Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	1
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	2
	4	Study objectives and hypotheses	2
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	n/a
Participants	**6**	Eligibility criteria	2
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	2
	8	Where and when potentially eligible participants were identified (setting, location and dates)	n./a
	9	Whether participants formed a consecutive, random or convenience series	2
Test methods	**10a**	Index test, in sufficient detail to allow replication	2
	10b	Reference standard, in sufficient detail to allow replication	2
	11	Rationale for choosing the reference standard (if alternatives exist)	n/a
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	n/a
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	n/a
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	n/a
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	n/a
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	2-3
	15	How indeterminate index test or reference standard results were handled	2-3
	16	How missing data on the index test and reference standard were handled	2-5
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	2-5
	18	Intended sample size and how it was determined	n/a
RESULTS			
Participants	**19**	Flow of participants, using a diagram	n/a
	20	Baseline demographic and clinical characteristics of participants	n/a
	21a	Distribution of severity of disease in those with the target condition	n/a
	21b	Distribution of alternative diagnoses in those without the target condition	n/a
	22	Time interval and any clinical interventions between index test and reference standard	2-5
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	n/a
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	2-5
	25	Any adverse events from performing the index test or the reference standard	n/a
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	5
	27	Implications for practice, including the intended use and clinical role of the index test	6
OTHER INFORMATION			
	28	Registration number and name of registry	2
	29	Where the full study protocol can be accessed	n/a
	30	Sources of funding and other support; role of funders	6

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

ICD hospital diagnostic codes to identify SARS-CoV-2 infections are valid both in reference to PCR results and with latent class analyses assuming no reference standard

Cristiano S. Moura PhD[1], Autumn Neville BA[2], Fangming Liao MSc[3], Bijun Wen PhD [3], Fahad

Razak MD, MSc, FRCPC [3,4,5], Surain Roberts PhD[3,5], Amol A. Verma MD, MPhil, FRCPC[3,4,5],

Sasha Bernatsky MD PhD[1,2],

[1] McGill University, Montreal, Quebec, Canada

[2]Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada

[3]St. Michael's Hospital, Unity Health Toronto, Toronto, ON

[4]Department of Medicine, University of Toronto, Toronto, ON

[5]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON

Abstract

Background: In 2020, International Classification of Diseases (ICD-10) codes were created for lab-confirmed SARS-CoV-2 infections. We assessed the operating characteristics of ICD-10 discharge diagnostic code U07.1 within a large unselected general population

Methods: The General Medicine Inpatient Initiative (GEMINI) assembles hospitalization data (including administrative ICD-10 discharge diagnostic codes, lab results, demographics, and more) from hospitals in Ontario, Canada. We studied adults (18+) admitted during 2020 and tested at least once for SARS-CoV-2 via polymerase chain reaction (PCR) during (or within 48 hours before) hospitalization. With PCR results as the reference standard, we calculated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for ICD-10 code U07.1 hospital discharge diagnostic codes. Analyses were stratified by demographics, calendar period, and timing of the first test (within/after 48h of hospital admission). Latent class analyses (LCA) were also performed as an alternate way to estimate sensitivity and specificity without assuming a reference standard.

Results: In 11,852 hospitalizations with at least one SARS-CoV-2 PCR test, 444 (3.7%) were positive. The sensitivity of code U07.1 to identify SARS-CoV-2 infection was 97.8%, specificity was 99.5%, PPV was 88.2%, and NPV was 99.9%. Operating characteristics were similar in most stratified analyses, but the

1

specificity and PPV were lower if the first SARS-CoV-2 test was done >48 hours after admission. With LCA, the sensitivity of code U07.1 was 97.5%, and the specificity was 100%.

Interpretation: Sensitivity, specificity, PPV, and NPV of code U07.1 were high. Our results support using this code to identify SARS-CoV-2 infection in hospitalization data.

2

Background

Reliable, standardized case definitions are needed to make optimal use of routinely collected health data, particularly administrative healthcare records. In early 2020, the World Health Organization (WHO) released a new International Classification of Diseases Revision 10 (ICD-10) code for the identification of confirmed SARS-CoV-2 infection (U07.1, virus identified). If this new code can reliably identify SARS-CoV-2 infection within health data, it could expedite important research and surveillance activities. A handful of studies have assessed ICD-10 code U07.1 in North America and elsewhere. Given the potential for variability across different jurisdictions (of coding practices, patient populations, and other factors), we aimed to confirm this code's validity within a broad patient population with universal health care. Moreover, no other study has evaluated the new ICD-10 code using latent class analyses (assuming no reference standard). In the current study, we assessed the operating characteristics of code U07.1 within hospital data, using both a reference standard (SARS-CoV-2 polymerase chain reaction, PCR results) and by using LCA with no reference standard (a novel approach, particularly in this setting).

Methods

Data

The General Medicine Inpatient Initiative (GEMINI) is a hospital research collaborative collecting clinical and administrative data from hospitals in the province of Ontario, Canada. Data include inpatient and emergency department (ED) care, including demographics, administrative discharge diagnostic codes, vital signs, and the results of laboratory test results and imaging [1,2]. In Canada, upon discharge, trained medical clerks at each hospital assign administrative discharge diagnostic codes (one most responsible diagnosis and up to 25 additional codes) [3]. We studied all adults (18+) admitted to one of seven GEMINI participating hospitals between January and December 2020 with at least one SARS-CoV-2 PCR test at admission (or the 48 hours preceding) or during hospitalization. During this period in Ontario, rapid antigen testing was generally unavailable for the general population, and confirmation of suspected SARS-CoV-2 by PCR was mandated by public health authorities.

We characterized demographics (sex, age, and urban vs rural residence as per Statistics Canada [4]), the Charlson Comorbidity Index score (CCI, based on ICD discharge diagnostic codes for current and prior hospitalizations), and specific pulmonary comorbidities (asthma, pulmonary fibrosis, and chronic obstructive pulmonary disease). We also evaluated the length of hospital stay, intensive care unit (ICU) admission, mechanical ventilation, and hospital death.

Main Analysis

Continuous variables were described using means, standard deviations, and quantiles. Categorical variables were described using counts and percentages.

3

To evaluate the performance of the U07.1 code using PCR test results as the reference standard, we identified all SARS-CoV-2 PCR tests performed 48 hours before admission and all tests performed during hospitalization. Readmissions were treated as independent observations, meaning that any tests performed in previous encounters were generally not accounted for in the current admission. We assumed cases as lab-confirmed (SARS-CoV-2 identified) if at least one test was positive in the observation period and a confirmed non-case (no SARS-CoV-2 identified) if there were no positive PCR tests (and at least one negative test result). We then assigned:

    i)   TP as lab-confirmed SARS-CoV-2, with discharge code U07.1

    ii)   FN as lab-confirmed SARS-CoV-2, without discharge code U07.1

    iii)   FP as no SARS-CoV-2 identified, with discharge code U07.1

    iv)   TN as no SARS-CoV-2 identified, without discharge code U07.1

Additional analyses

Stratified analyses were carried out for sex and age (<50 years, 50-75 years, and > 75 years), urban versus rural residence, calendar period of admission (January-April, May-August, and September-December), and timing of first PCR test (before admission, within 24h of the admission, between 24-48h of the admission, or after 48h of admission).

In additional analyses, we assessed the operating characteristics of U07.1 without assuming the PCR test result as the reference standard. For that, we used latent class analyses (LCA), a method well suited to the evaluation of case definitions in the real world when there is no completely accurate gold standard (including PCR, where clinical performance has been said to approach 80% sensitivity and 98-99% specificity)[5]. In this analysis, we estimated operating characteristics (sensitivity and specificity) by combining three different potential indicators of SARS-CoV-2 infection PCR test results, ICD-10 discharge diagnosis code U07.1, and clinical presentation (fever AND at least one additional feature: oxygen saturation < 93% OR oxygen supplementation OR mechanical ventilation). We treated disease status as a two-class latent variable and assumed non-informative priors in the estimation of model parameters. In the main LCA analyses, we excluded individual hospitalizations with missing data. To assess the implications of potential bias from missing data, we performed secondary analyses where i) all records with missing data were considered as non-cases of SARS-CoV-2 infection; ii) all missing data were considered as SARS-CoV-2 infection cases. LCA was performed using the 'poLCA' R package [6].

All analyses in this paper were performed using R version 4.1.2 [7].

Results

There were 11,852 hospitalizations occurring between Jan. 23, 2020, and Dec. 31, 2020, associated with at least one SARS-CoV-2 PCR test (regardless of result). In 444 of these 11,852 hospitalizations, we found at least one positive test, representing a frequency of 3.7% for confirmed SARS-CoV-2. Among the 444 PCR-confirmed cases, 434 had an ICD-10 discharge diagnosis code U07.1.

The sensitivity of code U07.1 was 97.8% (95% CI 95.9-98.9%), the specificity was 99.5% (95% CI 99.3-99.6%), PPV was 88.2% (95% CI 85.0-90.9%), and NPV was 99.9% (99.8-100.0%) Stratified analyses are presented in Tables 1-2. Operating characteristics were similar across sex, age groups, and calendar periods (Table 1). When considering the timing of the first PCR test, the specificity of ICD code U07.1 was slightly lower when the test was done more than 48 hours after admission (Table 2).

Results of the LCA are presented in Table 3. Data on vital signs and PCR were missing in 773 (6.0%) and 952 (7.4%) admissions, respectively. In the main analyses (when any individual with missing data was removed), the sensitivity of the U07.1 code was 97.5%, and the specificity was 100%. Results were similar in secondary analyses that considered missing data alternatively as indicative of all non- SARS-CoV-2 or as all SARS-CoV-2 cases.

Interpretation

Our results demonstrate the high sensitivity, specificity, and positive predictive value of ICD-10 code U07.1 in identifying lab-confirmed SARS-CoV-2 infection in hospital data. These operating characteristics were similar across sex, age groups, calendar periods, and comorbidities. The sensitivity of the ICD code was higher when the test was done between 24-48 hours of hospitalization. Latent class analyses (which do not assume a gold standard to estimate sensitivity/specificity) were comparable to our main findings.

Our results are consistent with studies in other jurisdictions evaluating the reliability of ICD-10 code U07.1 in identifying SARS-CoV-2 infection in hospitalization data. In the only other Canadian study, Wu et al. assessed the validity of SARS-CoV-2 ICD codes in Alberta provincial health databases from emergency admissions and inpatients in two cohorts linked to administrative health records [8]. They found that the sensitivity of ICD-10 code U07.1 for inpatients with positive PCR tests was 94.2% (95% CI 93.5-94.8%), and the PPV was 94.5% (95% CI 93.8-95.2%) [8].

Kadri et al. examined the reliability of ICD-10 code U07.1 in American administrative hospitalization data early in the pandemic [9]. Using a positive PCR test as the gold standard, they estimated the sensitivity of U07.1 at 98.0% (95% CI 97.6-98.4%) specificity at 99.0% (95% CI 98.9-99.1%) and PPV at 91.5% (95% CI 90.8-92.3%) [9]. They concluded that hospitals accurately code SARS-CoV-2 diagnoses, though they

5

advocated for continuing to assess the reliability of this code over time [9]. Subsequent studies found similar results [10-12]. Lynch et al. in the United States (US) reviewed Veterans Affairs medical records containing ICD diagnostic code U07.1 and calculated PPV, with PCR testing as the reference standard [13]. They also found high PPV in hospitalized patients (93.8%, 95% CI 91.8–95.6%) [13]. Bhatt et al. evaluated hospital discharge diagnoses between April 1, 2020, and July 31, 2020, in the Mass General Brigham health system (which includes Massachusetts General Hospital, Brigham and Women's Hospital, and other allied hospitals across Massachusetts) [14]. Compared to all other studies, they found a much lower sensitivity (49.2%, 95% CI 47.1-51.3%) for the hospital ICD code U07.1 compared to PCR test results; they estimated high specificity (99.4%, 95% CI 99.3-99.5%) and a PPV of 90.0% (95% CI 88.2-91.6%) [14]. They attributed the lower sensitivity to scribing delays at discharge, changes to testing criteria and interpretation differences when looking at test results and symptom presentation [14]. Bodilsen et al. also confirmed a high PPV for SARS-CoV-2 codes in Danish hospital data [15].

One of the motivations for using Canadian data for our study is that in this county, individuals have universal access to health care, so administrative records hold discharge diagnoses for all hospitalizations. Since only one other Canadian study has been published on this topic, and since it is well-known that ICD code use (and their validation) can vary greatly over jurisdictions, our study is important. Moreover, since the sensitivity and specificity of case definitions from administrative data can vary greatly over demographic and other variables, our stratified analyses offer some unique information. Finally, LCA has not often been used to validate ICD diagnostic codes from electronic health data, although often there is no easily available 'gold standard' for a disease state [16]. LCA analyses recognize that there is no true gold standard; as mentioned previously, even though PCR testing has largely been considered the reference standard for identifying SARS-CoV-2), its clinical sensitivity approaches 80% sensitivity. This is because someone with only a small amount of the virus in their nasal passages may have a negative test even if they have the virus somewhere else (like the gastrointestinal tract)[17]. Further, hospitalization for SARS-CoV-2 disease may occur later in the disease course when only a relatively small amount of the virus is present.

Overall, our findings confirm the validity of the new ICD-10 code. This is reassuring for research and surveillance activities relying on administrative hospitalization data. Moreover, our results were consistent even when no gold standard was assumed. However, our study has some potential limitations. First, PCR tests performed outside Ontario hospitals were unavailable, which may have caused some individuals tested for SARS-CoV-2 in outpatient settings to be identified as false positives (i.e., with a U07.1 diagnosis but without a positive test). However, this limitation is common in many published studies on this topic. Additionally, our study uses data from the first year of the pandemic, and as others have suggested, repeat analyses in years to come may yield further insights [18].

6

In conclusion, in a large Canadian electronic hospital database, ICD-10 code U07.1 accuracy of these codes can facilitate ongoing research initiatives, including determining how to manage future pandemics and costs or new SARS-CoV-2 waves and determining the effectiveness of immunization.

References

1. Verma AA et al. (2017). "Patient characteristics, resource use and outcomes associated with general internal medicine hospital care: the General Medicine Inpatient Initiative (GEMINI) retrospective cohort study." CMAJ Open 5(4): E842-E849.
2. Verma AA et al. (2020). "Assessing the quality of clinical and administrative data extracted from hospitals: The General Medicine Inpatient Initiative (GEMINI) experience." JAMIA. doi: 10.1093/jamia/ocaa225.
3. Canadian Coding Standards for Version 2018 ICD-10-CA and CCI. https://secure.cihi.ca/free_products/CodingStandards_v2018_EN.pdf
4. Statistics Canada. *Postal CodeOM Conversion File Plus (PCCF+) Version 7B, Reference Guide*. Ottawa: Statistics Canada, 2019.
5. How Good are COVID-19 (SARS-CoV-2) Diagnostic PCR Tests? https://www.cap.org/member-resources/articles/how-good-are-covid-19-sars-cov-2-diagnostic-pcr-tests
6. Linzer DA, Lewis JB (2011). "poLCA: An R Package for Polytomous Variable Latent Class Analysis." *Journal of Statistical Software*, **42**(10), 1–29. https://www.jstatsoft.org/v42/i10/.
7. R version 4.1.2 (https://www.R-project.org/)
8. Wu G, D'Souza AG, Quan H, Southern DA, Youngson E, Williamson T, Eastwood C, Xu Y. Validity of ICD-10 codes for COVID-19 patients with hospital admissions or ED visits in Canada: a retrospective cohort study. BMJ Open. 2022 Jan 21;12(1):e057838. doi: 10.1136/bmjopen-2021-057838. PMID: 35063962; PMCID: PMC8787827.
9. Kadri SS, Gundrum J, Warner S, Cao Z, Babiker A, Klompas M, Rosenthal N. Uptake and Accuracy of the Diagnosis Code for COVID-19 Among US Hospitalizations. JAMA. 2020 Dec 22;324(24):2553-2554. doi: 10.1001/jama.2020.20323. PMID: 33351033; PMCID: PMC7756233.
10. Moll K, Hobbi S, Zhou CK, Fingar K, Burrell T, Hernandez-Medina V, Obidi J, Alawar N, Anderson SA, Wong HL, Shoaibi A. Assessment of performance characteristics of COVID-19 ICD-10-CM diagnosis code U07.1 using SARS-CoV-2 nucleic acid amplification test results. PLoS One. 2022 Aug 18;17(8):e0273196. doi: 10.1371/journal.pone.0273196. PMID: 35980905; PMCID: PMC9387790.
11. Kluberg SA, Hou L, Dutcher SK, Billings M, Kit B, Toh S, Dublin S, Haynes K, Kline A, Maiyani M, Pawloski PA, Watson ES, Cocoros NM. Validation of diagnosis codes to identify hospitalized COVID-19 patients in health care claims data. Pharmacoepidemiol Drug Saf. 2022 Apr;31(4):476-480. doi: 10.1002/pds.5401. Epub 2022 Jan. 3. PMID: 34913208.
12. Rao S, Bozio C, Butterfield K, Reynolds S, Reese S, Ball S, Steffens A, et al. Accuracy of COVID-19-Like-Illness Diagnoses in Electronic Health Record Data. Pre-Print. JMIR Public Health Surveillance. May 3, 2022. https://preprints.jmir.org/preprint/39231
13. Lynch KE, Viernes B, Gatsby E, DuVall SL, Jones BE, Box TL, Kreisler C, Jones M. Positive Predictive Value of COVID-19 ICD-10 Diagnosis Codes Across Calendar Time and Clinical

7

Setting. Clin Epidemiol. 2021 Oct 27;13:1011-1018. doi: 10.2147/CLEP.S335621. PMID: 34737645; PMCID: PMC8558427.

14. Bhatt AS, McElrath EE, Claggett BL, Bhatt DL, Adler DS, Solomon SD, Vaduganathan M. Accuracy of ICD-10 Diagnostic Codes to Identify COVID-19 Among Hospitalized Patients. J Gen Intern Med. 2021 Aug;36(8):2532-2535. doi: 10.1007/s11606-021-06936-w. Epub 2021 Jun. 7. PMID: 34100236; PMCID: PMC8183587

15. Bodilsen J, Leth S, Nielsen SL, Holler JG, Benfield T, Omland LH. Positive Predictive Value of ICD-10 Diagnosis Codes for COVID-19. Clin Epidemiol. 2021 May 25;13:367-372. doi: 10.2147/CLEP.S309840. PMID: 34079379; PMCID: PMC8164665.

16. Bernatsky S, Lix L, Hanly JG, Hudson M, Badley E, Peschken C, Pineau CA, Clarke AE, Fortin PR, Smith M, Bélisle P, Lagace C, Bergeron L, Joseph L. Surveillance of systemic autoimmune rheumatic diseases using administrative data. Rheumatol Int. 2011 Apr;31(4):549-54. doi: 10.1007/s00296-010-1591-2. Epub 2010 Jul 28. PMID: 20665025.

17. Grassia R, Testa S, Pan A, Conti CB. SARS-CoV-2 and gastrointestinal tract: The dark side of the pandemic. Dig Liver Dis. 2020 Jul;52(7):700-701. doi: 10.1016/j.dld.2020.04.028. Epub 2020 Apr 29. PMID: 32423849; PMCID: PMC7188635.

18. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, Del Campo R, Ciapponi A, Sued O, Martinez-García L, Rutjes AW, Low N, Bossuyt PM, Perez-Molina JA, Zamora J. False-negative results of initial RT-PCR assays for COVID-19: A systematic review. PLoS One. 2020 Dec 10;15(12):e0242958. doi: 10.1371/journal.pone.0242958. PMID: 33301459; PMCID: PMC7728293.

8

Table 1 Operating characteristics of ICD*U07.1 (lab-confirmed COVID) within hospital diagnostic codes (PCR as reference) stratified by sex, age and urban-versus-rural residence

Parameter	Estimate (95% CI)
Female	
Sensitivity	97.2% (93.6%-99.1%)
Specificity	99.5% (99.2%-99.6%)
Positive predictive value	86.1% (80.5%-90.5%)
Negative predictive value	99.9% (99.8%-100.0%)
Male	
Sensitivity	98.1% (95.7%-99.4%)
Specificity	99.5% (99.3%-99.7%)
Positive predictive value	89.7% (85.6%-92.9%)
Negative predictive value	99.9% (99.8%-100.0%)
< 50 years	
Sensitivity	95.1% (87.8%-98.6%)
Specificity	99.5% (99.1%-99.7%)
Positive predictive value	84.6% (75.5%-91.3%)
Negative predictive value	99.8% (99.6%-100.0%)
50-75 years	
Sensitivity	98.5% (95.6%-99.7%)
Specificity	99.5% (99.2%-99.6%)
Positive predictive value	87.1% (81.9%-91.2%)
Negative predictive value	99.9% (99.8%-100.0%)
> 75 years	
Sensitivity	98.2% (94.8%-99.6%)
Specificity	99.6% (99.3%-99.8%)
Positive predictive value	91.5% (86.4%-95.2%)
Negative predictive value	99.9% (99.8%-100.0%)
Urban	
Sensitivity	97.6% (95.6%-98.8%)
Specificity	99.5% (99.3%-99.6%)
Positive predictive value	87.8% (84.5%-90.7%)
Negative predictive value	99.9% (99.8%-100.0%)
Rural	
Sensitivity	100.0% (39.8%-100.0%)
Specificity	99.7% (98.4%-100.0%)
Positive predictive value	80.0% (28.4%-99.5%)
Negative predictive value	100.0% (99.0%-100.0%)

*ICD=international classification of diseases PCR=polymerase chain reaction

9

Table 2 Operating characteristics of ICD*U07.1 (lab-confirmed COVID) within hospital diagnostic codes (PCR as reference) stratified by sex, age and urban-versus-rural residence

Parameter	Estimate (95% confidence interval)
Timing of first PCR	
Before admission	
Sensitivity	98.7% (96.2%-99.7%)
Specificity	99.9% (99.7%-100.0%)
Positive predictive value	97.8% (95.0%-99.3%)
Negative predictive value	99.9% (99.8%-100.0%)
Between 0-24h of admission	
Sensitivity	95.8% (90.4%-98.6%)
Specificity	99.7% (99.6%-99.9%)
Positive predictive value	88.3% (81.4%-93.3%)
Negative predictive value	99.9% (99.8%-100.0%)
Between 24-48 of admission	
Sensitivity	100.0% (63.1%-100.0%)
Specificity	99.0% (96.5%-99.9%)
Positive predictive value	80.0% (44.4%-97.5%)
Negative predictive value	100.0% (98.2%-100.0%)
Beyond 48h of admission	
Sensitivity	97.8% (92.4%-99.7%)
Specificity	94.2% (92.1%-95.9%)
Positive predictive value	71.4% (62.7%-79.1%)
Negative predictive value	99.7% (98.8%-100.0%)
Calendar period of admission	
Jan-Apr 2020	
Sensitivity	99.4% (97.0%-100.0%)
Specificity	98.6% (97.9%-99.2%)
Positive predictive value	90.5% (85.5%-94.2%)
Negative predictive value	99.9% (99.6%-100.0%)
May-Aug 2020	
Sensitivity	96.2% (90.5%-99.0%)
Specificity	99.5% (99.2%-99.7%)
Positive predictive value	79.5% (71.5%-86.2%)
Negative predictive value	99.9% (99.8%-100.0%)
Sep-Dec 2020	
Sensitivity	96.8% (92.8%-99.0%)
Specificity	99.7% (99.6%-99.9%)
Positive predictive value	92.2% (87.0%-95.8%)
Negative predictive value	99.9% (99.8%-100.0%)

10

Table 3 Operating characteristics of ICD*U07.1 (lab-confirmed COVID) within hospital diagnostic codes, using latent class analyses

Main analyses (subjects with missing data removed)	
Parameter	**Estimate (95% confidence interval)**
Sensitivity	97.5% (95.9-99.0%)
Specificity	100% (100-100%)
Alternative 1 (missing data considered as non-COVID cases)	
Parameter	**Estimate (95% confidence interval)**
Sensitivity	100.0% (100-100%)
Specificity	99.4% (98.0-100%)
Alternative 2 (missing data considered as COVID cases)	
Parameter	**Estimate (95% confidence interval)**
Sensitivity	96.2% (94.4-98.0%)
Specificity	99.9% (99.8-100%)

11