

Quality of Electronic Health Records: Encountering Misclassification

Dena H Jaffe, Sam Montgomery, and Cheryl Grosso

Acknowledgement: The authors thank Nathan Vavroch, Stacey Purinton, and Rob Taylor for their review and insights.

Background

Electronic health records (EHRs) include encounter records that document the patient-provider interaction and contain information for patient care, billing, and research. The encounter visit describes the location of healthcare service summarized in the broader category of classification, such as inpatient or outpatient, as well as a more specific category of type, such as admission for same day surgery or nursing care facility.

Changes in documentation over time and variable documentation practices and data capture across healthcare systems can lead to misclassification and potentially, information bias in observational studies.^{1,2}

Objective

This study uses EHRs to identify and describe misclassification related to encounter classification.

Methods

Data source

- Data were from the Oracle EHR Real-World Data (ORWD), a cloud-based, de-identified, and Health Insurance Portability and Accountability Act (HIPAA)-complaint database.³ This multicenter EHR database aggregates data from healthcare systems across the United States.
- This study utilized data from the ORWD from 135 healthcare systems.
- Data were examined over an approximate 10-year period (1/2013 to 9/2022).

Analysis

- Encounter classifications were assessed in relation to the Health Level 7 Fast Healthcare Interoperability Resources (HL7 FHIR) 2020 value set: inpatient, ambulatory, observation, emergency, virtual, and home health.⁴
- Descriptive analyses were performed.

Measures

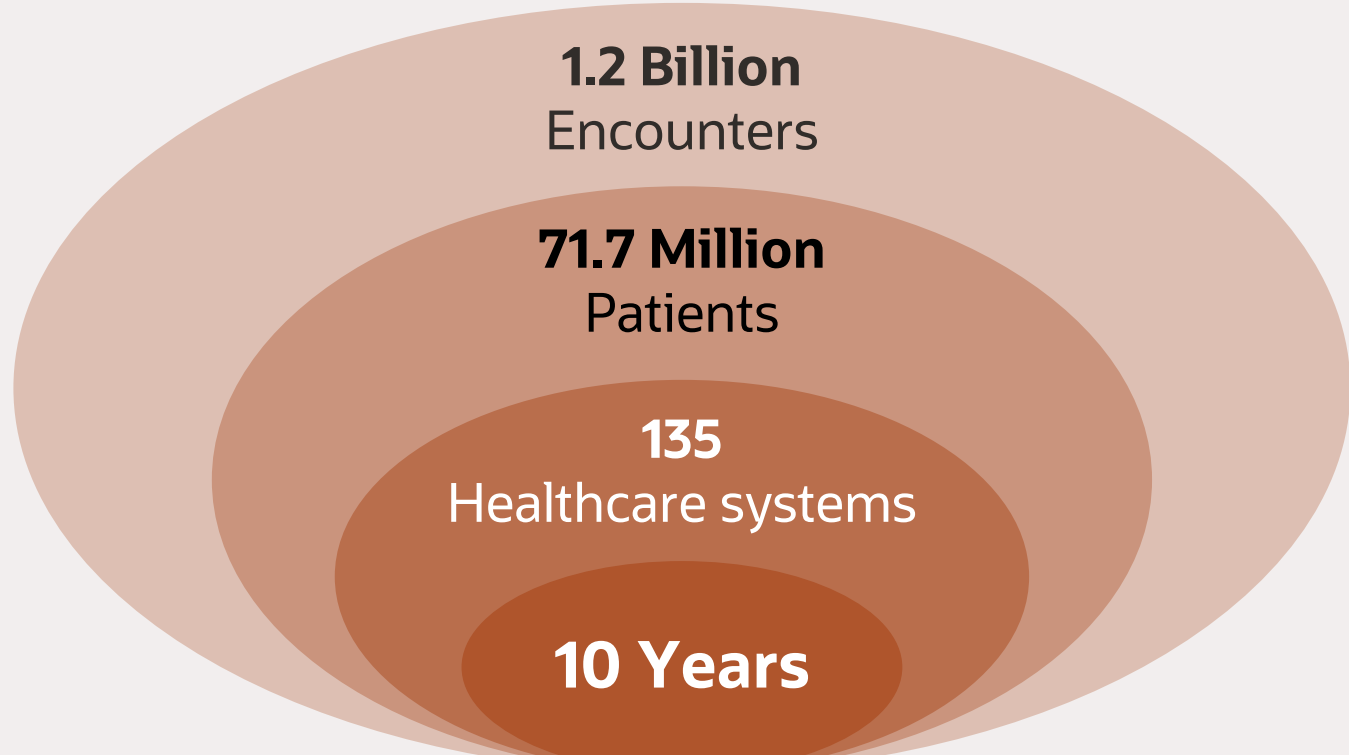
- Encounter identification (ID) numbers are used to associate all encounter-level details within the encounter.
- All patients in the ORWD have at least one encounter ID for inclusion in the database.
- Encounter classification identifies the general location of the healthcare service, such as inpatient, outpatient, and emergency.
 - Distinct encounter classifications included meaningful (i.e., outpatient or inpatient) and non-meaningful (i.e., null or unknown) values.
 - Multiple encounter classifications, typically a result of transition of care, were identified using the encounter ID.
- Encounter type identified the more granular location of the healthcare service, such as walk-in clinic, follow-up encounter, and encounter for medication refill.

Results

Over approximately 10 years, 1.2 billion unique encounters IDs were identified for 71.1 million patients across 135 healthcare systems (Fig 1). Among these unique encounter IDs:

- 80.7% had a single distinct encounter classification-type combination, including meaningful and non-meaningful encounter classifications.

Figure 1: Encounters, Patients, and Healthcare Systems in the ORWD (2013-2022)



Longitudinal contribution of data across 135 healthcare systems were examined with the majority of healthcare systems contributing annually to the EHR.

- During the 9.8-year assessment period, healthcare systems provided an average of 8.3 years (standard deviation [SD] = 2.1; median = 9.8) of encounter data.
- Healthcare systems had on average 1.0 million encounters per year (SD = 1.5 million; median = 0.4 million).

1.5 billion encounter classifications were identified during the assessment period.

- Most encounter classifications were for outpatient (69.2%), preadmit (15.7%), and emergency (7.5%) visits (Table 1).
- Non-meaningful values represented 17.3% of encounter classifications.

Table 1. Encounter Classifications in the ORWD (2013-2022)

Encounter Classifications	% of Total Meaningful Encounters ^a
Outpatient	69.2%
Preadmit	15.7%
Emergency	7.5%
Inpatient	3.5%
Recurring patient	1.6%
Encounter by computer link	1.0%
Admitted for Observation	0.6%
Telephone consultation	0.4%
History taking	0.2%
Research administrative procedure	0.1%
Total other (e.g., other, radiology, home health)	0.1%

^a Meaningful encounters classifications (n=1.2 billion) were defined as having a non-null or non-unknown value.

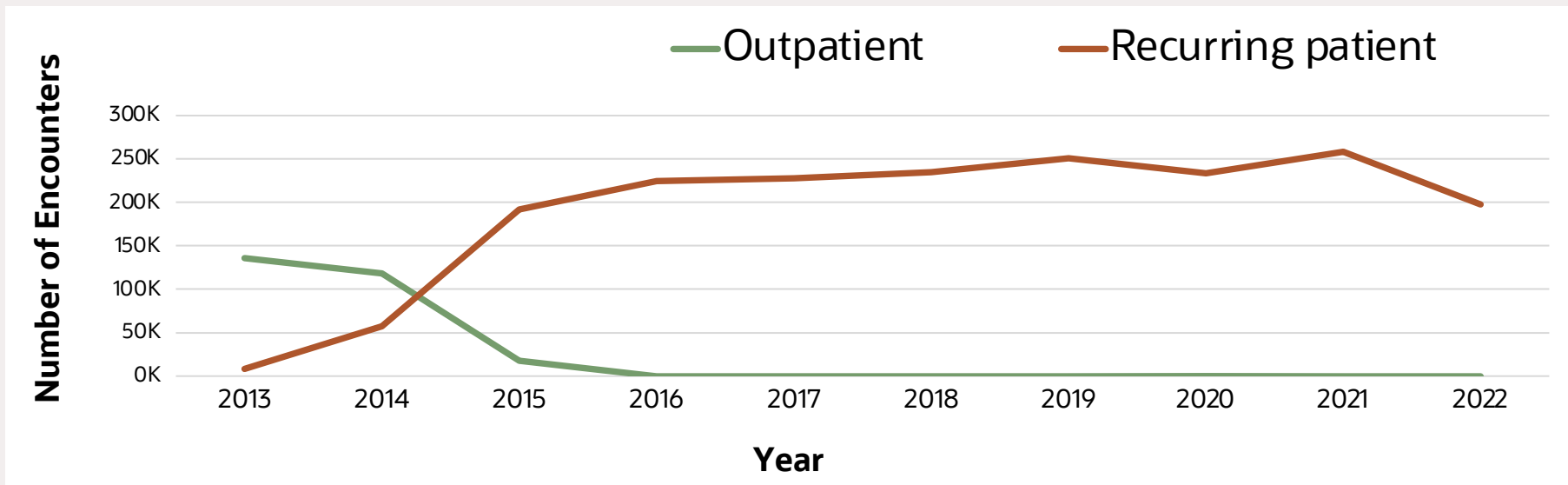
Potential for misclassification error using encounter data in EHRs was identified in four non-mutually exclusive areas (Fig 2).

Temporal changes in coding practice within a healthcare system

Changes in coding practices for encounter classification and type designation were observed within healthcare systems over time.

Example: In a single healthcare system, encounter type of 'recurring patient' was coded as encounter classification = 'outpatient' during the years 2013 to 2014 and as encounter classification = 'recurring patient' from 2015 to 2022 (Fig 3).

Figure 3. Changes in Coding Practices for Encounter Classification in a Single Healthcare System



Differential standard coding practices between healthcare systems

Coding practices for encounter classifications and types may be determined by individual healthcare systems.

Example: During the observation period, encounter type = 'urgent care' (encounter or facility) was utilized by 15.6% (21/135) of healthcare systems. Of these 21 healthcare systems, 3 had <10 encounters for the encounter type of urgent care. Among the remaining 18 healthcare systems, encounter classification was coded consistently over time by each healthcare system as either outpatient, emergency, or null (Fig 4).

Figure 2. Encounter Classification and Sources of Misclassification in Electronic Health Records

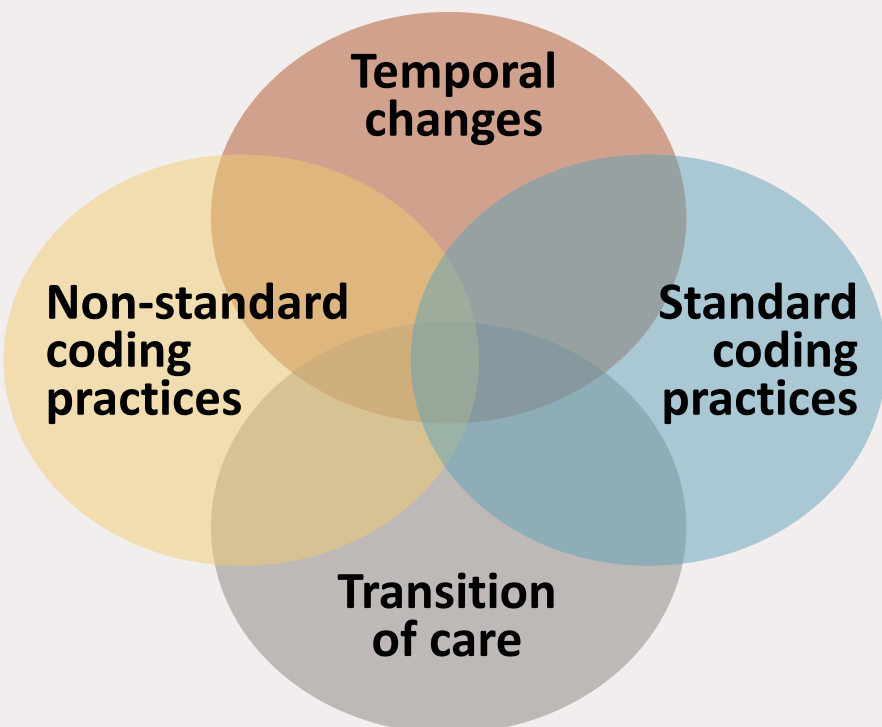
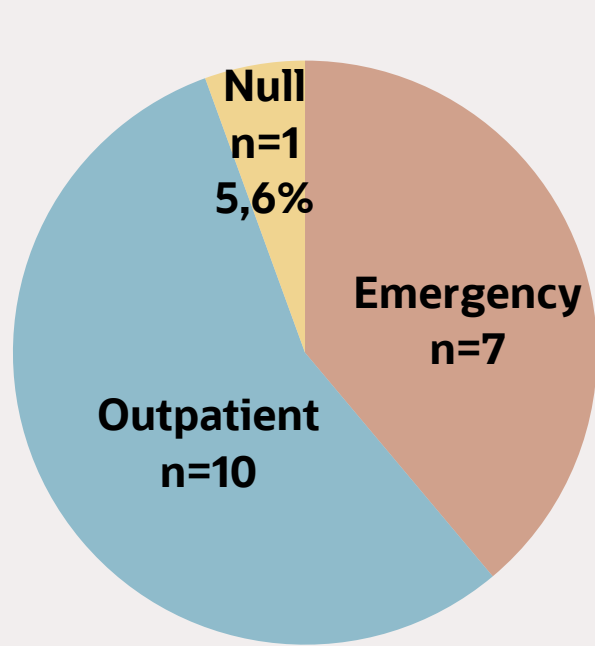


Figure 4. Encounter Classification Assignment for Encounter Type of Urgent Care



Inadequate capture of care transition during an encounter

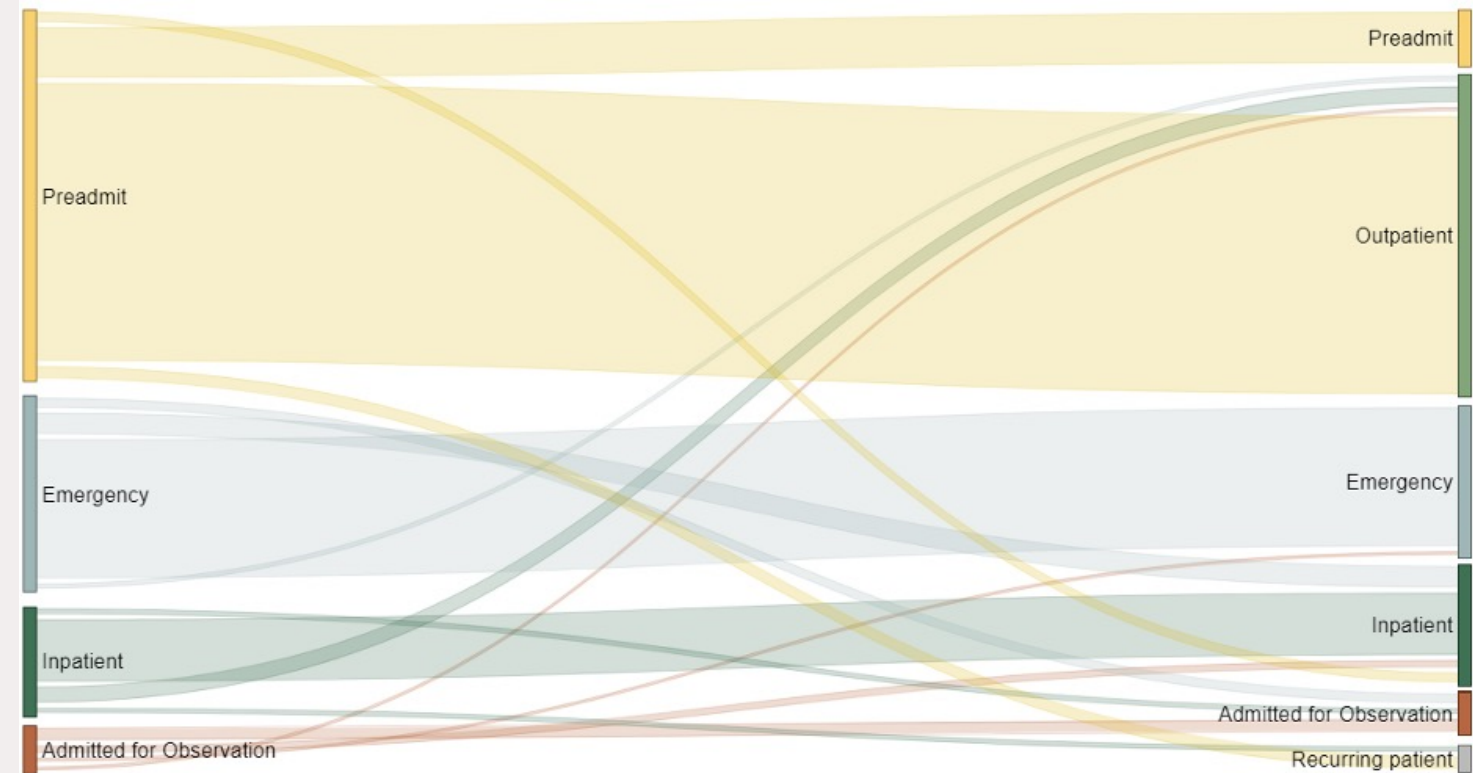
Understanding a healthcare system's coding practices is required to ascertain the transition from one department or location to another. Differential practices include:

- Multiple encounter classifications on a single encounter ID. The assessment of transitions between care requires ordering of the date-time stamp associated with the encounter classification (e.g., first, next, last).

- Consecutive unique encounter IDs for a related event. A transitioning encounter is identified using a time-dependent algorithm associating each encounter ID within a specific time period.

Example: Multiple encounter classifications on a single encounter ID indicate transitioning encounters classifications when the first and next encounter classifications differ (Fig 5).

Figure 5. Multiple Encounter Classifications Associated with a Single Encounter ID



Non-standard coding practice

While standardized coding conventions are recommended, some healthcare systems have created unique encounter classifications.

Example: HL7 FHIR standard codes for encounter classifications include inpatient, ambulatory, observation, emergency, virtual, and home health, but some healthcare systems have non-FHIR standard coding encounter classification concepts:

- HL7 FHIR defines virtual encounter classifications as "...Where the patient and the practitioner(s) are not in the same physical location. Examples include telephone conference, email exchange, robotic surgery, and televideo conference."⁴ However, in the ORWD there are unique, non-FHIR standard classifications for 'encounter by computer link' and 'telephone consultation' only.
- 'Total Other' encounter classifications (see Table 1) comprise 0.1% of all non-null encounter classifications (1.6 million) and include:

- Attending clinic
- Encounter for blood donation
- Home health
- Private nursing service
- Seen in clinic
- Seen in radiology department
- Skilled nursing specialty
- Other

Strengths and limitations

- The ORWD is a large, national EHR-agnostic database that aggregates patient data from all EHR providers within a healthcare system.
- Data mapping is limited to standard fields and thus might result in missing data.
- Facility and physician-level variability in encounter classification coding may also impact misclassification.

Conclusions

- We identified the potential for measurement error in four areas related to encounter classification coding in EHRs: temporal changes, standard coding practices, transition of care, and non-standard coding practices.
- Direct or indirect causes of measurement errors in encounter classification may be the result of numerous factors including the natural growth within a healthcare system, such as onboarding additional facilities or third-party EHRs, and updated regulatory standards, vocabulary, and value sets, such as the United States Core Data for Interoperability (USCDI) or HL7 FHIR value sets.^{4,5}
- The example of non-standard coding practice illustrates the consequences of a flexible versus rigid EHR platform. Thus, while healthcare system-specific codes may create efficiencies in care, these changes have implications for interoperability and downstream research.
- Measurement error can be considered a type or subset of missing data,⁶ and while missing data requires attention with regard to data management and bias assessment, it was not the focus of the present study.
- Large, multisystem databases are at risk for misclassifying encounter classifications during data collection, management, and analysis and can result in information bias. In an effort to reduce this bias, investigators are urged to examine and manage these potential sources when generating real-world evidence.

References

- Young JC, Conover MM, Jonsson Funk M. Measurement Error and Misclassification in Electronic Medical Records: Methods to Mitigate Bias. *Current Epidemiology Reports* 2018 5:4. 2018;5(4):343-356. doi:10.1007/s40471-018-0164-X
- Hernan MA, Cole SR. Invited Commentary: Causal Diagrams and Measurement Bias. *Am J Epidemiol*. 2009;170(8):959-962. doi:10.1093/aje/kwp295
- Elwerthemuepha L, Carlson K, Moog R, et al. Cerner real-world data (CRWD) - A de-identified multicenter electronic health records database. *Data Brief*. 2022;42:108120. doi:10.1016/j.dib.2022.108120
- FHIR. HL7 FHIR CI-Build: Resource Encounter - Content . HL7.org 2011+. FHIR R6 hl7.fhir.core#6.0.0-cibuild. Published September 2, 2023. Accessed September 10, 2023. https://build.fhir.org/encounter.html
- Interoperability Standards Advisory (ISA). United States Core Data for Interoperability (USCDI). Published July 2023. Accessed September 10, 2023. https://www.healthit.gov/isa/united-states-core-data-interoperability-uscdi
- Howe CJ, Cain LE, Hogan JW. Are All Biases Missing Data Problems? *Curr Epidemiol Rep*. 2015;2(3):162-171. doi:10.1007/s40471-015-0050-8