EMPOWERING PEOPLE TO IMPACT HEALTH

A Realistic Approach to Achieving a 1% Duplicate Record Error Rate

AHIMA
American Health Information Management Association®

STAFF

Julie A. Pursley, MSHI, RHIA, CHDA, FAHIMA
Director, Health Information Thought Leadership
AHIMA

Lesley Kadlec, MA, RHIA, CHDA
Director, Policy & State Advocacy Engagement
AHIMA

Lauren Riplinger, JD
Vice President, Policy & Government Affairs
AHIMA

Mary Butler
Senior Editor, *Journal of AHIMA*
AHIMA

ABOUT AHIMA

The American Health Information Management Association (AHIMA) is a global nonprofit association of health information (HI) professionals. AHIMA represents professionals who work with health data for more than one billion patient visits each year. AHIMA's mission of empowering people to impact health drives our members and credentialed HI professionals to ensure that health information is accurate, complete, and available to patients and clinicians. Our leaders work at the intersection of healthcare, technology, and business, and are found in data integrity and information privacy job functions worldwide.

ACKNOWLEDGEMENTS

AHIMA would like to acknowledge and thank members for the patient identification and matching (PIM) roundtable listed in **Appendix A** who contributed their time and expertise to the development of the recommendations included in this report.

INTRODUCTION

Quality health information (HI) is fundamental to the provision of safe and effective patient care[1]. In other words, data collected during the course of a patient's journey through the healthcare system must be accurate, timely, relevant, valid, and complete to ensure the reliability and overall integrity of the information.[2] For years, these key principles have been the foundation of health data quality management and remain relevant today and are critical to our healthcare infrastructure.

Patient identification and matching (PIM) is vital to delivery of safe and efficient care, and **misidentification errors have been a recurring challenge in healthcare, resulting in administrative inefficiencies, serious injuries, and even death.** A 2016 National Patient Misidentification Report noted that 86 percent of respondents said they have witnessed or know of a medical error that was the result of patient misidentification.[3] **Healthcare organizations' revenue cycles also face challenges associated with misidentification, costing the average healthcare facility $17.4 million per year in denied claims and lost revenue.**[4]

AHIMA recommends a cycle approach to achieve a 1% duplicate record error rate.

The COVID-19 pandemic has further illuminated the undeniable and overwhelming need for accurate, available, and trustworthy health information as well as the need for strong governance practices. Recently, it has been estimated that 40 percent of patients' demographic data are missing from commercial laboratory testing for COVID-19.[5] Contact tracers rely on accurate and comprehensive data to locate patients, and public health reporting relies on consistent data that is reliable, and reproducible by different users across applications. A widespread and safe vaccination program also demands a reliable and accurate means of identifying individuals.

Technology to address the root cause of data integration and searching across multiple disparate databases is not yet in widespread use in healthcare. For that reason, AHIMA believes a standardized national approach to patient identification is paramount.

However, a uniform, national approach to PIM could take time due to ongoing policy considerations. Since 1999, Congress has prohibited the US Department of Health and Human Services (HHS) from spending federal dollars to promulgate or adopt a national unique patient identifier (UPI). Narrow interpretation of the language creates a barrier to public-private sector collaboration in advancing a nationwide patient identification strategy.

AHIMA continues to advocate to remove this ban to enable the industry and the HHS to advance an end-to-end solution.[6]

1. Davoudi, Sion et al. "Data Quality Management Model (2015 Update) – Retired." *Journal of AHIMA* 86, no.10 (October 2015): expanded web version. https://library.ahima.org/PB/DataQualityModel.
2. Ibid.
3. Imprivata. 2016 National Patient Misidentification Report. https://www.imprivata.com/patient-misidentification.
4. ECRI Institute. "Patient Identification: Executive Summary." August 2016. https://www.ecri.org/Resources/Whitepapers_and_reports/PSO%20Deep%20Dives/Deep%20Dive_PT_ID_2016_exec%w20summary.pdf.
5. Cidon, Dan. "Why Patient Identity Management is Critical for COVID-19 Surveillance." *Journal of AHIMA*. June 1, 2020. https://journal.ahima.org/why-patient-identity-management-is-critical-for-covid-19-surveillance.
6. Ibid.

In the meantime, there are proven people, processes, and technology approaches[7] that HI professionals use that could be leveraged to achieve and maintain a low duplicate record error rate.

PATIENTS AND PROVIDERS NEED ACCESS TO QUALITY HEALTH DATA

Collecting patient demographics is the starting point to trusted, reliable data throughout the patient's health journey.

Regardless of where a patient receives their care, patient demographic data quickly cascades through a web of connected and potentially disparate electronic systems—whether internal or external to the healthcare entity. Matching errors can lead to a separation of a patient's visit from prior visit documentation. Because of this, record components such as diagnoses, testing, allergies, medications, contact information, social history, family history, and advance directives may not be inextricably linked to the patient's unique health record, which could inhibit a provider's ability to access critical information. In turn, this may result in improper care, additional testing, or jeopardized patient safety based on inaccurate patient information made available to the provider.

7. Butler, Mary. "The HIM View for 2020." *Journal of AHIMA*. January 1, 2020. https://journal.ahima.org/the-him-view-for-2020.

COMPOUNDING EVENTS LEADING TO MISIDENTIFICATION

This scenario highlights compounding events leading to patient misidentification and matching.

Anne Anderson, DOB 2/12/70 is asymptomatic but has comorbities – she schedules a lab test at a COVID-19 drive through testing event.

Anne is provided a registration link to the laboratory's patient portal.

Anne enters requested data: first name, last name, DOB, address, gender, and cell phone number.

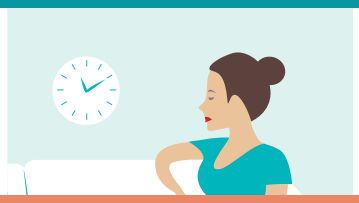
Anne is notified by the laboratory that her results were sent to the healthcare organization.

Two days later, the postive test results are sent to Anne's healthcare organization.

Anne's test results from the laboratory are unfortunately merged with the health record of **"Ann Andersen, DOB 12/2/70."** The two patient's results are now comingled within one health record.

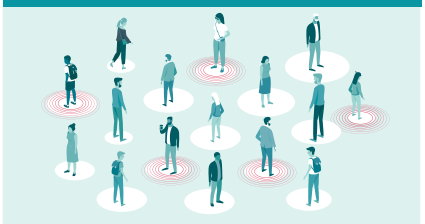
Due to the mix up, **Ann Andersen** is contacted by her healthcare provider. They inform her of the positive COVID-19 test result.

Ann Andersen informs the care provider that she has not had a test and that a mistake has been made.



In the meantime, **Anne Anderson** has been told her results were sent to her care provider, but she has not received a verifying call.

Anne erroneously believes that she does not have a positive COVID-19 test result. Anne is unkowningly spreading the virus at her essential workplace.



Because Anne Anderson's contacts have not been notified through contact tracing, they may have been exposed and could be infecting others, creating community spread.

Other examples of patient identification mix-ups can be found at https://patientidnow.org/impact/

DEFINING DUPLICATES, OVERLAYS AND OVERLAPS

Duplicates, overlays, and overlaps exist in most health information technology systems and unfortunately can jeopardize patient safety.

Duplicate:

- *A duplicate record is created when two or more medical record numbers are created for the same person, causing them to have two or more records.*[8]
- *Example: Patient registration does not find the patient's record using the information provided by the patient. A new medical record number is generated, creating a duplicate.*

Overlay:

- *An overlay occurs when the incorrect patient is registered, admitted, or documented on another patient's record.*[9]
- *Example: The database inadvertently merges John Clark, DOB 10/5/81, with the record belonging to John Clark, DOB 5/10/81. Both individuals' information is now in the same record, comingled.*

Overlap:

- *An overlap occurs when there is more than one UPI for the same person across two or more facilities in the enterprise.*[10]
- *Example: Hospital A acquires Hospital B through a merger/acquisition. Overlaps are created when records from the two different facilities are not connected at an enterprise level.*

ACHIEVING AND MAINTAINING A 1% DUPLICATE ERROR RATE

In AHIMA's 2020 Patient Identification Survey, 22 percent of respondents reported they achieved a 1 percent or less duplicate error rate in their electronic health record (EHR). But how does an organization get there? A 1 percent duplicate record error rate is achievable through a structured cycle where dedicated resources, time and effort are prioritized and supported. Organizations that achieve and maintain a 1 percent duplicate error rate demonstrate a commitment to patient safety while supporting enhanced and seamless access to health information.

While achieving a 1 percent rate will require dedication and effort, it is important to recognize that this minimal rate still represents wrong-patient errors.[11] In other words, an organization with 500,000+ patients in their master patient index/enterprise master patient index (MPI/EMPI) that maintains a 1 percent duplicate error rate still leaves 5,000 patient lives' at risk for potential misdiagnosis, duplicate testing, and/or medical errors.

Another significant finding of the survey indicated that 29 percent of survey respondents were not aware of their duplicate error rate, which underscores the need for organizations to prioritize and address PIM.

RECOGNIZING VARIABILITY IN PIM

AHIMA recognizes industry variability in PIM methods and processes. For that reason it is important to identify and understand how database environments, duplicate error and creation rate calculations, workforce factors, and the types of patient matching algorithms that play an important role in achieving and maintaining a low duplicate record error rate.

---

8. Harris, Shannon and Shannon H. Houser. "Double Trouble—Using Health Informatics to Tackle Duplicate Medical Record Issues." *Journal of AHIMA* 89, no. 8 (September 2018): 20–23. http://library.ahima.org/doc?oid=302567.
9. Landsbach, Grant. "Study Analyzes Causes and Consequences of Patient Overlay Errors." *Journal of AHIMA* 87, no.9 (September 2016): 40-43. https://bok.ahima.org/doc?oid=301860.
10. AHIMA Work Group. "Managing the Integrity of Patient Identity in Health Information Exchange (2014 update)." *Journal of AHIMA* 85, no.5 (May 2014): expanded web version. https://library.ahima.org/PB/PatientIdentityHIE.
11. Lyons, Maureen. "New Joint Commission advisory on accurate patient identification." October 2, 2018. https://www.jointcommission.org/en/resources/news-and-multimedia/news/2018/10/new-joint-commission-advisory-on-accurate-patient-identification.

MPI/EMPI Database Environments

SINGLE
A single MPI is typically found where the registration system and the EHR system are combined into one database.

MULTIPLE
Two or more MPI databases (e.g. EHR and registration) with potential for overarching EMPI database. Example: One registration MPI, one EHR MPI, and an overarching EMPI for the health system.

COMPLEX
Multiple MPI/EMPI databases to manage a regional or national infrastructure.

DUPLICATE ERROR AND CREATION RATE CALCULATIONS

Error rates differ depending upon how an organization's technology calculates it. While some technologies offer quality and benchmarking reports to manage the MPI/EMPI, others do not. The same holds true for calculating a duplicate error and creation rates.

To accurately calculate the duplicate record error rate and creation rate, AHIMA recommends the following calculations for a single MPI database:

Duplicate Record Error Rate:

- Total number of confirmed duplicate records divided by the total number of patient records in the MPI database.
- Example: 5,000 duplicate pairs (two records) that involve 10,000 individual records. Database contains 100,000 individual records: 10,000/100,000 = 1% duplicate error rate.

Creation Rate:

- Total number of confirmed duplicate records for a defined time period divided by the total number of registration events within the MPI during the same time period.
- Example: 3,000 duplicate patient records were confirmed in the third quarter. There were 200,000 registration events within the time period: 3,000/200,000 = 1.5% creation rate.

WORKFORCE FACTORS

Iterative PIM training is critical for the entire workforce. Errors can occur anywhere in the patient's journey from initial contact to post-encounter. Every staff member using the health record needs awareness training in order to identify, prevent, and resolve PIM errors.

Staffing may or may not be adequately equipped to handle the volume of work needed to achieve and maintain a low duplicate error rate. While some organizations may have one or two dedicated staff members, others inlcude a centralized data integrity team. According to AHIMA's 2020 patient identification survey, 27 percent of respondents cite lack of resources to correct duplicates as a challenge they face in managing their respective MPIs/EMPIs.

TYPES OF PATIENT MATCHING ALGORITHMS[12]

Algorithms are mathematical formulas that combine weighted data elements to determine the probability of a duplicate in order to identify potential duplicate entries. They are a critical component of any successful MPI/EMPI solution.[13]

Algorithms contained within the database environment differ according to an organization's technology and build. Edits or modifications to the matching algorithms may also not be allowed due to a system's proprietary technology. Common matching algorithms include:

Deterministic: Involves a unique identifier, sometimes coupled with a limited number of nonunique identifiers, such as DOB for additional validation, that are compared to identify exact matches. Considered a basic record matching algorithm. Comparisons are usually made based on name, DOB, Social Security Number (SSN) and sometimes gender.

Rules-based: Each data element receives a "weight" for how essential it is to match a record. Even if every data element does not match exactly, the records are considered "matched" so long as enough data elements are identical. For example, records are considered matched if first name, last name, DOB, and gender match or if last name, address, and DOB match.

Probabilistic: Compares several (nonunique) field values between records, assigning a weight to reflect how closely the two-field values match. The weights are then added across the fields to indicate the probability of an actual match. May be considered an intermediate or advanced algorithm.

ORGANIZATIONAL GOALS NECESSARY TO ACHIEVE AND MAINTAIN DATA INTEGRITY

Earlier this year, AHIMA convened a roundtable of HI professionals who work with MPI and EMPI databases on a daily basis. The roundtable participants serve as part of a data integrity team which fulfills the critical role of **identifying, cleaning, measuring, mitigating and remediating** patient demographic data within a multitude of HI systems while educating and reinforcing the importance of linking each patient to their unique health record.

Participants included individuals from various US regions and care settings that manage data in seven different EHRs. As part of the roundtable discussion, ***the workgroup identified several impact areas to address PIM including: governance and leadership, data collection, and data integrity.***

An organizational goal checklist was created around the above impact areas.

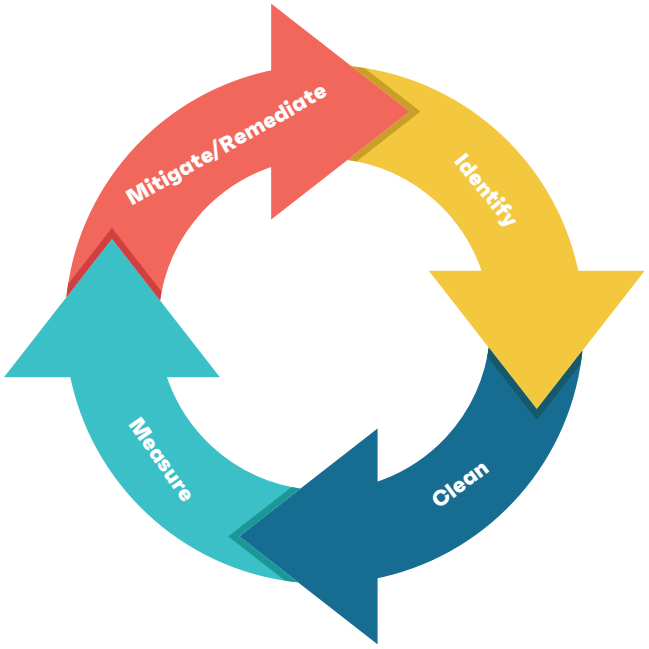
- This resource lists *organizational goals* within three areas
  - Governance and leadership
  - Data collection
  - Data integrity
- The *organizational goal* is an enterprise's overarching goal
- *Health information leadership initiative(s)* are key action items that an HI leader must lead to achieve the *organizational goal*

12. Riplinger, Lauren. "It's Not Another Number: Technical Strategies to Accurately Match Patients with Their Data." http://patientidnow.org/wp-content/uploads/2020/06/Alternatives-to-UPI.pdf.
13. AHIMA. "Fundamentals for Building a Master Patient Index/Enterprise Master Patient Index (2010 update)." *Journal of AHIMA* (Updated September 2010). http://library.ahima.org/doc?oid=106227.

ICMMR CYCLE

AHIMA recommends the "ICMMR cycle" approach to achieve a 1 percent duplicate record error rate.

The *ICMMR Cycle* includes Identifying, Cleaning, Measuring, Mitigating and Remediating. This PIM data lifecycle must be performed iteratively. Provided below are a set of sample considerations for each cycle element. It recommended that organizations use the **"organizational goal checklist"** as a companion piece to the cycle in creating a tailored plan for their organization in achieving a 1 percent duplicate record error rate.



Identify:

- Assess the current state of duplicates in the MPI/EMPI by identifying and running all available duplicate reports
- Run a report showing the number of potential duplicates in the MPI/EMPI database
- **Calculate the initial duplicate record error rate percentage**
- If there is a lack of confidence in these results, perform a data analysis of the MPI/EMPI database which will provide a higher level of detail on the number and types of potential duplicates
- Proactively meet with MPI/EMPI technology vendors to understand how their **patient matching algorithms** function and are calculated.

Clean:

- Create a plan for cleaning the database to reduce the number of duplicates
- Explore cleaning the database using internal/external technology partners and/or data integrity staff resources
- While cleaning, ensure that multiple records with the same medical record number are identified and resolved at the same time

Measure:

- Run duplicate report to identify number of potential duplicates in the MPI/EMPI database and calculate duplicate record error rate percentage
- Benchmark against the duplicate record error rate quarterly
- Set stretch goals
- Allow and plan for disruption
- Implement grace periods for situations where data is ingested into the MPI/EMPI from outside sources due to a merger/acquisition or other data integration need
- Complex bidirectional interfaces can increase the duplicate creation rate; monitor closely

Mitigate/Remediate:

- Iterative staff education and training to identify, prevent, and resolve errors
- Incorporate processes for front-end and back-end database functions including, but not limited to, naming conventions, feedback loops, and duplicate prevention
- Merging record workflows should underscore the importance of accuracy first
- Operationalize stringent, daily process for working the duplicate error queue
- These steps must be concurrent with cleaning the database

CONCLUSION

Recognizing the importance of a low duplicate error rate and the fact that an end-to-end PIM solution is still on the horizon, it is more important than ever for HI professionals to have the tools and resources to accurately identify and uniquely match patients to their health information.

Despite variability in PIM methods and processes—coupled with the importance of identifying and understanding existing database environments, duplicate error and creation rate calculations, workforce factors, and types of patient matching algorithms-achieving and maintaining a 1 percent duplicate record error rate is feasible in today's healthcare environment. Organizations should strive to meet the organizational goals in the checklist by using the ICMMR Cycle to create a tailored plan.

APPENDIX A: ROUNDTABLE PARTICIPANTS

AHIMA would like to thank and acknowledge the following roundtable participants who contributed their time and expertise to the development of this white paper.

Keely N. Aarnes, PMP

Ermeise Baker, RHIT, CCS-P

Laritha Boone, MBA-HCM, RHIA

Patricia Campola, RHIA

Shonda S. Cannon

Necole Coots, RHIA

Michele D'Ambrosio, MBA, RHIA

Sally C. Deming, RHIT

Marlenda Jackson Dusz

Brandi Hunsaker, RHIT

Riana Jackson, RHIT

Melissa Paige Key, MSM, RHIA

Doreen Koch, RHIA

Michelle Layton

Lee Luftman, RHIA

Susan H. Mason, RHIT, CCS

Wendi Melgoza, MSHIM, RHIA, CPHI

Sharon Meyer, RHIA

Zianna Lei Peters, RHIA

Linda D. Rader, RHIT

Ashley Running

Janice Dansby Washington, RHIA

Terry M. Walker, RHIT

Rebecca Way, RHIA