by sending your completed comment form to feedback@sequoiaproject.org.

Context and Next Steps

1. We intend that these 'minimal acceptable principles' will be introduced into the health IT community gradually in order to avoid the occurrence of breaking changes and to provide opportunity for vendors and health IT implementers to adapt and improve their patient matching approaches and success rates. The recommendations will initially be proposed as guidelines and evolve over time to become official policy, and then finally become part of testing programs. The proposed phases are described in more detail, as follows:

    Phase 1: Adopt principles as guidance. During Phase 1, these patient matching rules will be considered guidance, and will not be enforced through

512		testing programs and will not be a condition of joining or participating in
513		health data sharing networks.
514		**Phase 2:** Principles will become official policy but will not be tested. In this
515		phase, the "MAY/SHOULD" constraints will change to "MUST" constraints,
516		other than as noted below.
517		**Phase 3:** Principles will become an enforceable condition of testing and
518		onboarding processes for new or existing data sharing partners. At this
519		point, the list of patient matching rules will become part of the PASS/FAIL
520		testing criteria. Hence, organizations not meeting these criteria will not be
521		allowed to claim adoption of Cross-Organizational Patient Identity
522		Management Maturity Model Level 1, nor enter into production until the
523		identified deficits are remediated. Please keep this in mind when assessing
524		the rules. The testing criteria should be developed by the community via an
525		open, inclusive, consensus-based process and approved via applicable data
526		sharing connections or network formal change management processes,
527		including associated testing of organizations currently in production.
528	2.	These practices should be adopted as soon as is practical. It is expected that
529		some of these rules can be implemented at any
530		time, with little to no negative impact to exchange
531		partners. Other rules, such as those based on
532		workflow, will likely require partner coordination in
533		order to more effectively facilitate adoption.
534	3.	We anticipate that adoption and additional factors
535		will generate feedback to iteratively improve and
536		refine these patient matching practices. This list of Level 1 rules will likely evolve
537		with implementation and lessons learned.
538	4.	Organizations adopting these principles must obligate their internal patient
539		matching data sources, consumers, and systems to do the same. Similarly, a
540		data sharing network, must legally bind its network participants, who, in turn
541		legally bind their participants. An organization must legally bind its vendors and
542		systems.
543		

> Workflow rules will require partner coordination before adoption to avoid breaking changes.

544 Adopters of these principles are generally expected to deploy them initially as "SHOULD"
545 type of constraints, with the intention to change them to "MUST" constraints as quickly
546 as possible. In a similar manner, "SHOULD NOT" constraints will be changed to "MUST
547 NOT" constraints, with exceptions as noted below. The below proposed principles are
548 specifically intended for those organizations exchanging using the IHE International Cross-
549 Community Patient Discovery (XCPD) standard. However, it is expected that many of
550 these practices will also assist those exchanging using other standards. Feedback from
551 the community is requested—especially on similar rules for other standards, such as HL7
552 v2 query messages, HL7 FHIR and other RESTful approaches, etc.

Traits & Identifiers

554 Below are proposed rules to improve the use of traits and identifies for cross-
555 organizational patient identity management and matching. These are intended to be
556 implemented by technical staff who manage and maintain patient identity matching
557 systems.
558

559 1. Patient Discovery Initiating Gateways SHOULD query using all traits required by
560   the underlying specifications. In addition, where optional traits are known to be
561   of high quality, then participants SHOULD query using all possible optional traits.
562   [*Note to reviewers*: the term "high quality" is not defined. Can and should it be
563   defined? Is it possible to define?]
564 2. Patient Discovery Initiating Gateways SHOULD NOT require the use of any
565   specific identifier or value such as an SSN unless such a trait is required by the
566   applicable specification or standard. Any existing policy or statutory
567   requirements related to the use of SSNs for patient matching still apply. It is
568   anticipated that networks and members of those networks will either directly
569   reference IHE XCPD, or have a more constrained implementation specification.
570   For example, Carequality would reference the Query Implementation Guide, the
571   eHealth Exchange would reference the Patient Discovery Specification.
572 3. Patient Discovery Responding Gateways SHOULD NOT require the use of any
573   specific identifier or value such as a SSN unless such a trait is required by the
574   applicable specification or standard. Any existing policy or statutory
575   requirements related to the use of SSNs for patient matching still apply. It is
576   anticipated that networks and members of those networks will either directly
577   reference IHE XCPD, or have a more constrained implementation specification.
578   For example, Carequality would reference the Query Implementation Guide, the
579   eHealth Exchange would reference the Patient Discovery Specification.
580 4. Patient Discovery Initiating Gateways SHOULD NOT transmit any temporary or
581   default value for any patient trait as this can contaminate the partner gateway's
582   patient traits and/or result in false negative matches. (See appendix for a
583   definition of temporary values.) Although XCPD Initiating Gateways SHOULD NOT
584   transmit temporary values, if it is known that this operation will not harm any
585   exchange partner, then XCPD Initiating Gateways MAY transmit temporary
586   values. An example of harm would be if their partners add patient records to
587   their system based on inbound XCPD queries.
588 5. Patient Discovery Responding Gateways SHOULD NOT reply with any temporary
589   value for any patient trait as this can contaminate the partner gateway's patient
590   traits and/or result in a false negative matches. Although XCPD Responding
591   Gateways SHOULD NOT transmit temporary values, if it is know that this
592   operation will not harm any exchange partner, then XCPD Responding Gateways

593 MAY transmit temporary values. An example of harm would be if their partners
594 add patient records to their system based on XCPD query responses.
595 6. Patient Discovery Initiating Gateways SHOULD NOT make any assumptions about
596 how long a partner's patient identifier will be valid. Organizations that maintain
597 internal correlations between internal patient identifiers and external patient
598 identifiers, SHOULD implement the behavior described in Exception Handling #2
599 and #3 below. Alternatively, organizations that maintain internal correlations
600 SHOULD implement their systems so that they always issue a XCPD request,
601 before contemporaneous XCA Query for Documents/Retrieve Documents
602 requests.
603 7. Patient Discovery Responding Gateways SHOULD NOT make any assumptions
604 about how long a partner's patient identifier will be valid. Organizations that
605 maintain internal correlations between internal patient identifiers and external
606 patient identifiers, SHOULD implement the behavior described in Exception
607 Handling #2 and #3 below.
608 8. Patient Discovery Responding Gateways SHOULD NOT require identical
609 demographic traits on subsequent requests, as were used on the initial request,
610 if the same identifier provided on the initial correlation is re-used on subsequent
611 requests. By identical traits, we are referencing the exact same number of traits
612 and the exact same text in each trait supplied.
613 9. If applicable to their internal architecture, Patient Discovery Responding
614 Gateways MAY return multiple ambiguous matches per Assigning Authority.
615 Also note that this "MAY" constraint will remain a "MAY" constraint after the
616 remainder of these rules change to "MUST" constraints.   Responding gateways
617 SHOULD handle multiple ambiguous matches per Assigning Authority. [*Note to*
618 *reviewers*: Should rule #9 and rule #12 be adjudicated as they likely conflict for
619 some initiating systems resulting in potential duplicates? Rule #9 is intended to
620 facilitate disambiguation workflows with human staff review. Rule #11 and #12
621 are designed for fully automated systems. [*Note to reviewers*: Do we need to
622 discern these two cases?]
623 10. Patient traits transmitted by Patient Discovery Gateways to other Patient
624 Discovery Gateways SHOULD NOT be truncated. [*Note to reviewers*: should we
625 limit truncating systems to only query for patients where the data are known to
626 not be truncated? What, if any, harm does truncation cause to a partner
627 initiating or responding gateways?]
628 11. Patient Discovery Initiating Gateways SHOULD NOT supply more than one
629 patient identifier, per assigning authority. [*Note to reviewers*: should we require
630 XCPD Initiating Gateways to use a single gateway-wide PIX manager or other
631 patient identity management system and only provide a single patient ID for all
632 requests? What would be the positive and negative consequences of this
633 decision? What is the impact to federated XCPD Gateways or XCPD Gateways
634 fronting a multi-participant data center?]

635   12. Patient Discovery Responding Gateways SHOULD NOT return duplicate patient
636   records, or return the same patient record in such a way that a duplicate record
637   will be created by the XCPD Initiating Gateway. Internally, the systems behind
638   XCPD/XCA Exchange SHOULD NOT return the same patient with the identical
639   data using different assigning authorities or identifiers. The same patient should
640   use the same identifier for each request or response. This is important to
641   prevent duplicate patients from being created.
642   13. Patient identifiers SHOULD be consistent, not reused,
643   unchanging, and should prevent the creation of duplicate
644   patients at partner sites. A patient
645   identifier should not be constructed
646   in such a way that it dynamically
647   changes based on the known identity
648   of that patient at that time. Systems should not be allowed, for
649   example, to simply concatenate a list of all patient identities
650   together and return that value as the patient ID, since that list of all known
651   patient identities can change at any time.

> *Patient identifiers SHOULD be consistent, not reused, unchanging, and should prevent the creation of duplicate patients at partner sites*

Matching Algorithms

653 Below are proposed rules to improve matching algorithms for cross-organizational
654 patient identity management and matching. These are intended to be implemented by
655 technical staff who manage and maintain matching algorithms.
656

657   1. Patient Discovery Responding Gateways SHOULD track patient identity trait
658   changes and SHOULD respond based on prior or current (historical)
659   demographics.
660   2. Patient Discovery Responding Gateways SHOULD match based on normalized
661   traits.
662   3. Patient Discovery Responding Gateways SHOULD use case insensitive matching.
663   [*Note to reviewers*: should The Sequoia Project curate a list of normalization best
664   practices? A possible problem with this list is that normalization practices should
665   perhaps belong inside of each organization.]
666   4. Patient Discovery Responding Gateways SHOULD NOT use exact character-by-
667   character matching.
668   5. Other than immediately above, these rules will not define the specific algorithms
669   to be used, or avoided, since specific algorithms are system, vendor, data, and
670   organization dependent. [*Note to reviewers*: Do you agree with this approach?]

Exception Handling

Below are proposed rules to improve exception handling for cross-organizational patient identity management and matching. These are intended to be implemented by technical staff who configure and manage exception handling for patient identity systems.

1. Patient Discovery Responding Gateways MAY return an error indicating that additional patient consent may allow different, presumably more, information to be returned. This change will be implemented as per a timeline determined by the Patient Identity Management Maturity Model Level 1 Adopter. Note that this behavior only applies to Responding Gateways that would deny access based on lack of consent, and it only applies if returning such an error itself is not an impermissible disclosure.

2. An organization's patient identifiers SHOULD NOT be reused for different patients but the identifiers are allowed to be permanently decommissioned and a new identifier may be assigned the same patient. If a patient is merged, unmerged, linked, unlinked, or undergoes a similar transaction, the XCPD and XCA Responding Gateway SHOULD permanently decommission the identifier or identifiers formerly used to represent the patients subject to the merge, unmerge, link, or unlink. The XCA Gateway SHOULD generate an error for all subsequent Query for Documents or Retrieve Documents requests using that decommissioned patient identifier. Systems are not required to decommission identifiers if their internal logic is such that correct and complete patient data are returned for that identifier.

3. XCA Initiating Gateways SHOULD have logic in place to correctly process Query for Documents or Retrieve Documents errors indicating that a patient identifier has been decommissioned such that this triggers a new XCPD Patient Discovery request. [*Note to reviewers*: The purpose of this to make sure that initiators and responders are both exhibiting proper behavior in terms of decommissioned identifiers. The XCPD/XCA profiles assume this behavior, however, we are aware that many organizations have not implemented it – resulting in undefined behavior. One organization has expressed that this should be optional behavior, but we believe it should be required to avoid a patient safety issue. Please comment.]

4. Patient Discovery Initiating Gateways MAY use the XCPD "revoke" transaction to indicate that a previous correlation made by a partner SHOULD BE revoked.

5. Patient Discovery Responding Gateways MAY accept the XCPD "revoke" transaction and, if they do, they MUST revoke the correlation.

708 **Comment Submission Form:**
709 **Cross-Organizational Patient Matching Minimal Acceptable**
710 **Principles**

711 *Table 5: Cross-Organizational Patient Matching Minimal Acceptable Principles*
712 *public comment response form*

Comment Number	Referencing Document Line Number	Original Text	Suggested Revised Text	Comments
1				
2				
3				

CHAPTER 5: IN CLOSING

Many organizations are dedicated to implementing the highest possible quality patient matching. At the same time, many of those same organizations have (or will) experienced significant and unacceptable error rates when matching patients across organizational boundaries. The enclosed case study, proposed maturity model, and proposed minimal acceptable principles are meant to spur constructive dialogue through the public comment process.

It is your engagement with the process that will determine the strength of the final, adopted proposals and, ultimately, the increased accuracy, usefulness, and consistency of patient identity matching and health data sharing across the country. Please join us in helping to create a framework that can be freely used by all to help assess and improve patient matching.

APPENDIX

Term Definitions

Error Rate: The combined rate of incorrect negative matches, and incorrect positive matches.

HL7: Health Level 7, an international standards development organization.

IHE: Integrating the Healthcare Enterprise, an international standards development organization.

Initiating Gateway: A system sending outbound XCPD and XCA requests.

Match Rate: The combined rate of correct negative matches, and correct positive matches.

Responding Gateway: A system receiving inbound XCPD and XCA requests.

Temporary values: (sometimes known as a default value) is a trait associated with a patient that is known to be incorrect due to lack of information. Temporary values are often created when clinical IT systems require that a value be entered even if that value is not available, such as if a patient has not been identified. Temporary values are often short-term in nature. Examples of temporary values are an SSN of all 1s (111-11-1111) or a newborn name of "Baby Jones." A pseudonym is not a temporary value as it is intended to be a substitute patient identifier with a specific purpose, such as to protect the privacy of a public figure.

XCA: IHE Cross-Community Access is an international standard in Final Text status. XCA is focused on standards-based sharing of clinical documents.

XCPD: IHE Cross-Community Patient Discovery, which is an international standard in Final Text status. XCPD is a standards-based method of discovering mutually known patients between different communities.