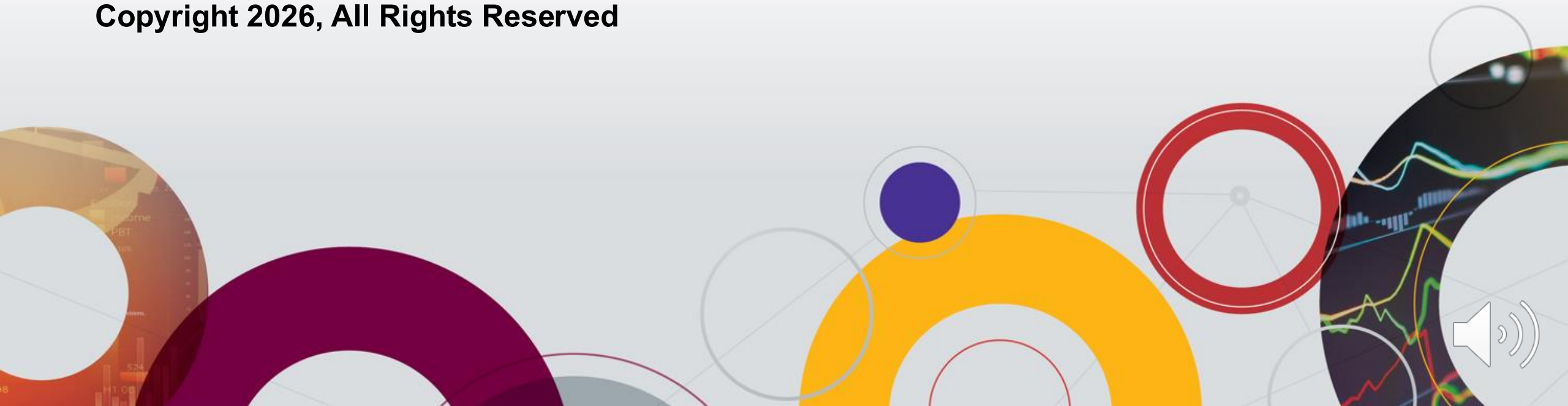


4.2 Identifier and Transaction Standards

10x10 – Introduction to Biomedical Informatics & Artificial Intelligence
William Hersh, MD, FACMI, FAMIA, FIAHSI
Copyright 2026, All Rights Reserved



Identifiers

Various approaches exist (or have been proposed) for

- Patients
- Providers
- Employers
- Health Plans

Patient identifiers

Risks

Easy linkage of records

Facilitate health information exchange

Reduce errors and costs arising from duplicate records

Benefits

Easy linkage of records

Potentially compromise privacy and confidentiality

Challenges with inadequate identification in EHRs

Identifier errors compromise quality of care and can be costly (Riplitinger, 2020)

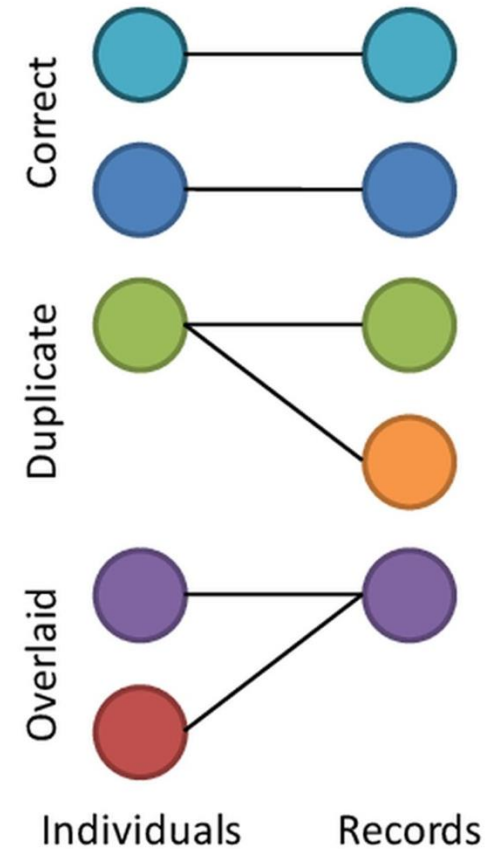
- Time and cost to correct errors from duplicate and overlaid records

Duplicate records more likely to be associated with missed abnormal test results (Joffe, 2009)

Study of five large academic centers found (McCoy, 2013)

- Large numbers of patients with matching names (17-41%), though reduced with other data
- Highly variable policies for preventing, detecting, and removing duplicate records, and for mitigating errors

Improved patient identification better identifies adverse drug reactions (Sragow, 2020)



Patient identifiers – key attributes (Riplitinger, 2020)

Unique – only one person has a particular identifier

Non-disclosing – discloses no personal information

Permanent – will never be re-used

Ubiquitous – everyone has one

Canonical – each person has only one

Invariable – will not change over time

One solution is government-issued patient identifiers

Most industrialized countries have them, e.g.,

- [New Zealand National Health Index \(NHI\)](#)
- Singapore issues National Registration Identity Card (NRIC) to all citizens and Foreign Identification Number (FIN) to all long-term visitors, which are used as identifiers in healthcare
- Ireland adopted as part of national eHealth strategy and implemented in 3 years (Munro, 2016)
- Most other countries with advanced economies use them without controversy, some using general unique identifying number (UIN) and others using specific unique health identifier (UHI) (Mills, 2019)

Government-issued patient identifiers in the US?

HIPAA mandated creation of patient identifiers but public pressure forced postponement

- Congress banned funding for national health identifier in 1998 but repealed it in 2019 (Grannis, 2019; VanHouten, 2021)

Could/should we use the social security number as a national health identifier?

- Technical problems: many duplicates, no check digit for checksum validation (Winkler, 2009)
- Other problems: used for too many other purposes, allows re-identification in public health data sources (Acquisti, 2009)

Alternative to a national identifier – probabilistic matching

Use of probabilistic matching algorithms to link patient records based on various attributes, e.g., name, address, date of birth, phone, etc.

Many methods show relatively high level of accuracy (Grannis, 2003; Tromp, 2011; Sayers, 2016)

- Methods use widely in health information exchange (Kho, 2015)

Research still required for problems with non-standardized (“dirty”) data (Randall, 2013) and missing data (Ong, 2014)

ONC state of patient record-matching (Morris, 2014)

Imperative as a patient safety, care coordination, and data quality issue (among others)

Current state-of-art works well, but would benefit from standardizing patient-identifying attributes in record, such as

- First/given, middle/second given, and last/family names
- Suffix – e.g., Jr./Sr., II/III/etc., MD/RN/PhD, Esq., etc.
- Date of birth – YYYYMMDD, with HHMMSS if available
- Current and historical addresses – in some international format
- Phone number – all known
- Gender – from HL7 value set; M, F, UN

Also need process for handling changes across healthcare system

Advocates further research to evaluate algorithms and additional attributes as well as dissemination of best practices

Recent work in patient matching

Unlikely for two individuals to share same first name, last name, and date of birth in US database of 85 million individuals, with 98.8% of individuals uniquely identified using only these three items (Zech, 2016)

Any three of name, address, date of birth, and gender gave very high rate of matching, and as well as social security number alone (Ansolabehere, 2017)

Most important factor for matching in Indiana HIE was standardizing patient address (Grannis, 2019)

More recommendations going forward

Framework for Cross-Organizational Patient Identity Management (Heflin, 2018)

Patient-Empowered Approaches to Improving Record Matching (RAND, Pew) (Rudin, 2018)

Mandated report from 21st Century Cures Act: *Approaches and Challenges to Electronically Matching Patients' Records across Providers* (GAO, 2019)

Targets for improving patient matching, aggregation, and linking (ONC, 2019)

- Improvements to matching algorithms
- Improvements to data quality
- Expanded data sharing
- Data standardization

Recent initiatives

PatientIDNow

- Framework for national strategy around patient identity
- Aim for confidentiality, integrity, availability (CIA) without compromising privacy and security
- Also aiming for standardization, interoperability, and equity

Project US@

- ASTP/ONC standard for patient addresses
- Rules for formats, e.g., upper-case, acceptable abbreviations, numbers and numerical ranges, diacritical characters, etc.

Element	Description
Business/Firm Name	Only to be used for patient addresses containing businesses
Street Address Line	<PRIMARY ADDRESS NUMBER><PREDIRECTIONAL><STREET NAME><SUFFIX><POSTDIRECTIONAL><SECONDARY ADDRESS IDENTIFIER><SECONDARY ADDRESS>
Last Line	<CITY><STATE><ZIP+4>

Provider identifiers

Universal Physician Identifier Number (UPIN) was maintained by the US government for physicians who treated Medicare patients

National Provider Identifier (NPI) now assigned to all US physicians

- Issued by the National Provider System (NPS), overseen by the Department of HHS
- 10-digit number with last digit serving as check digit
- CMS no longer processes claims without NPI

Employer and health plan identifier standards

Employers – [National Standard Employer Identifier \(EIN\)](#)

Health plans – [Health Plan Identifier \(HPID\)](#) and Other Entity Identifier (OEID) mandated in Affordable Care Act (ACA)

Transactions

ASC X12N standards are designed to encourage electronic commerce for health claims, simplifying current situation of 400+ different formats

- HIPAA mandated use of these standards for healthcare business electronic data exchange
 - “Administrative simplification”
- Original version of HIPAA ASC X12 standards was called version 4010 and was superseded by [Version 5010](#), which had deadline for compliance of January 1, 2012 (Moynihan, 2010)

ASC X12N transactions – Version 5010

Health insurance claims and equivalent encounter information (837)

Enrollment and disenrollment in a health plan (834)

Eligibility for a health plan (request 270/response 271)

Health care payment and remittance advice (835)

Health plan premium payments (820)

Health claim status (request 276/response 277)

Referral certification and authorization (278)

Coordination of benefits (837)



You don't believe these are important?

Ask Former President Obama!

Part of reason for early failure of Healthcare.gov was due to health plans not properly having implemented 834 standard for enrollment and disenrollment in health plans

- News media: “Obamacare’s most important number: 834” (Kliff, 2013)
- More technical description (Laszewski, 2013)
- Many other informatics lessons learned as well (Blumenthal, 2013)

