

Health Information Management Best Practices for Quality Health Data During the COVID-19 Global Pandemic

Mary H. Stanfill, MBI, RHIA, CCS, CCS-P, FAHIMA, Kathy Giannangelo, MA, RHIA, CCS, CPHIMS, FAHIMA, and Susan H. Fenton, PhD, RHIA, CPHI, FAHIMA

Introduction

The emergence of the deadly COVID-19 virus, resulting in the current world-wide pandemic requires effective information management and accurate data reporting to enable the healthcare community's efforts to learn as much as possible about this specific strain of coronavirus and halt the pandemic. Therefore, it is essential that health information management (HIM) professionals ensure COVID-19 documentation, data capture, data analysis and reporting, as well as coding are accurate and reliable to support clinical care, organizational management, public health reporting, population health management, and scientific research. However, many health information (HIM) professionals in inpatient and ambulatory care settings are struggling to provide direction on COVID-19 data capture due to evolving guidance ranging from what needs to be captured to how the data should be coded and reported to how to accurately present the data both internally and externally. Detailed clinical documentation for COVID-19 data collection is also a challenge due to unprecedented demands on healthcare providers.

For example, though the first COVID-19 case in the US was confirmed on January 21¹, the first coding guidance was not available until late February. The National Center for Health Statistics (NCHS) developed interim COVID-19 coding advice in a supplement to the ICD-10-CM Official Coding Guidelines that was effective February 20, 2020.² This guidance provided advice on which existing coronavirus and infection codes to use until a specific code for COVID-19 could be implemented in the US. A new ICD-10-CM code (U07.1, COVID-19) was subsequently released, effective for discharges on or after 4/1/2020.³ The American Hospital Association and the American Health Information Management Association released further clarification and guidance on correct coding for COVID-19 in March, 2020.⁴ This initial uncertainty and evolving guidance and coding change in the midst of the pandemic is likely to result in extreme variability in the clinical and administrative data spanning the full course of the pandemic. Administrative data are patient-identifiable data used for administrative, regulatory, healthcare operations, and payment (financial) purposes.⁵ Clinical data is information that is recorded about a patient and their care in a patient's medical chart, in an electronic health record, or in a clinical data registry.⁶

This variability and the complexity of the data that will be needed for clinical care, public health reporting, population health management, and scientific research makes it imperative that HIM professionals look beyond administrative data, such as the ICD-10-CM/PCS and HCPCS/CPT code sets used for financial and other purposes, to more detailed clinical data. Administrative data alone is unlikely to be adequate for the comprehensive data required for the COVID-19 pandemic. HIM professionals should take the lead in assisting healthcare organizations to ensure that both clinical and administrative data capture is accurate and reliable for internal and external purposes. This paper presents best practices from three authors with a collective 112 years of health information management experience across a wide range of organizational and public health settings. These best practices for ensuring quality health data, include proactive steps

during the pandemic, as well as retrospective data validation practices to apply to both clinical and administrative data.

Data Quality Management Best Practices

General best practices for interactive data quality management in both inpatient and ambulatory care settings involve the iterative steps found below. While the steps may differ in actual practice between the electronic health record (EHR) and paper record, the basics remain the same.

1. Identify specific data elements that can be or should be collected for a specific data capture need. This includes identifying all relevant data fields, data types, data definitions, and the associated data values using the applicable data dictionary. Create a template containing these data elements. Each data element listed should have a standard definition and standard value. The goal is to have a template that enables structured data capture and a standard way to collect the data elements. For coded data specifically, identify all relevant code values.
2. Consider the expected data results (i.e. data output) for the identified data elements that are specifically associated with the data collection need. For coded data, that may include identifying changes to codes during a reportable time period or documentation and reporting guidelines such as code sequencing and bundling or unbundling to determine normative coded data patterns.
3. Run a time-limited report, either on the full set of the template data elements or on selected data elements, to obtain data output for a time-limited subset of data.
4. Evaluate the data output to determine if the results are consistent with what is expected.
5. Analyze the result. For example, does the data output indicate under-reporting due to lack of an available specific coded value?
6. Document findings and explanations in a data quality issue log for future reference when reports are run using the identified data elements.
7. Follow up to resolve any aberrant data patterns identified. For example, work with Clinical Documentation Improvement (CDI) staff to develop scenarios to increase awareness around the importance of data quality. Likewise, an investigation may be needed to address any missing data elements, especially if clinical data is missing from the EHR.

This process of pulling a small set of data to validate that documentation and data entry (i.e. data capture) is consistent with expectations should be done following specific identified transition points and at appropriate intervals to ensure data is accurate and reliable.

The rest of this paper discusses how to employ these data quality best practices more specifically for the purpose of ensuring COVID-19 data quality in administrative as well as clinical data during the current pandemic.

Administrative Data Quality Reporting Best Practices

In order to ensure the accuracy and reliability of administrative COVID-19 data capture for both internal and external uses, HIM professionals should follow best practice steps. The application of best practices in the steps below give examples using data coded with the standard transaction

code sets. However, these steps are repeatable to explore logical, expected coded data patterns associated with the COVID-19 relevant data elements used for administrative reporting. For example, the concepts here could also be used to validate use of a new condition code, DR, Disaster Related, established for reporting on the Uniform Bill X12 837I Version 5010A.⁷

1. Identify all codes from the standard transaction code sets (e.g. ICD-10-CM/PCS, HCPCS Level II, CPT) that might be associated with a COVID-19 episode, starting from a potential COVID-19 exposure to a death due to complications from infection with COVID-19. Sources to identify all codes include the CDC and NCHS^{2,3}, AHA⁴, CMS⁸, and the AMA.⁹ A sample list of ICD-10-CM codes is included in Appendix A.
2. Consider the anticipated COVID-19 data for the health system. Consult the local public health department, the facility Medical Director or Infection Control Manager to determine the onset of COVID-19 within the health system. Speak with more than one person to validate the date and establish a shared understanding of the “beginning” of COVID-19 for data analysis purposes. The health system’s start date might be January 21st in the state of Washington, or on/after January 27, 2020, the date that the US Department of Health and Human Services declared the COVID-19 crisis as a federal public health emergency.¹⁰
 - a. For example, ask “In order to track these cases, we need to define, as a system, the window of time that we are going to call our COVID-19 treatment window. I need a starting date so that I can compartmentalize the data to provide accurate reporting.”
3. Run pre-COVID-19 reports to identify whether and if so, how the identified existing codes were used in the organization to report cases unrelated to COVID-19. The report should be prior to the defined begin date for COVID-19 in the organization.
 - a. Evaluate the data output to determine whether or not there was a significant number of patients using any of the codes prior to the beginning of the health system’s COVID-19 window. Codes utilized significantly before COVID-19 began (e.g., reported for 15 or more patients within the last three months) may not be useful for uniquely identifying COVID-19 patients during the pandemic. If this is the case, the health system may need to rely more on clinical data, encoded in SNOMED CT for example (see the section below on clinical data quality practices).
 - b. Analyze the results to determine why the codes were used prior to the emergence of COVID-19 in the health system to ensure a good understanding of the data in order to inform appropriate use and interpretation of administrative data. If codes are not used significantly before COVID-19 began (e.g., reported for less than 15 patients within the last three months), data collected after the organization begin date can be reasonably presumed to be informative for identifying cases related to the COVID-19 pandemic.
 - c. Document findings and explanations in a data quality issue log for future reference when reports are run using the identified data elements.
 - d. Follow up to resolve any aberrant data patterns identified.
4. Following the implementation of the new COVID-19 code (U07.1) for discharges/date of service on and after 4/1/2020, run a report to ensure data is accurate and reliable.

- a. Analyze the report to determine if the data follows the expected pattern. For example, if the organization has a significant number of COVID-19 patients, the use of code B97.29 should drop off, while the use of code U07.1 should increase on and after 4/1/2020. Appendix A includes a table summarizing expected diagnosis coded data patterns before and on/after 4/1/2020.
 - b. If the data does not follow the expected pattern, conduct an investigation to determine the root cause and, as previously described, document findings/explanations in a data quality issue log and follow up to resolve the problem and correct the data.
5. Take additional steps to reconcile administrative and clinical data. Once the use of transaction codes sets (e.g. ICD-10-CM diagnosis codes) is validated in the administrative data, compare the identified cases with the cases identified using clinical data elements. Analyzing the SNOMED CT encoded clinical data found in the electronic health record for example may reveal additional COVID-19 cases (see below).

Clinical Data Quality Reporting Best Practices

HIM professionals also have a key role in ensuring the accuracy and reliability of clinical COVID-19 data capture for both internal and external uses. The following steps illustrate the use of clinical data quality reporting best practices.

1. Examine the clinical information systems data dictionaries to identify data elements and fields relevant to identifying exposure to, symptoms of, testing for, diagnosis of, and/or treatment of COVID-19.
 - a. If no data dictionary is present, collaborate with the appropriate clinical professionals within the organization to identify and collate these data elements for COVID-19. Refer to the CDC COVID-19 Data Dictionary [11] and CDC COVID-19 Patient Impact & Hospital Capacity Module Form[12] as a starting point.
 - b. If the data dictionary is present, identify relevant data elements for COVID-19.
 - i. Compare the relevant fields selected to CDC and CMS reporting requirements to identify gaps and differences in field definitions. Where gaps exist, add to the data dictionary. Reconcile any differences.
 - c. Determine which data elements are structured versus unstructured. For structured fields, identify the allowable code values to understand data output. Identify all identifiers from terminologies captured in clinical systems (e.g., SNOMED CT, LOINC) associated with a COVID-19 episode. [13, 14] A sample list of SNOMED and LOINC identifiers is included in Appendix B.
2. As described for administrative data, run pre-COVID-19 reports to identify whether and if so, how the identified existing clinical data elements and clinical concept identifiers were utilized in the organization prior to the defined begin date for COVID-19 in the organization.
 - a. Review reports to determine whether or not there was a significant number of patients using any of the clinical identifiers prior to the beginning of the health system's COVID-19 window.
 - b. Document findings and explanations in a data quality issue log for future reference when reports are run using clinical data elements.

- c. Conduct an investigation to determine the root cause of aberrant data and follow up to resolve the problem.
3. Run reports within the health system's active COVID-19 window to determine if clinical information systems (e.g., the electronic health record), are capturing expected clinical COVID-19 data for identified data elements, including SNOMED-CT and LOINC concept identifiers.
 - a. Relevant clinical data may be found for example in new screening tools to capture information on suspected cases (e.g., travel to different countries and other exposure risks) and SNOMED-CT concept identifiers are likely to be used for example in the problem list. Appendix B includes a table with examples of expected SNOMED CT coded data patterns.
 - b. Evaluate the data output to determine if the clinical data follows the expected pattern. For example, if care providers are not completing data fields in new screening tools, the report will show little or no data. This might indicate a need to establish or re-visit documentation priorities. Minimal reporting or skipping important documentation elements will impact an organization's ability to fully analyze and understand the COVID-19 pandemic now and in the future.
 - c. Analyze the results. Use the reports to confirm SNOMED and LOINC identifiers are in use in clinical information systems and ensure a good understanding of the data in order to inform appropriate use and interpretation of clinical data to track COVID-19 cases. Reconcile cases identified with SNOMED and LOINC with cases identified with administrative data (e.g. ICD-10-CM).
 - d. Document findings and explanations in a data quality issue log for future reference when reports are run using identified clinical data elements.
 - e. Follow up to resolve any aberrant data patterns identified and resolve variances in administrative and clinical data. Consult with the medical director and/or infection control manager and CDI staff if warranted to revise or implement COVID-19 documentation guidance for the health system. If SNOMED and LOINC identifiers are not captured by the health system's clinical information systems, follow up with system owners to determine when they might become available to ensure complete COVID-19 data capture.

Conclusion

Utilizing these recommended best practices increases the HIM professional's understanding of the clinical and administrative data collected and reported for a COVID-19 episode. They also provide a mechanism for data validation to ensure that clinical and administrative data is captured, reported, used and interpreted appropriately.

The best practices described in this article should be taken proactively and iteratively, at the start of and during the pandemic, in order to validate COVID-19 data and ensure consistent data capture. It is imperative that HIM professionals follow these best practices and stay abreast of coding and data reporting changes as they are announced to have the high-quality data needed, as well as consider how both administrative and clinical data can be appropriately used and interpreted for clinical care, public health reporting, population health management, trending and scientific research to address COVID-19. With these best practices in place, HIM professionals

can be a trusted advisor during a global pandemic and influence patient health outcomes now and into the future.

References:

1. CDC. First Travel-related Case of 2019 Novel Coronavirus Detected in the United States. <https://www.cdc.gov/media/releases/2020/p0121-novel-coronavirus-travel-case.html>
2. NCHS. ICD-10-CM Interim Coding Guidance for COVID-19. February 20, 2020 <https://www.cdc.gov/nchs/data/icd/ICD-10-CM-Official-Coding-Gudance-Interim-Advice-coronavirus-feb-20-2020.pdf>
3. NCHS. New ICD-10-CM Code for the 2019 Novel Coronavirus (COVID-19). March 18, 2020 <https://www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-3-18-2020.pdf>
4. AHA Frequently asked questions regarding ICD-10-CM Coding for COVID-19. March 26, 2020 <https://www.codingclinicadvisor.com/faqs-icd-10-cm-coding-covid-19>
5. AHIMA. "Fundamentals of the Legal Health Record and Designated Record Set." *Journal of AHIMA* 82, no.2 (February 2011): Expanded Online Version
6. Office of the National Coordinator for Health Information Technology. Health IT Playbook, Glossary. <https://www.healthit.gov/playbook/glossary/#c>
7. National Uniform Billing Committee. NUBC Guidance: Claims for COVID 19 Treatment. March 23, 2020. https://www.nubc.org/system/files/media/file/2020/03/NUBC%20Announcement%20for%20COVID-19%20claims_1.pdf
8. CMS HCPCS Code for lab test: <https://www.cms.gov/newsroom/press-releases/cms-develops-additional-code-coronavirus-lab-tests>
9. AMA: <https://www.ama-assn.org/practice-management/cpt/covid-19-coding-and-guidance>
10. Department of Health and Human Services. Secretary Azar Declares Public Health Emergency for United States for 2019 Novel Coronavirus. January 31, 2020. <https://www.hhs.gov/about/news/2020/01/31/secretary-azar-declares-public-health-emergency-us-2019-novel-coronavirus.html>
11. CDC 2019-nCoV PUI and CRF Data Dictionary <https://www.cdc.gov/coronavirus/2019-ncov/downloads/data-dictionary.pdf>
12. CDC. COVID-19 Patient Impact & Hospital Capacity Module. <https://www.cdc.gov/nhsn/acute-care-hospital/covid19/index.html>

13. SNOMED International. SNOMED CT Coronavirus Content.

<https://confluence.ihtsdotools.org/display/snomed/SNOMED%2BCT%2BCoronavirus%2BContent>

14. LOINC. SARS Coronarvirus 2. <https://loinc.org/prerelease/>

Resources:

Centers for Disease Control and Prevention. Coronavirus (COVID-19). n.d.

<https://www.cdc.gov/coronavirus/2019-ncov/index.html>

Centers for Medicare and Medicaid Services. Current Emergencies. March 31, 2020.

<https://www.cms.gov/About-CMS/Agency-Information/Emergency/EPRO/Current-Emergencies/Current-Emergencies-page>

National Institutes of Health. Coronavirus (COVID-19). March 31, 2020.

<https://www.nih.gov/health-information/coronavirus>

National Library of Medicine. Value Set Authority Center. March 19, 2020.

<https://vsac.nlm.nih.gov/welcome>

Office of the National Coordinator. Interoperability for COVID-19 Coronavirus Pandemic. 2020.

<https://www.healthit.gov/isa/covid-19>

APPENDIX A

Sample List of Relevant ICD-10-CM Codes for COVID-19

ICD-10-CM Code	Code Description
B97.29	Other Coronavirus as the cause of disease classified elsewhere
B34.2	Coronavirus infection, unspecified
Z03.818	Encounter for observation for suspected exposure to other biological agents ruled out
Z11.9	Encounter for screening for infectious and parasitic disease, unspecified
Z20.828	Contact with and suspected exposure to other viral communicable disease
U07.1	COVID-19 (effective date 4/1/2020)

Examples of Expected ICD-10-CM Coded Data Patterns

Scenario	Before 4/1/2020	On and After 4/1/2020
Pneumonia due to COVID-19	J12.89 Other viral pneumonia B97.29 Other coronavirus as the cause of disease classified elsewhere	U07.1 COVID-19 J12.89 Other viral pneumonia
Acute bronchitis due to COVID-19	J20.8 Acute bronchitis due to other specified organism B97.29 Other coronavirus as the cause of disease classified elsewhere	U07.1 COVID-19 J20.8 Acute bronchitis due to other specified organism
ARDS due to COVID-19	J80 Acute respiratory distress syndrome B97.29 Other coronavirus as the cause of disease classified elsewhere	U07.1 COVID-19 J80 Acute respiratory distress syndrome

Appendix B

Sample List of Relevant SNOMED and LOINC Concepts for COVID-19

Identifier	Terminology	Term/Name
840539006	SNOMED CT	Disease caused by severe acute respiratory syndrome coronavirus 2
840544004	SNOMED CT	Suspected disease caused by severe acute respiratory coronavirus 2
840546002	SNOMED CT	Exposure to SARS-CoV-2
840534001	SNOMED CT	Severe acute respiratory syndrome coronavirus 2 vaccination
94558-4	LOINC	SARS coronavirus 2 Ag [Presence] in Respiratory specimen by Rapid immunoassay
94559-2	LOINC	SARS coronavirus 2 ORF1ab region [Presence] in Respiratory specimen by NAA with probe detection

Examples of Expected SNOMED CT Coded Data Patterns

Scenario	SNOMED CT 2020-03-09 Release	SNOMED CT 2020-07-31 Release
Pneumonia due to COVID-19	441590008 Pneumonia caused by Severe acute respiratory syndrome coronavirus (disorder)	882784691000119100 Pneumonia caused by SARS-CoV-2 (disorder)
Acute bronchitis due to COVID-19	233601004 Acute viral bronchitis (disorder) 415360003 Severe acute respiratory syndrome-related coronavirus (organism)	138389411000119105 Acute bronchitis caused by SARS-CoV-2 (disorder)
ARDS due to COVID-19	67782005 Acute respiratory distress syndrome (disorder) 415360003 Severe acute respiratory syndrome-related coronavirus (organism)	674814021000119106 Acute respiratory distress syndrome caused by SARS-CoV-2 (disorder)