Patient Identification Techniques — Approaches, Implications, and Findings

Lauren Riplinger¹, Jordi Piera-Jiménez², Julie Pursley Dooling³

¹ AHIMA, Washington DC, USA
² AHIMA International, Barcelona, Spain; Open Evidence Research Group, Universitat Oberta de Catalunya, Barcelona, Spain
³ AHIMA, Chicago, IL, USA

Summary

Objectives: To identify current patient identification techniques and approaches used worldwide in today’s healthcare environment. To identify challenges associated with improper patient identification.

Methods: A literature review of relevant peer-reviewed and grey literature published from January 2015 to October 2019 was conducted to inform the paper. The focus was on: 1) patient identification techniques and 2) unintended consequences and ramifications of unresolved patient identification issues.

Results: The literature review showed six common patient identification techniques implemented worldwide ranging from unique patient identifiers, algorithmic approaches, referential matching software, biometrics, radio frequency identification device (RFID) systems, and hybrid models. The review revealed three themes associated with unresolved patient identification: 1) treatment, care delivery, and patient safety errors, 2) cost, and resource considerations, and 3) data sharing and interoperability challenges.

Conclusions: Errors in patient identification have implications for patient care and safety, payment, as well as data sharing and interoperability. Different patient identification techniques ranging from unique patient identifiers and algorithms to hybrid models have been implemented worldwide. However, no current patient identification techniques have resulted in a 100% match rate. Optimizing algorithmic matching through data standardization and referential software should be studied further to identify opportunities to enhance patient identification techniques and approaches. Further efforts to improve patient identity management include adoption of patients’ photos at registration, naming conventions, and standardized processes for recording patients’ demographic data attributes.

Keywords

Patient identity, patient matching, identification errors, patient identification systems, patient records linkage

1 Introduction

Patient identification is the process of “correctly matching a patient to appropriately intended interventions and communicating information about the patient’s identity accurately and reliably throughout the continuum of care” [1]. Patient identification encompasses not only the physical identification of the patient but also technologies able to enhance the accuracy of patient identification [1]. The main attributes of ideal patient identifiers have been described as unique, ubiquitous, and unchanging in their nature [2]. Different technological approaches and practices and operational processes that optimize accurate patient identification are necessary to meet the increasing and diverse demands for the use and reuse of data by various stakeholders [3].

Today, lack of widely adopted operational principles and limitations in processes and technologies result in inaccurate patient identification [1]. Lack of accurate patient identification can affect clinical decision-making, treatment, patient outcomes, patient privacy, and results in duplicative testing and increased costs [4-7]. When a patient is incorrectly matched to another patient’s record, patient care and safety are jeopardized as incorrect data can cascade to a multitude of internal and external systems and databases such as laboratory, radiology and health information networks [8], potentially leading to laboratory, imaging, and medication errors as well as wrong-site surgeries. Similarly, with the existence of multiple records for a single patient, clinicians can miss critical information because it is in the duplicate record [9]. In both scenarios, care decisions are based on an incomplete or erroneous picture of the patient’s medical history because data is not accurate or reliable.

Healthcare organizations vary in how they collect and identify patients. Globally, some countries use unique patient identifiers (UPIs) to assist with patient identification while the United States (US) prohibits the use of a national UPI due to privacy and security concerns [9].

As the volume, velocity, and variety of health data continue to grow, accurately matching patients to their health data is a challenge that must be met to ensure care coordination across the continuum, enhance patient safety, ensure the appropriate use of resources, and foster successful data sharing and interoperability [10, 11].

This paper explores current patient identification techniques used worldwide and the problems and challenges associated with unresolved patient identification and it offers recommendations for further study.

2 Objectives

• To identify the current patient identification techniques and approaches used worldwide;
• To identify challenges associated with reliable patient identification.

3 Methods

To inform the paper, the authors conducted a literature review for relevant peer reviewed and grey literature articles focusing on: 1) patient
identification techniques adopted worldwide and 2) consequences and implications of unresolved patient identification issues published from January 2015 to October 2019 in Scopus, PubMed, Web of Science, the Association for Computing Machinery (ACM) Digital Library, and the American Health Information Management Association (AHIMA) Body of Knowledge. Searches for grey literature (i.e. evidence not published in commercial publications, including but not limited to government publications, conference papers, and white papers) [12] on the topic were conducted using Google. Articles for the review were drawn from peer-reviewed journals, conference papers, consumer studies, health professional studies, research performed by independent research institutions, as well as systematic and narrative reviews of the various topics. A snowball approach was undertaken to identify additional sources of information.

4 Results

4.1 Patient Identification Approaches and Considerations

Ensuring quality health information has and will increasingly become complicated as new data streams are utilized and as organizations share records electronically using different information systems and standards [13, 14]. Worldwide, there are several different patient identification approaches, techniques, and solutions including UPIs and algorithms. At the same time, technological and methodological innovations have introduced new approaches such as referential matching, biometrics, and radio frequency identification device (RFID) technologies as ways to further improve techniques for patient identification [4].

Below we examine a variety of techniques utilized to address patient identification.

4.1.1 Unique Patient Identifiers (UPIs)

Many healthcare systems outside of the US have adopted nationwide initiatives to implement UPIs. Such approaches are common in Europe including England, Wales and the Isle of Man [15], Denmark [16, 17], Estonia [18, 19], and more recently Spain [20] and Ireland [21]. Outside of Europe, UPIs are common in such countries as New Zealand [22], China [23], and Israel [24], all of which use their National Identifier (ID) as their patient identification number. However, even though UPIs are widely implemented and constitute a preferred method due to reduced reliance on patient attributes for patient matching, the challenges of generating and implementing UPIs often lead to limited implementations within institutions compared to other care settings [4, 14, 25, 26]. Privacy concerns are also often cited as a major concern in implementing UPIs [25].

In Singapore, the National Registration Identity Card—the UPI identified for use—was not sufficiently ubiquitous to be used as an identifier as 28% of the Singaporean population are non-residents and therefore not required to have an identity card [34]. In Canada, because healthcare is funded and governed at the province-level and every province has different regulatory frameworks, the deployment of a nationwide UPI proved difficult [34]. As a result, patient registries, which support the centralized storage and retrieval of patient identification data, are used to identify patients in conjunction with an enterprise-wide master patient index (MPI), data quality remediation, and governance policies [15].

In the US, long-standing policy barriers have hindered adoption of a unique patient identifier. In 1996, the Congress passed legislation calling for the adoption of a UPI [27]. However, due to privacy and security concerns, the Congress included statutory language in an annual appropriations bill limiting the promulgation or adoption of a UPI [28]. Since then, the statutory language has remained and hampered the US Department of Health and Human Services (HHS) from engaging in rulemaking to adopt a UPI standard. Despite this ongoing limitation, the Congress recently included report language in the 2020 spending agreement encouraging HHS to provide technical assistance to private sector-led patient identification initiatives and directed the Office of the National Coordinator for Health Information Technology (ONC) to 1) submit a report to the Congress on the technological and operational methods that improve identification of patients [29], and 2) recommend actions increasing the likelihood of accurately matching patients to their health data. Inclusion of the report language may continue to help foster private sector-led initiatives focused on patient identification and help identify and implement important policy levers to further advance a nationwide patient identification strategy in the US.

4.1.2 Algorithmic Approaches

Algorithms are another common approach to matching patients to their health information using demographic characteristics including: first name, last name, gender, date of birth, social security number (in the US), and address [30]. Algorithms range from basic, i.e., deterministic matching involving a unique identifier coupled with a limited number of non-unique identifiers (e.g., date of birth) that are compared to identify exact matches, to more sophisticated probabilistic matching techniques that use threshold limits [7, 31]. The complexity of such algorithms varies widely, may be vendor-specific and dependent on the customization added to the base installation of an electronic health record (EHR) system [32]. While matching algorithms can achieve matching rates approaching 90%, they are not perfect and do not represent a 100% accurate patient matching solution [30, 33, 34].

Inaccurate or incomplete patient demographic information can hinder the enhancement of algorithmic accuracy [35]. Occurrences of inaccurate or incomplete demographic data can be the result of: lack of best practices in collecting demographic data at registration, variations in organizational and health information technology (IT) vendor policies and processes in how demographic data are collected, failure by patients to provide the correct information at registration (e.g., register with nickname versus legal name), and failure to update patients’ information when their demographic information changes (e.g., address, phone, email, and name change) [36]. Transcription and free text errors can also limit algorithm accuracy including spelling variations, phonetic variations, double last names, double first names, and alternate first names [37, 38].
Standardized data elements are also generally needed to optimize matching algorithm accuracy (e.g., telephone number, date of birth, address) [39, 33]. Lack of standardized demographic data elements across institutions and care settings can exacerbate the “syntactic heterogeneity of data,” making matching algorithms less effective [32, 33]. Evidence suggests that standardizing certain demographic data elements could improve match rates [33].

4.1.3 Referential Matching Software
Beyond algorithmic methodologies, a growing number of organizations are implementing add-on technologies including referential matching software to increase odds of identifying patients correctly. Referential matching software is a data augmentation where a third-party service provider adds an additional layer of demographic data (typically from outside of healthcare) including datasets from credit reporting and public utilities that are routinely updated and maintained to enhance patient matching [31]. Companies involved in such technologies report match rates as high as 98 and 100%, however such rates have not been independently verified [31, 40]. Concerns have been raised that referential matching could lead to clinicians and payers having access to personal and financial information, including credit information. However, existing referential matching approaches do not appear to share patient health information outside of the healthcare institution [31].

Accuracy of non-health information such as data from the US Postal Service and the Social Security Administration used by matching software is also a concern for clinicians and patients [31]. Referential matching also has limitations related to certain patient populations including children, homeless individuals, and undocumented immigrants because data sources used for referential matching do not contain or have limited information on these populations [31].

4.1.4 Biometric Identification Systems
Limitations with UPIs and matching algorithms have led to the exploration of other add-on technologies including biometric identification technologies. Such approaches include fingerprints, palm vein scanning, iris scanning, and facial recognition [41]. Biometric identification technologies are advantageous because biometrics data is more difficult to “steal, exchange or forget” [25]. Biometric technologies are increasingly being used to identify patients in the US [2]. However, they are not without limitations. Research suggests that biometric methods to date for infant identification including eye scanning, ear and face recognition, and finger and palm-based methods solutions are not as effective as with adults since such features are difficult to capture and subject to change during child development [42, 43]. In Europe, biometric technologies also raise ethical and legal considerations [44]. The General Data Protection Regulation (GDPR) identifies biometric information as a “special” category of personal information merit ing protection that is explicitly prohibited from use for the purpose of identifying an individual unless for medical diagnosis or in the treatment or management of healthcare provider certain conditions are met [45].

4.1.5 Radio Frequency Identification
Additional emerging technologies, like radio frequency identification (RFID) are under exploration by hospitals to enhance patient identification. Unlike existing barcode technologies, RFID labels can hold more data than barcodes and be read automatically without user intervention [46]. Additionally, RFID technologies offer re-writeable functionality, allowing information to be modified, and require “no line of sight,” eliminating the need for human intervention [46]. RFID technologies also offer more advanced forms of data security than barcode technology, including encryption, allowing for patients’ health data to be kept more secure [46]. At the same time, application of RFID technology is not widespread as it can be cost-prohibitive [46], and lacks standards or guidelines for implementation within healthcare. RFID technologies also raise unique privacy concerns including the wireless transmission of patients’ health information after completion of healthcare services and collection of patients’ health data by third-party actors without patients’ approval or knowledge [47].

4.2 Discussion on Patient Identification Methods
The variety of patient identification methodologies is expansive and may include hybrid models that combine different methods described above. From the basic to the most complex methods, researchers agree that no perfect patient linkage solution exists and all the approaches present challenges that must be addressed [14, 25, 31].

In the US, these challenges are further hampered by existing legal and policy impediments. Lack of a perfect patient identity solution raises significant concerns. The literature review revealed three distinct themes associated with unresolved patient identification issues. The first theme focuses on how patient misidentification can lead to clinical errors or “near misses” which raises patient safety concerns. The second theme identifies financial, payment, and resource implications associated with patient misidentification. The third theme identifies the limitations patient misidentification places on data sharing and interoperability. Each theme is examined below.

4.2.1 Patient Safety
Failure to accurately identify patients raises patient safety and quality of care concerns across the care continuum from diagnostic testing to treatment [48]. A 2016 study classified over 7,600 out of 10,915 events from January 2013 to August 2015 as “wrong-patient events” involving patient identification [1]. Factors contributing to these events included admitting a patient under another patient’s medical record, creating a duplicate record at registration [49], pulling the record of a patient with a name similar to the intended patient, or asking about patient’s identity while failing to either check a patient’s identification band or ask for two acceptable forms of identification. More recently, the Joint Commission’s sentinel event statistics indicate 37 patient safety events out of 436 sentinel events have occurred due to patient identification errors in the second quarter of 2019 including surgical or invasive procedure events involving wrong patient, wrong procedure, and wrong site [50]. Failure to
accurately match patients to their health information can lead to “false positives” when the medical records of two different patients are mistakenly matched [35] or “false negatives” when the records for the same patient are not matched. Duplicates are the result of several factors including varying methods of matching patient records, lack of data standardization, lack of policies and procedures, and frequently changing demographic data [39].

The existence of duplicate records can also lead to duplicative testing and treatments because of inaccurate or unavailable data. For instance, one hospital reported 30% of clinicians surveyed reordered tests because of lack of access to previous records [51].

4.2.2 Financial and Resource Concerns
Challenges associated with accurately matching patients to their health information raise financial and resource concerns. Repetitive tests and treatments are likely to add costs and impact timeliness of care delivery. From a revenue cycle perspective, there may be claims denials [52, 53] and implementation of time consuming and costly processes to correct medical records [51]. Research indicates that between 10 to 15% of all health insurance denials are due to incorrect patient identification numbers [13].

4.2.3 Data Sharing and Interoperability
Patient misidentification raises data sharing and interoperability concerns as well. Historically, patient identification practices in the US have been fragmented and inconsistent [32]. Furthermore, hospitals report that difficulties in accurately matching patients to their health information across health IT systems limit health information exchange [54]. In turn, limited health information exchange can lead to clinical decision-making based on incomplete information resulting in increased chances of misdiagnosis, unsafe treatment, and duplicative testing [55].

Accurate patient identification may also have downstream implications for the secondary use of data for population health, quality improvement, public health, research and detection of waste, fraud, and abuse [11, 56].

5 Some Recommendations
As new data streams are increasingly integrated into the clinical setting, unique, ubiquitous identifiers will become increasingly necessary. Patient identity management challenges include accurately matching patients within and across organizations including for research and clinical trials. To address the shortcomings of both current and emerging approaches, we offer the following recommendations:

**Recommendation 1: Standardization of Data Elements**
Additional research and study should be undertaken by non-profit, non-government institutions to identify and evaluate data elements (such as demographic data) that could further optimize the accuracy of matching algorithms. For example, in the US, recent suggestions have included adoption of US Postal Service certified address standardization rules [57] and patients’ cell phone numbers [34, 56].

**Recommendation 2: Evaluate Effectiveness of Referential Matching Approaches**
Additional study and research are needed to evaluate the effectiveness of referential matching in improving algorithmic matching rates. Such research should include an analysis of improvement rates when using such software as well as identifying potential gaps in the use of such technology among certain patient populations (e.g., pediatrics, homeless, undocumented immigrants, etc.). Research should also include an examination of potential privacy risks to patients with the use of such technologies. Other areas for additional study include further study of RFID technologies and the potential use of big data analytics approaches (such as using Fuzzy algorithm (Levenshtein Distance) and MapReduce techniques) [3] for patient matching involving large repositories of data for better clinical decision-making.

**Recommendation 3: Guidance on Preventing Adverse Events Related to Patient Identification Errors**
Further steps can be taken today as regard to existing processes and practices to address patient identification errors. Practices that

6 Conclusion
Accurate and unique identification of patients along the care continuum is essential for patient care and safety, addressing cost and resource concerns, and enhancing data sharing and interoperability. Patient identification techniques ranging from UPIs to algorithms to biometric identification have been implemented worldwide—each accompanied by their own set of opportunities and challenges and resulting in no single solution with a 100% match rate. The volume, velocity, and variety of health data is expected to continue to grow, as is demand for new data streams to be incorporated into the electronic health record. The need to link electronic records, exchange and share data, and achieve interoperability is escalating. Without unique, unambiguous identifiers, the ability to merge new data streams into the medical record will become increasingly difficult. Opportunities exist for researchers and clinicians to play a role in enhancing existing and emerging approaches to patient identification.

References
2. Ranade-Kharkar P, Pollock SE, Mann DK, Thornton SN. Improving Clinical Data Integrity by using Data Adjudication Techniques for Data Received through a Health Information Exchange (HIE).

Correspondence to:
Lauren Riplinger, JD
AHIMA
Washington DC
USA
Tel: +1 202 839 1218
E-Mail: lauren.riplinger@ahima.org

more physical and biological traits: Research into the characteristics that are unique to an individual is addressing the need to correctly identify people in a variety of medical, social and security context. EMBO Rep 2016 Jan;17(1):22-6. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26666447

Riplinger et al.

Lauren Riplinger, JD
AHIMA
Washington DC
USA
Tel: +1 202 839 1218
E-Mail: lauren.riplinger@ahima.org

IMIA Yearbook of Medical Informatics 2020