Uptake and Accuracy of the Diagnosis Code for COVID-19 Among US Hospitalizations

Administrative claims may provide an understanding of the clinical and economic burden of novel coronavirus disease 2019 (COVID-19). However, their limited sensitivity in identifying conditions, such as sepsis, suggests the need to assess whether billing codes reliably capture COVID-19 discharges.

A new diagnosis code for COVID-19 (International Statistical Classification of Diseases, Tenth Revision, Clinical Modification [ICD-10-CM] code U07.1) was introduced on April 1, 2020, to facilitate billing and case monitoring. For discharges prior to April 1, 2020, the Centers for Disease Control and Prevention advised coding for the clinical syndrome (eg, pneumonia) and “other coronavirus as the cause of diseases classified elsewhere” (ICD-10-CM code B97.29).

We examined the uptake in COVID-19–specific coding (ICD-10-CM code U07.1), the transition from legacy coding, and the accuracy of the COVID-19–specific code using severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) polymerase chain reaction (PCR) testing as the reference standard.

Methods | The Premier Healthcare Database, an administrative all-payer repository that covers approximately 20% of all US hospitalizations from 48 states, was queried for inpatient discharges between January 1, 2020, and May 31, 2020, with COVID-19–specific coding or legacy coding. Only hospitals that reported data for every week during the study period were included. A descriptive analysis was performed by week to examine when and to what extent each hospital transitioned from legacy coding to COVID-19–specific coding.

Discharges from hospitals that also reported microbiology data through the TheraDoc clinical surveillance system also were queried for SARS-CoV-2 PCR test results from respiratory specimens. A positive SARS-CoV-2 PCR test result during or up to 4 weeks prior to the hospitalization was used as the reference standard against which the sensitivity, specificity, positive predictive value, and negative predictive value of ICD-10-CM code U07.1 were calculated. All analyses were performed using SAS version 9.4 (SAS Institute Inc). Given the use of deidentified data exclusively, the study was deemed not to require ethics board review based on the policy of the Office of Human Subjects Research Protections, National Institutes of Health, under the revised Common Rule.

Results | Between January 1, 2020, and May 31, 2020, 3,128,064 inpatient discharges were identified in 875 hospitals, of which 74,084 (2.37%) recorded specific and 8,656 (0.28%) recorded legacy coding for COVID-19. Legacy coding captured the early increase in putative COVID-19 discharges prior to the use of ICD-10-CM code U07.1, increasing from 706 discharges in February 2020 to 6,636 discharges in March 2020 (Figure). However, hospitals rapidly began using ICD-10-CM code U07.1 within 2 weeks of its release, at which point legacy coding decreased to prepandemic levels.

Among 150 hospitals reporting microbiology data, 51,907 inpatient discharges with SARS-CoV-2 PCR test results were identified during April and May 2020. Within this cohort, the sensitivity of ICD-10-CM code U07.1 was 98.01% (95% CI,
Table. Accuracy of the ICD-10-CM COVID-19-Specific Diagnosis Code in Capturing Inpatient Discharges Displaying PCR-Confirmed COVID-19 From 150 Hospitals Between April 1, 2020, and May 31, 2020

<table>
<thead>
<tr>
<th>Code used*</th>
<th>Positive</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code used</td>
<td>4866</td>
<td>451</td>
<td>5317</td>
</tr>
<tr>
<td>Code not used</td>
<td>99</td>
<td>645491</td>
<td>46590</td>
</tr>
</tbody>
</table>


Discussion

Despite its off-cycle release, the diagnosis code for COVID-19 (ICD-10-CM code U07.1) was widely adopted across study hospitals and rapidly replaced a stop-gap approach in coding for COVID-19 that relied on a preexisting nonspecific code for coronavirus. The COVID-19–specific code showed high sensitivity and specificity compared with the PCR test results. The lower positive predictive value reflects discharges with the COVID-19–specific ICD-10-CM code U07.1 and discharges with negative SARS-CoV-2 PCR test results. These patients may have been diagnosed outside the included hospitals or diagnosed more than 4 weeks prior to admission, or the diagnosis was based on clinical rather than laboratory grounds.

The World Health Organization has provided a specific code (ICD-10-CM code U07.2) for patients with negative SARS-CoV-2 PCR test results, but this code is not currently available for use in the US. Limitations include the potential for incomplete data due to ongoing data collection, selection bias in those with PCR test results available, and false-negative PCR test results. Furthermore, hospitals reporting microbiology data may not have been representative of all US hospitals.

The findings suggest that hospitals appear to provide reasonably accurate COVID-19 diagnosis codes in administrative data, and these codes may be a reasonable means to track inpatient discharges and costs associated with COVID-19. As the pandemic continues to evolve, it will be important to reassess the reliability of ICD-10-CM code U07.1 to capture COVID-19 cases.

Someer S. Kadri, MD
Jake Gundrum, MS
Sarah Warner, MPH
Zhun Cao, PhD
Ahmed Babiker, MBBS
Michael Klompas, MD
Ning Rosenthal, MD, PhD, MPH

Author Affiliations: Critical Care Medicine Department, National Institutes of Health Clinical Center, Bethesda, Maryland (Kadri, Warner); Premier Applied Sciences, Premier Healthcare Solutions Inc, Charlotte, North Carolina (Gundrum, Cao, Rosenthal); Division of Infectious Diseases, Emory University School of Medicine, Atlanta, Georgia (Babiker); Department of Population Medicine, Harvard Medical School, Boston, Massachusetts (Klompas).

Accepted for Publication: September 28, 2020.

Corresponding Author: Ning Rosenthal, MD, PhD, MPH, Premier Applied Sciences, Premier Healthcare Solutions Inc, 13034 Ballantyne Corporate Pl, Charlotte, NC 28277 (ning_rosenthal@premierinc.com).

Author Contributions: Ms Warner had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: Kadri, Gundrum, Cao, Rosenthal. Acquisition, analysis, or interpretation of data: Kadri, Gundrum, Warner, Babiker, Klompas, Rosenthal. Drafting of the manuscript: Kadri, Warner, Rosenthal. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Gundrum, Warner, Cao, Rosenthal. Obtained funding: Kadri, Rosenthal. Administrative, technical, or material support: Warner, Rosenthal. Supervision: Kadri, Rosenthal.

Conflict of Interest Disclosures: Mr Gundrum, Dr Cao, and Dr Rosenthal are employees and shareholders of Premier Healthcare Solutions Inc. Dr Klompas reported receiving grants from the Centers for Disease Control and Prevention and receiving grants and personal fees from UpToDate. No other disclosures were reported.

Funding/Support: This work was supported by Premier Healthcare Solutions Inc and the Intramural Research Program of the National Institutes of Health Clinical Center.

Role of the Funder/Sponsor: The funders/sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The findings and conclusions in this study are those of the authors and do not necessarily represent the official position of the National Institutes of Health.

Additional Contributions: We thank John P. Dekker, MD, PhD, National Institute of Allergy and Infectious Diseases, for his insightful feedback on laboratory testing for SARS-CoV-2. Dr Dekker did not receive financial compensation for his contribution.


97.62%-98.39%), the specificity was 99.04% (95% CI, 98.95%-99.13%), the positive predictive value was 91.52% (95% CI, 90.77%-92.27%), and the negative predictive value was 99.79% (95% CI, 99.75%-99.83%) (Table). Among patients tested between April and May 2020, the mean percentage of positive SARS-CoV-2 PCR test results was 9.57% (4965/51 907).

© 2020 American Medical Association. All rights reserved.