Use of chest imaging in COVID-19: a rapid advice guide

Web Annex A. Imaging for COVID-19: a rapid review

Chou R, Pappas M, Buckley D, McDonagh M, Totten A, Flor N, Sardanelli F, Dana T, Hart E, Wasson N, Nelson H

Pacific Northwest Evidence-Based Practice Center, Oregon Health and Science University, Portland, Oregon, USA

The named authors alone are responsible for the views expressed in this publication.

This publication forms part of the WHO rapid advice guide entitled *Use of chest imaging in COVID-19*. It is being made publicly available for transparency purposes and information, in accordance with the *WHO handbook for guideline development*, 2nd edition (2014).

WHO reference number: WHO/2019-nCoV/Clinical/Radiology_imaging/Web_Annex_A/2020.1

Contents

Tables

Figure

Key findings

- No study evaluated the diagnostic accuracy of chest imaging in asymptomatic patients possibly infected with SARS-CoV-2.
- In symptomatic patients in high COVID-19 prevalence cohorts, chest computed tomography (CT) appears to be associated with high sensitivity but low specificity, resulting in weak positive likelihood ratios and stronger negative likelihood ratios. This indicates that in these settings, negative imaging findings might be useful for ruling out COVID-19, but positive imaging findings are not useful for ruling in COVID-19.
- Evidence on the diagnostic accuracy of chest x-ray (CXR) was very limited, but suggests lower sensitivity and possibly higher specificity than chest CT for diagnosing COVID-19.
- Evidence on the diagnostic accuracy of lung ultrasound (LUS) was limited to one study that used chest CT findings as the reference standard.
- No study evaluated the effects of chest imaging on health outcomes.
- Studies on the utility of chest imaging for predicting health outcomes was limited and inconclusive.
- No study evaluated the diagnostic accuracy of chest imaging for diagnosis of pulmonary thromboembolism in patients with COVID-19.
- One study reported a prevalence of pulmonary thromboembolism of 30% in patients with COVID-19 and another study reported a prevalence of 23%, but included patients with diagnosed as well as suspected COVID-19.
- No study evaluated chest imaging in COVID-19 patients to inform decisions regarding discharge.

1 Background and Key Questions

A cluster of pneumonia cases in Wuhan, China was first reported to the World Health Organization (WHO) China Country office on December 31, 2019.[1] Soon thereafter, a novel coronavirus was identified as the causative agent.[2-4] This virus was named severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) and the associated disease was named coronavirus disease 2019 (COVID-2019).[5] Since December 2019, COVID-2019 has rapidly spread from Wuhan to other parts of China and throughout the world. On January 30, 2020, WHO declared the outbreak a Public Health Emergency of International Concern[6] and on March 11, 2020, WHO characterized the outbreak as a pandemic.[7]

A variety of chest imaging findings have been described in patients with COVID-19. Use of imaging could be useful for the diagnosis of patients with suspected COVID-19 and in patients diagnosed with COVID-19, to inform management. The purpose of this rapid review is to summarize the evidence on imaging with chest computed tomography (CT), chest x-ray (CXR), and lung ultrasound (LUS) for diagnosis of COVID-19, management of COVID-19 (including effects on health outcomes and prediction of outcomes), and diagnosis of pulmonary thromboembolism in patients with COVID-19.

The following key questions (KQs) guided our rapid review:

Key Question 1. In asymptomatic contacts of patients with COVID-19, and in contexts where laboratory testing (reverse transcription-polymerase chain reaction [RT-PCR]) is not available/results are delayed/results are initially negative, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used for the diagnostic workup of COVID-19?

Key Question 2. In symptomatic patients with suspected COVID-19, and in contexts where laboratory testing (RT-PCR) is not available/results are delayed/results are initially negative, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used for the diagnostic workup of COVID-19?

Key Question 3. In patients with suspected or confirmed COVID-19, not currently hospitalized and with mild symptoms, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used to support decision on hospital admission versus home discharge?

Key Question 4. In patients with suspected or confirmed COVID-19, not currently hospitalized and with moderate to severe symptoms, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used to support decision on regular ward admission versus intensive care unit admission?

Key Question 5. In patients with suspected or confirmed COVID-19, currently hospitalized and with moderate or severe symptoms, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used to modify the therapeutic management?

Key Question 6. In patients with suspected or confirmed COVID-19 and clinical deterioration and/or suspicion of pulmonary embolism, should imaging (including CT pulmonary angiography, lower extremity venous ultrasound) vs. no imaging be used to diagnose pulmonary embolism?

Key Question 7. In patients with COVID-19 whose symptoms resolved, should chest imaging (including CXR, CT scan, LUS) be added to vs. not added to laboratory criteria to support decision on discharge home vs. no discharge home?

2 Methods

The Pacific Northwest Evidence-based Practice Center team at Oregon Health & Science University conducted this rapid review. The rapid review approach utilizes systematic searches to identify studies but applies streamlined systematic review methods and focuses on the available evidence, due to time constraints.[8] For this review, these modified methods included:

- Study protocol not registered in a systematic review registry (e.g., PROSPERO)
- Search of grey literature limited to one website with non-peer-reviewed (pre-print) manuscripts
- Excluded non-English language case series
- Excluded case series with less than 20 patients
- Single reviewer assessment of study limitations and data abstraction, with a second reviewer verifying assessments and accuracy
- Formal quality assessment focused on studies reporting diagnostic accuracy
- Key methodological limitations summarized for studies on the association between imaging findings and health outcomes and the prevalence of pulmonary thromboembolism on imaging, given the low quality of the evidence

2.1 Search strategy

A medical librarian searched PubMed MEDLINE and Elsevier Embase (from 2003 through April 15, 2020) for relevant studies. Literature searches were reviewed by a second librarian. Search strategies are available in Annex A**.** We also searched the WHO Database of Publications on Coronavirus Disease,[9] a database of Chinese language COVID-19 studies maintained at the WHO Collaborating Centre for Guideline Implementation and Knowledge Translation, Lanzhou University, China, and the medRxiv preprint server[10] for preliminary reports; and reviewed reference lists of systematic reviews and included studies. Surveillance was conducted daily on MEDLINE and weekly on the other databases; this report includes results of surveillance through April 29, 2020.

2.2 Eligibility criteria

Table 1 summarizes the inclusion criteria used to select studies; detailed criteria for each KQ are described in Annex B. The population for KQs 1 and 2 was subjects suspected of having COVID-19 or SARS-CoV-2 infection (not necessarily meeting the case definition for COVID-19) and for KQs 3 to 7 it was patients diagnosed with COVID-19 or with SARS-CoV-2 infection. For all KQs, the imaging modalities were chest CT, CXR, and LUS. For studies reporting diagnostic accuracy, the comparison was one or more imaging modalities against a reference standard for COVID-19 and for studies reporting health outcomes, the comparison was no imaging or another imaging modality. We included randomized trials, cohort studies, case-control studies, and English language case series that reported diagnostic accuracy outcomes (e.g., sensitivity, specificity, positive predictive value, negative predictive value, or area under the receiver operating characteristic curve at receiver operating characteristic analysis [AUROC]) and randomized trials, cohort and case-control studies that reported the association between imaging findings and health outcomes. We excluded exploratory or descriptive studies

that described imaging findings in patients with COVID-19 but did not define a positive imaging test for COVID-19, which is necessary to estimate diagnostic accuracy. Studies on use of artificial intelligence to interpret imaging were included. Because few studies were on diagnostic accuracy of imaging for pulmonary thromboembolism in patients with COVID-19, we also included imaging series reporting the prevalence of pulmonary embolus in patients with COVID-19 who underwent contrast-enhanced CT pulmonary angiography.

Table 1. Inclusion criteria

	Inclusion	Exclusion
Population	KQ 1 and 2: Subjects suspected of having COVID-19 or SARS-CoV-2 infection KQ 3 to 7: Patients diagnosed with COVID-19 or SARS-CoV-2 infection	• Non-COVID-19 only
Interventions	• Chest computed tomography • CXR • Lung ultrasound • Note: For diagnosis of pulmonary thromboembolism, chest imaging could be performed in conjunction Doppler ultrasound or D-dimer	• Other imaging modalities
Outcomes	KQ 1, 2, 5, 6: Measure of diagnostic accuracy KQ 3 to 5, 7 • Length of ED stay (KQ 3) • Mortality • Treatment failure • Need for and length of hospital stay • Need for and length of ICU stay • Need for and length of respiratory support • Complications of imaging Secondary outcomes • Measures of association between imaging findings and health outcomes • COVID-19 transmission to health care workers • Measure of availability of care, access to care, and quality of care KQ 6: Prevalence of pulmonary thromboembolism (secondary outcome)	• Other outcomes
Study designs	• Randomized, nonrandomized and controlled clinical trials • Cohort studies • Case-control studies • Cross-sectional studies • Case series and imaging series (KQ 1, 2, 6)	• Systematic reviews (reference lists of relevant reviews checked for relevant primary studies) • Case reports • Case series with <20 patients • Modelling studies
Language	• Case series restricted to English language; otherwise no language restrictions	

Abbreviations: COVID-19=coronavirus disease 2019; CXR=chest x-ray; ED=emergency department; ICU=intensive care unit; KQ=Key Question; SARS-CoV-2=severe acute respiratory syndrome coronavirus-2

2.3 Screening process

A team of experienced systematic reviewers screened all titles and abstracts using the predefined inclusion and exclusion criteria. Each citation and abstract was reviewed by one team member for potential inclusion and full-text review. All citations were reviewed by a second team member; disagreements were resolved by consensus. Each full-text article was reviewed by one team member for potential inclusion; excluded articles were reviewed by a second team member to verify the exclusion decision. Literature screening was conducted using EndNote® software, version X9. Non-peer-reviewed articles were included, due to the recent nature of the COVID-19 pandemic and availability of relevant non-peer-reviewed articles. Potentially relevant non-English language studies were reviewed and translated by a native Chinese speaker at the WHO Collaborating Centre for Guideline Implementation and Knowledge Translation, Lanzhou University, China.

2.4 Outcomes

Outcomes were diagnostic accuracy and health outcomes (length of emergency department [ED] stay, mortality, treatment failure, need for and length of hospital stay, need for and length of intensive care unit [ICU] stay, need for and length of respiratory support, and harms of imaging) (Table 1). Secondary outcomes were COVID-19 transmission to health care workers, availability of care, access to care, and quality of care.

Outcomes for assessing the association between imaging findings and health outcomes were measures of diagnostic accuracy or risk estimates (relative risk, odds ratio, or hazards ratio). Adjusted risk estimates were utilized when available. For KQ 6, the prevalence of pulmonary thromboembolism on imaging was added to the protocol as a secondary outcome, due to the lack of evidence on effects of imaging for pulmonary embolus on health outcomes.

2.5 Data extraction

One experienced team member extracted data from included studies into standardized tables. A second reviewer checked the extracted data for completeness and accuracy. The data items extracted were: study author, year; eligibility criteria; country; population characteristics (sample size, age, sex, symptoms, interval since symptom onset); imaging test; criteria for a positive imaging test; imaging reader; and results. Selected CT parameters (CT reconstruction slice thickness, volumetric or spiral acquisition of images, use of dose reduction systems, and intravenous administration of iodinated contrast medium [for diagnosis of pulmonary thromboembolism]) were abstracted, with input from radiologists on the review team. For studies reporting diagnostic accuracy, we calculated the number of true positives, false positives, false negative, and true negatives from data in the study if necessary, and 95% confidence intervals (CIs) for sensitivity, specificity, positive predictive value, and negative predictive value were calculated using an online calculator if necessary.[11]

2.6 Critical appraisal

We critically appraised studies of diagnostic accuracy using criteria adapted from QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies –Version 2) and the U.S. Preventive Services Task Force.[12,13] Additional key limitations of studies were also noted, such as small sample sizes, inability to calculate diagnostic accuracy parameters, issues with selection of controls (e.g., controls from another country than cases) or use of inexperienced imaging readers. We summarized key limitations in studies on the association between imaging findings and health outcomes and the prevalence of pulmonary thromboembolism but did not use a formal risk assessment instrument, given the type of evidence available (e.g., imaging series for prevalence of pulmonary thromboembolism) and low quality of the studies.

2.7 Data synthesis

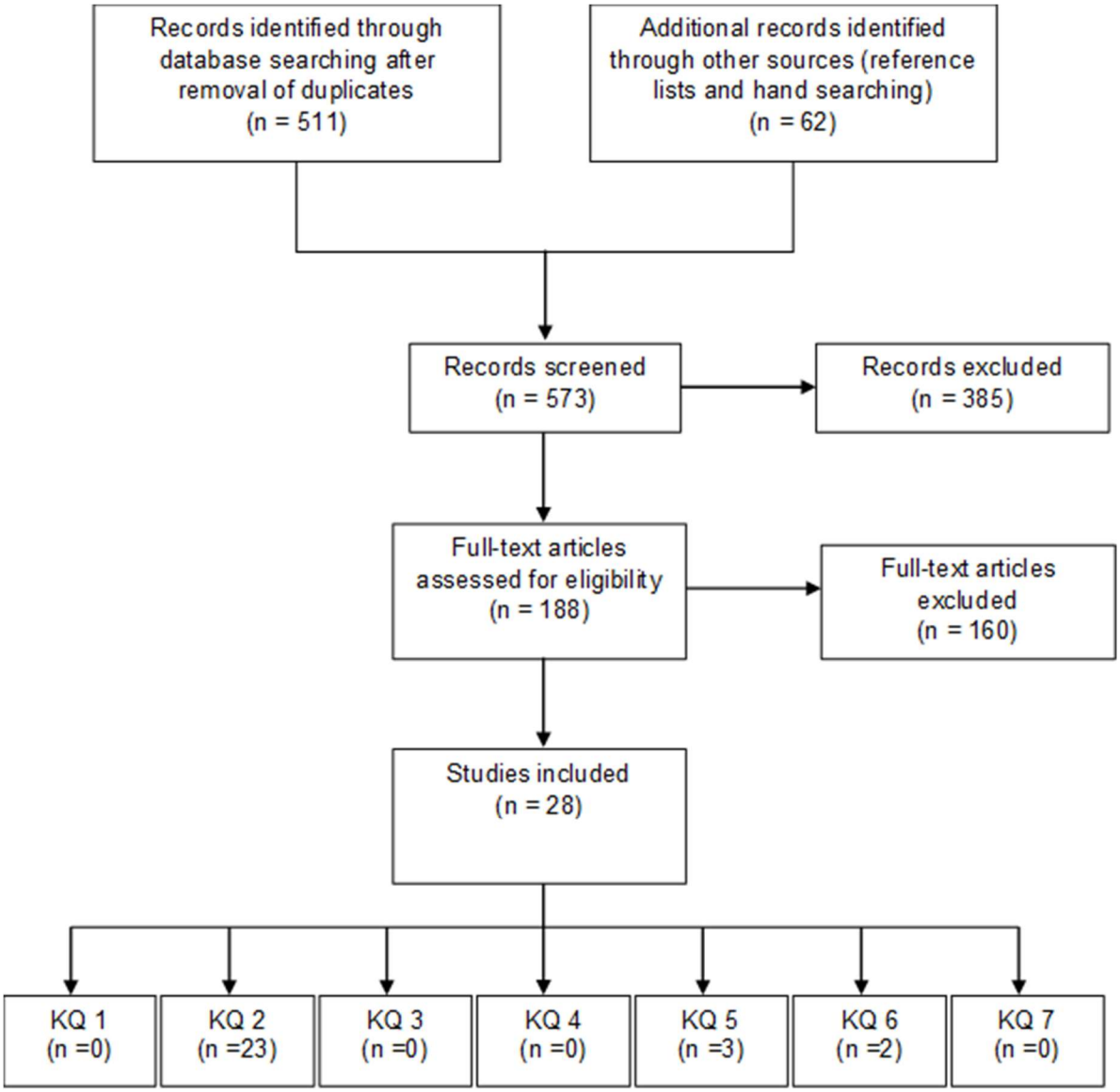
Results were synthesized narratively and in tabular form structured by included studies, following a best evidence approach. GRADE was used to summarize the strength of evidence by KQ. Quantitative synthesis was not possible due to methodological limitations in the studies, variability in study designs, and heterogeneity in populations, comparisons, and analytic methods.

3 Results

Our searches identified 28 studies that met inclusion criteria.[14-41] No study addressed the diagnostic accuracy of imaging in asymptomatic subjects, e.g. contacts of persons with COVID-19 (KQ 1). Twenty-three studies addressed the diagnostic accuracy of imaging in symptomatic patients suspected of having COVID-19 or SARS-CoV-2 infection (KQ 2). Of these, 19 studies evaluated chest CT, three studies evaluated CXR, and one study evaluated LUS. Seven studies were cohort studies, seven were case-control studies, and nine were case series. No study reported health outcomes associated with use of imaging. Three studies reported the association between imaging findings and health outcomes; all of these evaluated use of chest CT. No study reported the diagnostic accuracy of imaging for pulmonary thromboembolism in patients with COVID-19 (KQ 6). Two studies reported the prevalence of pulmonary thromboembolism on imaging in patients with COVID-19. One Chinese language study was translated into English.[39] Six studies were published as pre-peer review articles.

The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram summarizes the study selection process (Figure 1), including the reasons for exclusion. A list of included studies by KQ is provided in Annex C.

Figure 1: PRISMA flow diagram



3.1 Key Question 1. In asymptomatic contacts of patients with COVID-19, and in contexts where laboratory testing (RT-PCR) is not available/results are delayed/results are initially negative, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used for the diagnostic workup of COVID-19?

No study evaluated the diagnostic accuracy of imaging in asymptomatic contacts of patients with COVID-19.

3.2 Key Question 2: In symptomatic patients with suspected COVID-19, and in contexts where laboratory testing (RT-PCR) is not available/results are delayed/results are initially negative, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used for the diagnostic workup of COVID-19?

Twenty-three studies evaluated the diagnostic accuracy of imaging in symptomatic patients with suspected COVID-19 against a reference standard (Annex D).[14-20,22,24,25,27,29-40] Of 19 studies of chest CT, six were cohort studies (N=1846),[15,19,24,32,35,39] seven were case-control studies (total number of cases=975, total number of controls=925),[16,17,22,27,29,33,40] and six studies were case series (N=317).[14,25,30,31,37,38] Of three studies on CXR, one was a cohort study and two studies were case series.[20,34,36] One cohort study evaluated LUS.[18]

The studies had important methodological limitations (Annex E). Seventeen studies were assessed as being at high risk of bias and six studies at moderate risk of bias; no study was assessed as being at low risk of bias. All cohort studies except for one[19] were retrospective and the case-control studies were at risk for spectrum bias. In addition, the case-control studies did not match cases and controls on potential confounding factors such as age, sex, body mass index, or comorbidities. In some studies, cases and controls were selected from different populations (e.g., cases and controls from different countries or cases from patients in a study hospital and controls selected from an imaging database). A serious limitation of the case series of COVID-19 patients was that only sensitivity could be estimated, resulting in incomplete information regarding diagnostic accuracy, because sensitivity and specificity are related measures. There were also methodological limitations across study designs. The studies provided limited information regarding clinical presentation, such as the severity of symptoms at presentation. Few studies defined specific criteria for a positive imaging test for COVID-19. In 11 studies, it was unclear if the reference standard included serial RT-PCR or clinical follow-up to diagnose COVID-19. Relying on a single RT-PCR as the reference standard is likely to result in some misclassification of patients due to the potential for false-negative RT-PCR assays, particularly early in the disease course. The study of LUS compared performance using chest CT as the reference standard.

3.2.1 Chest CT

Nineteen studies evaluated the diagnostic accuracy of chest CT for COVID-19 (Table 2).[14-17,19,22,24,25,27,29-33,35,37-40] CT imaging slice thickness varied across studies, though most reported a reconstruction slice thickness of less than 1.25 mm. Nine studies reported use of radiation dose

reduction systems (e.g., automatic current tube modulation). Five studies evaluated use of artificial intelligence to interpret CT images.[17,27,29,33,40]

Among the non-artificial intelligence studies, six cohort studies (N=1846) evaluated the diagnostic accuracy of chest CT in patients with suspected COVID-19.[15,19,24,32,35,39] Three studies were conducted in China and one each in Italy, Belgium, and the Netherlands. The proportion of patients diagnosed with COVID-19 in the cohorts ranged from 39% to 85%, with the exception of one study in which the proportion diagnosed with COVID-19 was 19%.[39] The largest study (n=1014, COVID-19 prevalence 59%), from China (Wuhan), reported an overall sensitivity (based on CT read as positive by a radiologist) of 0.96 (95% CI 0.95 to 0.98), specificity of 0.25 (95% CI 0.22 to 0.30), positive predictive value of 0.65 (95% 0.62 to 0.58), and negative predictive value of 0.83 (95% CI 0.76 to 0.89).[15] Diagnostic accuracy was similar when patients were stratified by age less than 60 years versus 60 years or older, or by sex. The reference standard was a positive initial RT-PCR for SARS-CoV-2; re-classification of 15 patients as COVID-19 cases based on a subsequent positive RT-PCR had little impact on sensitivity (0.97) or specificity (0.26).

Three smaller cohort studies from Italy and China (n=103 to 274, COVID-19 prevalence 19% to 85%) also reported high sensitivity (0.91 to 0.97) and low specificity (0.53 to 0.68), based on imaging read as positive by a radiologist.[19,35,39] The positive and negative predictive values varied, due to differences in sensitivity and specificity as well as in the prevalence of COVID-19 (Table 2). In one of the cohort studies, using a scoring system to classify CT findings as positive or negative instead of or in addition to radiologist qualitative interpretation did not improve discrimination (ability to distinguish patients with COVID-19 from those without COVID-19), based on the AUROC.[39]

One cohort study (n=192, COVID-19 prevalence 43%) from Belgium that used a low radiation dose protocol (mean 0.58 mSV) was an outlier in that it reported both high sensitivity (0.87, 95% CI 0.80 to 0.99) and specificity (0.95, 95% CI 0.89 to 0.982) based on imaging read as positive by a radiologist, for a positive predictive value of 0.91 (95% CI 0.85 to 0.97) and negative predictive value of 0.90 (95% CI 0.85 to 0.96).[24] One other cohort study from the Netherlands (n=105, prevalence of COVID-19 50%) reported high discrimination for COVID-19 (AUROC 0.91, 95% CI 0.85 to 0.97) when images were categorized using the COVID-19 Reporting and Data System (CO-RADS), but did not report sensitivity and specificity at specific CO-RADS thresholds.[32]

Two case-control studies from China evaluated the diagnostic accuracy of CT for diagnosis of COVID-19 (Table 2).[16,22] The larger study (cases=219, controls=205) had serious methodological limitations, such as selection of cases with COVID-19 from China but controls with non-COVID-19 pneumonia from another country and exclusion of COVID-19 cases with negative CT scans.[16] It also reported high variability in diagnostic accuracy across seven radiologists (3 from China and 4 from the United States), with sensitivity ranging from 0.67 to 0.97 and specificity ranging from 0.07 to 1.00. A small case-cohort study (38 patients in validation set, number of cases and controls unclear) reported sensitivity of 1.0 and specificity of 0.37 using a multivariate logistic regression model based on radiological features to classify images as positive for COVID-19.[22]

Six case series (n=21 to 108) reported sensitivities for COVID-19 that ranged from 0.81 to 0.97 (Table 2).[14,25,30,31,37,38] All of the case series were from China. Reporting of clinical characteristics was limited and only one of the case series[25] clearly enrolled a consecutive series of patients.

Five studies evaluated the diagnostic accuracy of chest CT for diagnosis of COVID-19 using artificial intelligence to interpret imaging findings (Table 2).[17,27,29,33,40] All of the studies used a case-control design and each evaluated a different artificial intelligence algorithm. One study[17] selected cases and controls from China and the United States and in one study[27] cases were from China and some controls were selected from an international imaging database. Information about the clinical characteristics of controls was limited. The sensitivity of CT using artificial intelligence ranged from 0.75 to 0.95 and the specificity ranged from 0.86 to 0.96. In two studies, the diagnostic accuracy of artificial intelligence was superior to the accuracy of image reading radiologists; the other studies did not compare performance of artificial intelligence to radiologists.[17,33]

Table 2. Diagnostic Accuracy of Chest Computed Tomography to Diagnose COVID-19

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging Reader	Sensitivity (95%CI) Specificity (95%CI) AUROC (95%CI)
Cohort studies				
Ai, et al., 2020[15]	Suspected of COVID-19; underwent both chest CT imaging and SARS-CoV-2 RT-PCR; time interval between CT and RT-PCR ≤7 days.	Imaging read as positive for COVID-19	2 radiologists who came to consensus	Sensitivity: Overall: 0.96 (0.95 to 0.98) <60 years: 0.96 (0.94 to 0.98) >=60: 0.97 (0.94 to 0.99) Female: 0.97 (0.95 to 0.99) Male: 0.96 (0.93 to 0.98) Specificity: Overall: 0.25 (0.22 to 0.30) <60 years: 0.26 (0.22 to 0.32) ≥60: 0.22 (0.16 to 0.31) Female: 0.30 (0.25 to 0.37) Male: 0.19 (0.14 to 0.25) AUROC: Not reported
Caruso, et al., 2020[19]	Suspected COVID-19 patients with fever and respiratory symptoms such as cough, and dyspnea; patients with mild respiratory symptoms and close contact with a confirmed COVID-19 patient; or patients with a previously positive test result. Patients who underwent chest CT with contrast for vascular indication were excluded.	CT positive for viral pneumonia using clinically available dedicated application (Thoracic VCAR v13.1, GE)	Two radiologists in consensus evaluated images using a clinically available dedication application for diagnosis of viral pneumonia	Sensitivity: 0.97 (0.88 to 0.99) Specificity: 0.56 (0.45 to 0.66) AUROC: Not reported
Dangis, et al., 2020[24]	Possible COVID-19 infection and both SARS-CoV-2 RT-PCR and low-dose chest CT at presentation.	Imaging classified as positive for COVID-19 (scored based on the presence of findings as presented by Ng et al and Shi et al)	Two radiologists with 8 and 7 years of experience	Sensitivity: 1: 0.87 (0.80 to 0.98) 2: 0.96 (0.91 to 0.999) Specificity: 1: 0.94 (0.89 to 0.982) 2: 0.93 (0.88 to 0.98) AUROC: Not reported

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging Reader	Sensitivity (95%CI) Specificity (95%CI) AUROC (95%CI)
Prokop, et al., 2020[32]	Presenting to the emergency department with suspected COVID-19 based on lower respiratory tract infection symptoms including cough and clinically relevant dyspnea requiring hospital admission with or without fever >38 degrees C; CT performed and SARS-CoV-2 RT-PCR within 5 days of CT.	Based on categorization using the COVID-19 Reporting and Data System, threshold not utilized (only AUROC reported)	Average of 8 radiologists (4 had <5 years experience; the remainder had 5 to 27 years experience)	Sensitivity: Not reported Specificity: Not reported AUROC: 1: 0.91 (0.85 to 0.97) 2: 0.95 (0.91 to 0.99)
Wen, et al., 2020[35]	Under investigation for COVID-19; excluded persons with fever >14 days but no acute respiratory infection signs or symptoms or exposure history; acute respiratory infection signs or symptoms >14 days but no exposure history; and acute respiratory infection symptoms in the last 14 days but no exposure history, laboratory tests, or other examination sufficient to exclude COVID-19. All patients were hospitalized >=2 weeks.	CT read as positive for COVID-19; Fleischner Society lexicon used	3 radiologists with 8 to 15 years of experience; disagreements resolved through discussion and consensus	Sensitivity: 0.93 (0.86 to 0.97) Specificity: 0.53 (0.27 to 0.79) AUROC: Not reported
Yang, et al., 2020[39]	Evaluated for possible COVID-19 with RT-PCR for SARS-CoV-2 and CT.	A: Imaging read as positive B: Imaging total score>=2 C: Imaging read as positive and score >=2 D: Imaging read as positive or score >=2	2 radiologists jointly reviewed CT images	Sensitivity: A: 0.91 (0.79 to 0.97) B: 0.89 (0.77 to 0.96) C: 0.79 (0.66 to 0.89) D: 1.0 (0.93 to 1.0) Specificity: A: 0.68 (0.62 to 0.74) B: 0.32 (0.26 to 0.38) C: 0.50 (0.40 to 0.60) D: 0.24 (0.18 to 0.30) AUROC: A: 0.79 (0.86 to 0.73) B: 0.60 (0.52 to 0.68) C: 0.78 (0.85 to 0.71) D: 0.62 (0.69 to 0.54)

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging Reader	Sensitivity (95%CI) Specificity (95%CI) AUROC (95%CI)
Case-control studies				
Bai, et al., 2020[16]	Cases (COVID-19): Positive SARS-CoV-2 PCR and abnormal chest CT findings. Controls (possible viral pneumonia): Positive results on Respiratory Pathogen Panel and chest CT final impression with the word "pneumonia" within 7 days.	Imaging read as positive for COVID-19	A: Radiologist 1 (China) B: Radiologist 2 (China) C: Radiologist 3 (China) D: Radiologist 4 (United States) E: Radiologist 5 (United States) F: Radiologist 6 (United States) G: Radiologist 7 (United States)	Sensitivity: A: 0.72 (0.66 to 0.78) B: 0.72 (0.65 to 0.78) C: 0.94 (0.90 to 0.97) Age to matched sample (n=58) A: 0.80 (0.61 to 0.92) B: 0.67 (0.47 to 0.83) C: 0.97 (0.83 to 1.00) D: 0.93 (0.78 to 0.99) E: 0.83 (0.65 to 0.94) F: 0.73 (0.54 to 0.88) G: 0.70 (0.51 to 0.85) Specificity: A: 0.94 (0.89 to 0.97) B: 0.88 (0.83 to 0.92) C: 0.24 (0.13 to 0.30) Age to matched sample (n=58) A: 1.00 (0.88 to 1.00) B: 0.93 (0.76 to 0.99) C: 0.07 (0.01 to 0.24) D: 1.00 (0.88 to 1.00) E: 0.93 (0.76 to 0.99) F: 0.93 (0.76 to 0.99) G: 1.00 (0.88 to 1.00) AUROC: Not reported
Chen, et al., 2020[22]	Cases: Positive for SARS-CoV-2 RT-PCR. Controls: Pneumonia without COVID-19.	Classified as positive for COVID-19 by multivariate logistic regression model based on radiological features	2 senior radiologists	Sensitivity: 1.0 (CI Not Reported) Specificity: 0.37 (CI Not Reported) AUROC: 0.81 (0.67 to 0.95)

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging Reader	Sensitivity (95%CI) Specificity (95%CI) AUROC (95%CI)
Case series				
Ai, et al., 2020[14]	Positive SARS-CoV-2 RT-PCR and hospitalized.	"Relatively obvious imaging features of COVID-19"	Not described	Sensitivity: 0.89 (0.81 to 0.94) Specificity: Not reported AUROC: 0.74 (0.69 to 0.80)
Fang, et al., 2020[25]	Positive SARS-CoV-2 RT-PCR with 1) travel or residential history in Wuhan/local endemic area or contact with individuals from these areas within 14 days and 2) fever or acute respiratory symptoms of unknown case. Chest CT and RT-PCR testing within 3 days or less.	CT compatible with viral pneumonia (including typical and atypical CT manifestations)	Not described	Sensitivity: 0.95 (0.90 to 1.00) Specificity: Not reported AUROC: Not reported
Li and Xia., 2020[30]	Positive RT-PCR for COVID-19.	CT read as viral pneumonia	Read by 1 radiologist and checked by another	Sensitivity: 0.96 (0.87 to 0.995) Specificity: Not reported AUROC: Not reported
Long, et al., 2020[31]	Fever >38°C and COVID-19 pneumonia suspicion, underwent thin-section CT of the chest and RT-PCR examinations.	Abnormal CT consistent with COVID-19	2 radiologists; disagreement resolved by consensus	Sensitivity: 0.97 (0.85 to 0.999) Specificity: Not reported AUROC: Not reported
Xu, et al., 2020[37]	Positive RT-PCR for COVID-19 and CT scan available.	CT positive based on lesion distribution, lesion location, lesion density (ground glass opacity, consolidation, or mixed), thickness of interlobular and intralobular septa, enlarged lymph nodes within the mediastinum and pleural effusion	3 experienced radiologists; disagreements resolved by consensus	Sensitivity: 0.82 (0.69 to 0.91) Specificity: Not reported AUROC: Not reported

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging Reader	Sensitivity (95%CI) Specificity (95%CI) AUROC (95%CI)
Xu, et al, 2020[38]	Confirmed SARS-CoV-2 infection.	CT read as positive for COVID-19	2 radiologists with 8 to 10 years experience, disagreements resolved by 3rd radiologist	Sensitivity: Total: 0.81 (0.58 to 0.95) Severe: 1.0 Mild: 0.71 (0.42 to 0.92) Specificity: Not reported AUROC: Not reported
Artificial intelligence*				
Bai, et al., 2020[17]	Cases: Positive SARS-C0V-2 PCR and abnormal chest CT. Controls: Non-COVID-19 pneumonia and definite evidence of pneumonia on chest CT.	Imaging read as positive for COVID-19	1: Average of 6 radiologists 2: Artificial intelligence model 3: Radiologists with AI assistance	Sensitivity: 1: 0.79 (0.64 to 0.89) 2: 0.95 (0.83 to 1.0) 3: 0.88 (0.74 to 0.95) Specificity: 1: 0.88 (0.78 to 0.94) 2: 0.96 (0.88 to 0.99) 3: 0.91 (0.82 to 0.96) AUROC: Not reported
Jin, et al., 2020[27]	Cases: COVID-19 cases from 3 hospitals in Wuhan, criteria for diagnosis not reported. Controls: Healthy controls from 3 hospitals in Wuhan and non-COVID-19 controls (clinical status not reported) from 2 international databases.	Positive classification for COVID-19 by AI system	AI (based on deep convolutional neural network)	Sensitivity: 0.95 (0.95 to 0.96) Specificity: 0.94 (0.93 to 0.95) AUROC: 0.98 (0.98 to 0.98)
Li, et al., 2020[29]	Cases: Positive for SARS-CoV-2 RT-PCR . Controls: Community acquired pneumonia or non-pneumonia patients.	Positive classification for COVID-19 by AI system	AI (COVID-19 detection neural network [COVnet])	Sensitivity: 0.90 (0.83 to 0.94) Specificity: 0.96 (0.93 to 0.98) AUROC: 0.96 (0.94 to 0.99)

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging Reader	Sensitivity (95%CI) Specificity (95%CI) AUROC (95%CI)
Wang, et al., 2020[33]	Cases: Positive SARS-CoV-2 RT-PCR. Controls: Non-COVID-19.	Positive for COVID-19 by deep learning algorithm	A: AI (deep learning algorithm) B: Radiologist 1 C: Radiologist 2	Sensitivity: A: 0.75 (CI Not reported) B: 0.71 (CI Not reported) C: 0.73 (CI Not reported) Specificity: A: 0.86 (CI Not reported) B: 0.51 (CI Not reported) C: 0.50 (CI Not reported) AUROC: A. 0.81 (0.71 to 0.84) B. Not reported C. Not reported
Ying, et al., 2020[40]	Cases: COVID-19 cases from 2 hospitals in Wuhan, diagnostic criteria not reported. Controls: Patients with bacterial pneumonia from 1 hospital.	Positive classification for COVID-19 vs.: A: Healthy controls B: Bacterial pneumonia	AI: (details relation extraction neural network)	Sensitivity: A: 0.93 (0.76 to 0.99) B: 0.96 (0.81 to 0.999) Specificity: A: 0.96 (0.79 to 0.999) B: 0.77 (0.58 to 0.90) AUROC: A. 0.99 (CI Not reported) B. 0.95 (CI Not reported)

Abbreviations: AI=artificial intelligence; AUROC=area under the receiver operating characteristic; CI=confidence interval; COVID-19=coronavirus disease 2019; CT=computed tomography; PCR=polymerase chain reaction; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2

3.2.2 Chest X-Ray

A retrospective cohort study from Italy of patients admitted for suspicion of COVID-19 (n=110, COVID-19 prevalence 67%) reported sensitivity of CXR of 0.64 and specificity of 0.78 and 0.86, based on two radiologists, for positive predictive values of 0.85 to 0.90 and negative predictive values of 0.51 to 0.53 (Table 3).[20] Sensitivity was higher using an artificial intelligence algorithm (0.80) with similar specificity (0.81), for a positive predictive value of 0.89 (95% CI 0.82 to 0.94) and negative predictive value of 0.66 (95% CI 0.57 to 0.75).

Two case series reported the sensitivity of CXR for diagnosis of COVID-19 (Table 3). One study conducted in the United States (n=636) of patients presenting to urgent care reported a sensitivity of 0.42 (95% CI 0.38 to 0.46) based on a CXR read as "mildly, moderately, or severely" abnormal.[34] The sensitivity decreased to 0.11 (95% CI 0.08 to 0.14) based on a CXR read as "moderately or severely" abnormal. A Chinese study (n=64) of patients with COVID-19 from four regional and tertiary hospitals reported a sensitivity of 0.69 (95% CI 0.56 to 0.80).[36]

Table 3. Diagnostic Accuracy of Chest X-Ray to Diagnose COVID-19

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging Reader	Sensitivity (95%CI) Specificity (95%CI) AUROC (95%CI)
Castiglioni, et al., 2020[20]	Suspected of COVID-19 and admitted to 1 hospital.	Imaging classified as positive for COVID-19	1: Artificial intelligence (10 convolutional neural networks) 2: Radiologist 1 3: Radiologist 2	Sensitivity: 1: 0.80 (0.72 to 0.86) 2: 0.64 (0.52 to 0.74) 3: 0.64 (0.52 to 0.74) Specificity: 1: 0.81 (0.73 to 0.87) 2: 0.78 (0.61 to 0.90) 3: 0.86 (0.71 to 0.95) AUROC: Not reported
Weinstock, et al., 2020[34]	Presented to urgent care center and positive SARS-CoV-2 RT-PCR.	A: CXR read as mildly, moderately, or severely abnormal B: CXR read as moderately or severely abnormal	Each CXR read by 1 of 11 radiologists	Sensitivity: A: 0.42 (0.38 to 0.46) B: 0.11 (0.08 to 0.14) Specificity: Not reported AUROC: Not reported
Wong, et al., 2020[36]	Positive SARS-CoV-2 RT-PCR in patients from 4 regional and tertiary hospitals.	Abnormality on CXR, otherwise not defined; severity score calculated using adapted and simplified Radiographic Assessment of Lung Edema score	2 radiologists reviewed by consensus, disagreements were resolved by a 3rd radiologist if needed	Sensitivity: 0.69 (0.56 to 0.80) Specificity: Not reported AUROC: Not reported

Abbreviations: AUROC=area under the receiver operator curve; CI=confidence interval; COVID-19=coronavirus disease 2019; CXR=chest x-ray; RT-PCR= reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2

3.2.3 Lung Ultrasound

One case series from France (n=107) compared the accuracy of LUS versus chest CT in persons with COVID-19 (Table 4).[18] Accuracy against RT-PCR as the reference standard was not evaluated. Using a scoring system to classify chest CT and LUS findings, the sensitivity of LUS for identifying patients with an abnormal CT (defined as at least mild CT findings for COVID-19) was 0.95 and the specificity was 0.83, with an AUROC of 0.93 (0.95 in multivariate regression). For identifying patients with at least moderate CT imaging findings, the sensitivity was 0.92 and the specificity was 0.74, with an AUROC of 0.89 (0.90 in multivariate regression). Confidence intervals (CIs) were not provided and could not be calculated for diagnostic accuracy estimates; other limitations were exclusion of 17 patients without a CT severity score and performance of some ultrasounds by trainees.

Table **4. Diagnostic Accuracy of Lung Ultrasound to Diagnose COVID-19**

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging Reader	Sensitivity (95%CI) Specificity (95%CI) AUROC (95%CI)
Benchoufi, et al., 2020[18]	Suspected or diagnosed COVID-19 and underwent chest CT and lung ultrasound.	US score >=1 (for abnormal) or >=6 (for moderate or severe), based on sum of severity scores (0=up to 3 observed B-lines to 3=consolidation foci) at 8 points of the chest wall.	CT: Not described US: 1 physician for US (expert read 68% of US and trainee 32%)	Sensitivity: 1: 0.95 (CI Not reported) 2: 0.92 (CI Not reported) Specificity: 1: 0.83 (CI Not reported) 2: 0.74 (CI Not reported) AUROC: 1: 0.93 (0.95 in multivariate logistic regression) 2: 0.89 (0.90 in multivariate logistic regression)

Abbreviations: AUROC=area under the receiver operating characteristic; CI=confidence interval; COVID-19=coronavirus disease 2019; CT=computed tomography; US=ultrasound

3.3 Key Question 3. In patients with suspected or confirmed COVID-19, not currently hospitalized and with mild symptoms, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used to support decision on hospital admission versus home discharge?

No study evaluated the effectiveness of chest imaging versus no imaging in patients with suspected or confirmed COVID-19 not yet hospitalized to support decisions on hospital admission versus home discharge on health outcomes. In addition, no study evaluated the association between chest imaging findings and health outcomes in such patients.

3.4 Key Question 4: In patients with suspected or confirmed COVID-19, not currently hospitalized and with moderate to severe symptoms, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used to support decision on regular ward admission versus intensive care unit admission?

See Key Question 3.

3.5 Key Question 5: In patients with suspected or confirmed COVID-19, currently hospitalized and with moderate or severe symptoms, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used to modify the therapeutic management?

No study evaluated the effectiveness of chest imaging versus no imaging on health outcomes in patients currently hospitalized with moderate or severe symptoms and suspected or confirmed COVID-19. Three retrospective cohort studies evaluated the association between chest CT findings and health outcomes in persons with COVID-19 in hospitalized patients (Table 5, Annex D).[21,23,41] All of the studies had methodological limitations. One study (n=249, 8.8% admitted to an ICU) of patients diagnosed with COVID-19 in China reported a very imprecise estimate for the association between presence of radiological findings on CT and likelihood of ICU admission (unadjusted odds ratio [OR] 4.46, 95% CI 0.62 to 31.9); presence of radiological findings was not included in a multivariate model for prediction of intensive care unit admission.[21] In addition, “radiological finding” was not defined. Another study conducted in China (n=27, 37.0% mortality) found a CT severity score of 24.5 or greater associated with sensitivity of 0.86, specificity of 0.84, and AUROC of 0.90 (95% CI 0.87 to 0.93) for predicting mortality in hospitalized patients with COVID-19.[41] The CT severity score was based on the sum of radiologic scores in six regions, based on the degree of attenuation and the amount of lung parenchyma affected (range 0 to 72). The study was limited by its small sample size; in addition, CIs were not reported for sensitivity or specificity and the CT severity score threshold was not pre-specified. The third study evaluated COVID-19 patients in Italy (n=236, 46% ICU admission or death).[23] It found that a model that included estimates of the amount of well-aerated lung on

chest CT in addition to clinical factors improved specificity compared with a model with clinical factors alone (specificity 0.80 to 0.81 with amount of well-aerated lung vs. 0.73 without). However, adding measures of the amount of well-aerated lung had little effect on discrimination (AUROC 0.83, 95% CI 0.78 to 0.88 based on clinical factors alone and 0.86, 95% CI 0.80 to 0.90 when the proportion of well-aerated lung was added). Limitations included no control for confounders and unclear severity of symptoms at baseline; in addition, the thresholds for the amount of well-aerated lung used to indicate high mortality risk were not pre-specified.

Table **5. Association Between Chest Computed Tomography Findings and Health Outcomes in Persons With COVID-19**

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging Reader	Results
Chen, et al., 2020[21,22]	Diagnosed with COVID-19.	Radiological lesion (not otherwise defined)	Not reported	Risk of ICU admission, radiological lesion vs. no radiological lesion: Unadjusted OR 4.46 (95% CI 0.62 to 31.9); not included in multivariate model
Colombi, et al., 2020[23]	SARS-CoV-2 RT-PCR positive, with imaging findings on chest CT.	1: Clinical model 2: Model with % lung well-aerated assessed visually and clinical parameters; threshold not pre-specified 3: Model with % lung well-aerated assessed with software and clinical parameters; threshold not pre-specified 4: Model with clinical parameters, well-aerated lung volume <2.9 L and adipose tissue are >262 cm2; threshold not pre-specified	1: Not applicable 2: 2 radiologists with 5 and 14 years of experience 3: Software to calculate CT parameters 4: Software to calculate CT parameters	Sensitivity 1: 0.75 (0.66 to 0.82) 2: 0.72 (0.63 to 0.80) 3: 0.75 (0.66 to 0.83) 4: 0.75 (0.66 to 0.83) Specificity 1: 0.73 (0.65 to 0.81) 2: 0.81 (0.73 to 0.88) 3: 0.80 (0.72 to 0.86) 4: 0.81 (0.73 to 0.88) Positive predictive value 1: 0.70 (0.61 to 0.78) 2: 0.76 (0.68 to 0.82) 3: 0.75 (0.68 to 0.81) 4: 0.77 (0.69 to 0.83) Negative predictive value 1: 0.78 (0.72 to 0.83) 2: 0.78 (0.73 to 0.83) 3: 0.80 (0.73 to 0.85) 4: 0.79 (0.74 to 0.84) AUROC 1: 0.83 (0.78 to 0.88) 2: 0.86 (0.81 to 0.90) 3: 0.86 (0.80 to 0.90) 4: 0.86 (0.81 to 0.90)
Yuan, et al., 2020[41]	Diagnosed with COVID-19 (SARS-CoV-2 RT-PCR positive) and discharged with recovered symptoms or died in hospital.	CT score >24.5; sum of radiologic score (1=normal attenuation, 2=ground glass, 3=consolidation) times lung parenchyma distribution score (1=<25% abnormality, 2=25-50%, 3=50-75%, 4=over 75%) for 6 lung zones (range 0 to 72)	2 radiologists, discrepancies resolved by consensus	Sensitivity: 0.96 (CI NR) Specificity: 0.84 (CI NR) AUROC: 0.90 (0.87 to 0.93)

Abbreviations: AUROC=area under the receiver operating characteristic; CI=confidence interval; COVID-19=coronavirus disease 2019; CT=computed tomography; ICU=intensive care unit; NR=not reported; OR=odds ratio; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2

3.6 Key Question 6. In patients with suspected or confirmed COVID-19 and clinical deterioration and/or suspicion of pulmonary thromboembolism, should imaging (including CT pulmonary angiography, lower extremity venous ultrasound) vs. no imaging be used to diagnose pulmonary thromboembolism?

No study evaluated the diagnostic accuracy of imaging (with or without measurement of d-dimer) for diagnosis of pulmonary thromboembolism in patients with suspected or confirmed COVID-19. Two studies (n=100 and 106) reported the prevalence of pulmonary embolus in imaging series of hospitalized patients in France who underwent CT pulmonary angiography (Table 6, Annex D).[26,28] The prevalence of pulmonary thromboembolism in the two studies was 23% and 30%. Patients with pulmonary embolus were more likely to be in the ICU and to be undergoing mechanical ventilation compared with those without pulmonary embolus. Methodological limitations included unclear selection of patients for CT pulmonary angiography and limited clinical information about the patients. In addition, the study that reported a pulmonary embolus prevalence of 23% was not restricted to patients with COVID-19 confirmed by RT-PCR; rather, it also included patients with suspected COVID-19 based on fever and/or acute respiratory symptoms and exposure to a patient with confirmed SARS-CoV-2 infection.[26]

Table **6. Prevalence of Pulmonary Embolus in Imaging Series**

Author, Year	Eligibility Criteria	Definition of Positive Test	Imaging reader	Results
Grillet, et al., 2020[26]	SARS-CoV-2 RT-PCR positive (n=97) or positive CT and negative RT-PCR (n=9), underwent contrast CT (performed when clinical features of severe disease were present).	CT with contrast read as positive for pulmonary embolus	2 radiologists with 6 and 11 years of experience	23.0% (23/100)
Leonard-Lorant, et al., 2020[28]	Underwent CT pulmonary angiography examination for suspicion or follow-up of SARS-CoV-2 infection.	CT with contrast read as positive for pulmonary embolus	1 radiologist	30.2% (32/106)

Abbreviations: COVID-19=coronavirus disease 2019; CT=computed tomography; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2

3.7 Key Question 7. In patients with COVID-19 whose symptoms resolved, should chest imaging (including CXR, CT scan, LUS) be added to vs. not added to laboratory criteria to support decision on discharge home vs. no discharge home?

No study evaluated the effects of chest imaging versus no chest imaging to support decision on discharge home, or the association between chest imaging findings at discharge and health outcomes following discharge.

4 Conclusions

4.1 Summary of key findings

This rapid review summarizes the evidence on chest imaging in persons suspected of or diagnosed with COVID-19. The GRADE table summarizing key findings is shown in Table 7. No study evaluated diagnostic accuracy of chest imaging in asymptomatic subjects possibly affected with COVID-19. In symptomatic patients, chest CT was associated with high sensitivity but low specificity in most studies, resulting in weak positive likelihood ratios but stronger negative likelihood ratios. These findings indicate that when used for diagnosis in symptomatic patients, negative CT results are more useful for diagnosis than positive results. For example, based on the largest cohort study,[15] the positive likelihood ratio was 1.28 and the negative likelihood ratio was 0.16.

Based on these likelihood ratios, in a symptomatic population with a pre-test probability for COVID-19 of 10%, a positive CT would increase the post-test probability slightly to 12.5%, but a negative CT would decrease the post-test probability to 1.7%. However, findings regarding diagnostic accuracy must be interpreted with caution. Studies had methodological limitations, including retrospective design, imperfect reference standard (single RT-PCR), no explicit criteria for classifying a CT as positive, and limited information regarding clinical presentation. In addition, the studies evaluated cohorts with high COVID-19 prevalence or used a case-control design, which could overestimate diagnostic accuracy. Diagnostic accuracy estimates from high COVID-19 prevalence, symptomatic populations are likely to overestimate diagnostic accuracy compared with lower prevalence or asymptomatic populations, though the degree of overestimation is unknown.

Evidence suggested that use of artificial intelligence might improve the accuracy (particularly specificity) of chest CT for diagnosis of COVID-19, but additional validation in well-conducted studies in clinical settings are needed. Evidence on the accuracy of CXR was very limited, but suggests sensitivity that is lower than CT while specificity may be higher than CT, depending on what CT study is taken into consideration. Evidence on LUS was limited to one study that evaluated diagnostic accuracy in comparison with chest CT.

No study evaluated the effects of chest imaging on health outcomes. Studies on the utility of chest imaging for predicting health outcomes were limited and inconclusive. No study evaluated the diagnostic accuracy of chest imaging for diagnosis of pulmonary embolus in patients with COVID-19. One study reported a prevalence of pulmonary embolus of 30% and another study reported a prevalence of 23%, but included patients with diagnosed as well as suspected COVID-19. No study evaluated chest imaging in COVID-19 patients to inform decisions regarding discharge.

Table **7. Summary of Evidence – GRADE**

Outcome	Number and type of studies	Risk of bias	Indirectness	Imprecision	Inconsistency	Other considerations	Summary findings	Certainty of evidence
KQ 1 (diagnostic accuracy in asymptomatic patients) No evidence								
KQ 2 (diagnostic accuracy in symptomatic patients)								
CT: Sensitivity*	Cohort: 5 (N=1741) Case-control: 2 (289 cases, 243 controls) Case series: 6 (317 cases)	Serious limitations	Serious indirectness^	Precise	Consistent	None	Cohort studies: 0.87 to 0.97 Case control studies: 0.67 and 1.0 Case series: 0.81 to 0.97	Low
CT: Specificity*	Cohort studies: 5 (N=1741) Case-control studies: 2 (289 cases, 243 controls)	Serious limitations	Serious indirectness^	Precise	Serious inconsistency	None	Cohort studies: 0.24 to 0.68; one outlier with specificity 0.95 Case control studies: 0.07 to 1.0	Very low
CXR: Sensitivity	Cohort study: 1 (n=110) Case series: 2 (700 cases)	Very serious limitations	Serious indirectness^	Precise	Consistent	None	Cohort study: 0.64 Case series: 0.42 and 0.69	Very low
CXR: Specificity	Cohort study: 1 (n=110)	Very serious limitations	Serious indirectness^	Precise	Unable to determine	None	Cohort study: 0.78 and 0.86 (two radiologists)	Very low
LUS: Sensitivity	Cohort study: 1 (n=107)	Very serious limitations	Serious indirectness^	Serious imprecision	Unable to determine	None	Cohort study: 0.92	Very low
LUS: Specificity	Cohort study: 1 (n=107)	Very serious limitations	Serious indirectness^	Serious imprecision	Unable to determine	None	Cohort study: 0.74	Very low
KQ 3 (effectiveness/prediction in mildly symptomatic non-hospitalized patients) No evidence								
KQ 4 (effectiveness/prediction in moderately or severely symptomatic non-hospitalized patients) No evidence								

Outcome	Number and type of studies	Risk of bias	Indirectness	Imprecision	Inconsistency	Other considerations	Summary findings	Certainty of evidence
KQ 5 (effectiveness/prediction in hospitalized patients)								
Mortality, ICU admission	Cohort studies: 3 (N=512)	Very serious limitations	Serious indirectness^	Serious imprecision	Unable to determine	None	The amount of well-aerated lung did not improve discrimination for mortality or ICU admission versus clinical factors alone, based on 1 study. Otherwise evidence was too limited to determine the association between imaging findings and health outcomes	Very low
KQ 6 (pulmonary embolus)								
Prevalence	Imaging series: 2 (N=206)	Very serious limitations	Serious indirectness^	Serious imprecision	Consistent	None	Prevalence of pulmonary embolus 30% and 23%	Very low
KQ 7 (effectiveness/prediction to inform discharge decisions) No evidence								

Abbreviations: CT=computed tomography; CXR=chest x-ray; ICU=intensive care unit; KQ=Key Question; LUS=lung ultrasound
*Restricted to non-artificial intelligence studies
^Serious indirectness due to limited information about clinical presentation including symptom severity

4.2 Applicability

Several factors impact the applicability of findings. Cohort studies on diagnostic accuracy were conducted in high prevalence and symptomatic populations, which could reduce applicability of estimates to lower prevalence or asymptomatic populations. Information regarding the clinical characteristics of imaged patients was also limited, making it difficult to determine in whom results would apply. No study evaluated patients specifically with mild symptoms and few studies classified symptom severity at the time of imaging. Details about CT imaging techniques were limited from some studies, and some studies reported reconstruction slice thickness of 1.25 mm or greater, which could reduce diagnostic accuracy compared with thinner slice thickness. However, it was not possible to determine how imaging parameters impacted diagnostic accuracy. Most studies were conducted in China, which could limit applicability of findings to other countries.

4.3 Limitations in the evidence

The evidence had important limitations. For diagnostic accuracy, limitations included use of a case-control design, resulting in potential spectrum bias. A number of studies utilized a case series design, which provide incomplete information about diagnostic accuracy because only sensitivity can be estimated. Cohort studies evaluated symptomatic patients with a high proportion of COVID-19, which could limit applicability to lower prevalence or asymptomatic patients. Studies on diagnostic accuracy had methodological limitations related to methods used to select patients, description of baseline characteristics, application of clear criteria to interpret imaging tests as positive for COVID-19, and potential misclassification due to use of an imperfect reference standard (e.g., single RT-PCR). There was very limited evidence on the diagnostic accuracy of CXR or LUS and the utility of chest imaging for predicting clinical outcomes. No study evaluated effects of chest imaging on therapy decision-making or on health outcomes.

4.4 Limitations in the review

- Did not register protocol in PROSPERO
- Searched only one grey literature database
- Inclusion of non-peer reviewed studies
- Did not include descriptive or exploratory studies on imaging findings (e.g., reported various imaging findings in patients with COVID-19, but did not report diagnostic accuracy for imaging interpreted as positive for COVID-19)
- Did not translate non-English language case series, given the low quality of such evidence
- Summarized findings narratively

4.5 Research gaps

- Diagnostic accuracy studies of CT, ideally prospective, that evaluate clearly described patient cohorts (including mild and more severely symptomatic patients), apply well-defined criteria

for classifying an imaging test as positive for COVID-19, and utilize more reliable reference standards (e.g., serial RT-PCR assays or RT-PCR assays in combination with clinical follow-up).

- Diagnostic accuracy studies of CXR and LUS that adhere to the principles described above.
- Studies on the effectiveness of chest imaging versus no imaging on health outcomes for patient not currently hospitalized as well as those hospitalized.
- Studies on the association between chest imaging findings and health outcomes that control for potential confounders
- Studies on the accuracy of chest imaging for diagnosing pulmonary embolus in patients with COVID-19, the prevalence of pulmonary thromboembolism in well-defined (ideally prospective) cohorts of patients with COVID-19, risk factors for pulmonary thromboembolism, and effects of chest imaging for pulmonary thromboembolism on clinical management and health outcomes.

References

1. World Health Organization. Novel coronavirus - China. 2020. https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/. Accessed 30 March 2020.
2. Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020 Feb 22;395(10224):565-74. doi: 10.1016/s0140-6736(20)30251-8. PMID: 32007145.
3. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020 Mar;579(7798):270-3. doi: 10.1038/s41586-020-2012-7. PMID: 32015507.
4. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020 Feb 20;382(8):727-33. doi: 10.1056/NEJMoa2001017. PMID: 31978945.
5. World Health Organization. Novel coronavirus (2019-nCov) - Situation Report 22. 2020. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf?sfvrsn=fb6d49b1_2. Accessed 30 March 2020.
6. World Health Organization. Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). 2020. https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov). Accessed 30 March 2020.
7. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. 2020. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020. Accessed 30 March 2020.
8. World Health Organization, Alliance for Health Policy and Systems Research. Rapid reviews to strengthen health policy and systems: a practical guide. 2020. https://www.who.int/alliance-hpsr/resources/publications/rapid-review-guide/en/. Accessed May 1, 2020.
9. World Health Organization. Database of publications on coronavirus disease (COVID-19). 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov. Accessed 30 March 2020.
10. Cold Spring Harbor Laboratory. medRxiv: The Preprint Server of Health Sciences. https://www.medrxiv.org/. Accessed 30 March 2020.
11. MEDCALC. 2020. https://www.medcalc.org/calc/diagnostic_test.php. Accessed May 1, 2020.
12. University of Bristol Centre for Research Synthesis and Decision Analysis. The ROBINS-E tool (Risk Of Bias In Non-randomized Studies - of Exposures). http://www.bristol.ac.uk/population-health-sciences/centres/cresyda/barr/riskofbias/robins-e/. Accessed 30 March 2020.
13. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 30 March 2020.

14. Ai J, Gong J, Xing L, et al. Analysis of factors associated early diagnosis in coronavirus disease 2019 (COVID-19). medRxiv. 2020:2020.04.09.20059352. doi: 10.1101/2020.04.09.20059352.
15. Ai T, Yang Z, Hou H, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology. 2020 Feb 26:200642. doi: 10.1148/radiol.2020200642. PMID: 32101510.
16. Bai HX, Hsieh B, Xiong Z, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. Radiology. 2020 Mar 10:200823. doi: 10.1148/radiol.2020200823. PMID: 32155105.
17. Bai HX, Wang R, Xiong Z, et al. AI augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other etiology on chest CT. Radiology. 2020 Apr 27:201491. doi: 10.1148/radiol.2020201491. PMID: 32339081.
18. Benchoufi M, Bokobza J, Anthony Chauvin A, et al. Lung injury in patients with or suspected COVID-19 : a comparison between lung ultrasound and chest CT-scanner severity assessments, an observational study. medRxiv. 2020:2020.04.24.20069633. doi: 10.1101/2020.04.24.20069633.
19. Caruso D, Zerunian M, Polici M, et al. Chest CT Features of COVID-19 in Rome, Italy. Radiology. 2020 Apr 3:201237. doi: 10.1148/radiol.2020201237. PMID: 32243238.
20. Castiglioni I, Ippolito D, Interlenghi M, et al. Artificial intelligence applied on chest X-ray can aid in the diagnosis of COVID-19 infection: a first experience from Lombardy, Italy. medRxiv. 2020:2020.04.08.20040907. doi: 10.1101/2020.04.08.20040907.
21. Chen J, Qi T, Liu L, et al. Clinical progression of patients with COVID-19 in Shanghai, China. J Infect. 2020 Mar 19 doi: 10.1016/j.jinf.2020.03.004. PMID: 32171869.
22. Chen X, Tang Y, Mo Y, et al. A diagnostic model for coronavirus disease 2019 (COVID-19) based on radiological semantic and clinical features: a multi-center study. Eur Radiol. 2020 Apr 16 doi: 10.1007/s00330-020-06829-2. PMID: 32300971.
23. Colombi D, Bodini FC, Petrini M, et al. Well-aerated Lung on Admitting Chest CT to Predict Adverse Outcome in COVID-19 Pneumonia. Radiology. 2020 Apr 17:201433. doi: 10.1148/radiol.2020201433. PMID: 32301647.
24. Dangis A, Gieraerts C, Bruecker YD, et al. Accuracy and reproducibility of low-dose submillisievert chest CT for the diagnosis of COVID-19. Radiology: Cardiothoracic Imaging. 2020;2(2):e200196. doi: 10.1148/ryct.2020200196.
25. Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. Radiology. 2020 Feb 19:200432. doi: 10.1148/radiol.2020200432. PMID: 32073353.
26. Grillet F, Behr J, Calame P, et al. Acute pulmonary embolism associated with COVID-19 pneumonia detected by pulmonary CT angiography. Radiology. 2020 Apr 23:201544. doi: 10.1148/radiol.2020201544. PMID: 32324103.
27. Jin C, Chen W, Cao Y, et al. Development and evaluation of an AI system for COVID-19 diagnosis. medRxiv. 2020:2020.03.20.20039834. doi: 10.1101/2020.03.20.20039834.
28. Leonard-Lorant I, Delabranche X, Severac F, et al. Acute pulmonary embolism in COVID-19 patients on CT angiography and relationship to D-Dimer levels. Radiology. 2020 Apr 23:201561. doi: 10.1148/radiol.2020201561. PMID: 32324102.
29. Li L, Qin L, Xu Z, et al. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology. 2020 Mar 19:200905. doi: 10.1148/radiol.2020200905. PMID: 32191588.

30. Li Y, Xia L. Coronavirus disease 2019 (COVID-19): role of chest CT in diagnosis and management. AJR Am J Roentgenol. 2020 Mar 4:1-7. doi: 10.2214/ajr.20.22954. PMID: 32130038.
31. Long C, Xu H, Shen Q, et al. Diagnosis of the coronavirus disease (COVID-19): rRT-PCR or CT? Eur J Radiol. 2020 Mar 25;126:108961. doi: 10.1016/j.ejrad.2020.108961. PMID: 32229322.
32. Prokop M, van Everdingen W, van Rees Vellinga T, et al. CO-RADS - A categorical CT assessment scheme for patients with suspected COVID-19: definition and evaluation. Radiology. 2020 Apr 27:201473. doi: 10.1148/radiol.2020201473. PMID: 32339082.
33. Wang S, Kang B, Ma J, et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). medRxiv. 2020:2020.02.14.20023028. doi: 10.1101/2020.02.14.20023028.
34. Weinstock MB, Echenique A, Russell JW, et al. Chest x-ray findings in 636 ambulatory patients with COVID-19 presenting to an urgent care center: a normal chest x-ray is no guarantee. J Urgent Care Med. 2020;14(7):13-8.
35. Wen Z, Chi Y, Zhang L, et al. Coronavirus disease 2019: initial detection on chest CT in a retrospective multicenter study of 103 Chinese subjects. Radiology: Cardiothoracic Imaging. 2020;2(2):e200092. doi: 10.1148/ryct.2020200092.
36. Wong HYF, Lam HYS, Fong AH, et al. Frequency and distribution of chest radiographic findings in COVID-19 positive patients. Radiology. 2019 Mar 27:201160. doi: 10.1148/radiol.2020201160. PMID: 32216717.
37. Xu YH, Dong JH, An WM, et al. Clinical and computed tomographic imaging features of novel coronavirus pneumonia caused by SARS-CoV-2. J Infect. 2020 Apr;80(4):394-400. doi: 10.1016/j.jinf.2020.02.017. PMID: 32109443.
38. Xu ZF, Wu WX, Jin YB, et al. Key points of clinical and CT imaging features of 2019 novel coronavirus (2019-nCoV) imported pneumonia based on 21 cases analysis. medRxiv. 2020:2020.03.03.20030775. doi: 10.1101/2020.03.03.20030775.
39. Yang X, Wang Z, Liu X, et al. Screening for 274 suspected cases of novel coronavirus pneumonia. Chinese Journal of Infectious Diseases. 2020;38:Online pre-publishing. doi: 10.3760/cma.j.cn311365-20200218-00081.
40. Ying S, Zheng S, Li L, et al. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. medRxiv. 2020:2020.02.23.20026930. doi: 10.1101/2020.02.23.20026930.
41. Yuan M, Yin W, Tao Z, et al. Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China. PLoS One. 2020;15(3):e0230548. doi: 10.1371/journal.pone.0230548. PMID: 32191764.

Acknowledgements

The authors thank Xuan Yu at the WHO Collaborating Centre for Guideline Implementation and Knowledge Translation, Lanzhou University, China for assistance searching the Chinese language database and translation.

Appendix 1

1.1 Search strategies

PubMed MEDLINE
((((((((("COVID-19" [Supplementary Concept]) OR "COVID-19 diagnostic testing" [Supplementary Concept])) OR "severe acute respiratory syndrome coronavirus 2" [Supplementary Concept]) OR (("COVID 19"[Title/Abstract]) OR "COVID 19"[Other Term])) OR (("nCOV" OR "COV 2" OR "coronavirus 2"))) OR ("novel coronavirus" AND ("2019/01/01"[PDat] : "2020/12/31"[PDat]))))) AND ((((Diagnostic imaging[MeSH Subheading]) OR (image OR imaging)) OR ((("xray" OR "x ray" OR scan OR scans OR scanning OR ultrasound OR "CT" OR tomography OR tomographic OR “d dimer” OR “fibrin degradation”))))

((((((((((("COVID-19"[Supplementary Concept]) OR "COVID-19 diagnostic testing"[Supplementary Concept])) OR "severe acute respiratory syndrome coronavirus 2"[Supplementary Concept]) OR (("COVID 19"[Title/Abstract]) OR "COVID 19"[Other Term])) OR (("nCOV" OR "COV 2" OR "coronavirus 2"))) OR ("novel coronavirus" AND ("2019/01/01"[PDat] : "2020/12/31"[PDat])))))))) AND ((pulmonary embolism OR thromboembolism OR thromboembolic))

Elsevier Embase
('covid 19' OR 'severe acute respiratory syndrome coronavirus 2' OR 'ncov' OR 'cov 2' OR 'cov2' OR 'coronavirus 2' OR ('novel coronavirus' AND (2019:py OR 2020:py))) AND ('diagnostic imaging' OR image OR imaging OR 'xray' OR 'x ray' OR 'scan' OR 'scans' OR 'scanning' OR 'ultrasound' OR 'ct' OR 'tomography' OR 'tomographic' OR 'd dimer' OR 'fibrin degradation') [embase]/lim NOT ([embase]/lim AND [medline]/lim)

('covid 19' OR 'severe acute respiratory syndrome coronavirus 2' OR 'ncov' OR 'cov 2' OR 'cov2' OR 'coronavirus 2' OR ('novel coronavirus' AND (2019:py OR 2020:py))) AND (pulmonary embolism OR thromboembolism OR thromboembolic) [embase]/lim NOT ([embase]/lim AND [medline]/lim)

Appendix 2

2. 1 Inclusion Criteria (PICOS)

COVID-19 probability[1]	Disease severity	Risk factors for disease progression[2]	Resource constraints[3]
1. Suspected 2. Probable 3. Confirmed 4. Contact	1. Mild 2. Moderate to severe 3. Disease progression	1. YES 2. NO	1. YES 2. NO

[1]*WHO case classification*

[2]*Age group (children, adults, elderly >65), cardiovascular disease, cancer, diabetes, chronic pulmonary disease, hypertension, immunosuppressive conditions (HIV/AIDS, others), tuberculosis, obesity, other comorbidities. Particular considerations for pregnant women.*

[3]*Health workforce/qualified staff; personal protection equipment and other infection, prevention, and control measures; laboratory testing; hospital beds; ventilators; imaging equipment/devices.*

KQ 1: In asymptomatic contacts[1] of patients with COVID-19, and in contexts where laboratory testing (RT-PCR) is not available/results are delayed[1]/results are initially negative, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used for the diagnostic workup of COVID-19?

	Inclusion Criteria	Comments
Setting	• Need to rule in/rule out COVID19 in asymptomatic contacts, where laboratory testing (RT PCR) is not available/results are delayed[2]/results are initially negative	Need for early recognition
P	• Asymptomatic contacts of patients with COVID-19	
I	• Chest imaging (including CXR, CT scan, LUS)	
C	• Alternative chest imaging (including CXR, CT scan, LUS) • No chest imaging	

[1] *Contact: a person who experienced any one of the following exposures during the 2 days before and the 14 days after the onset of symptoms of a probable or confirmed case: (1) Face-to-face contact with a probable or confirmed case within 1 meter and for more than 15 minutes; (2) Direct physical contact with a probable or confirmed case; (3) Direct care for a patient with probable or confirmed COVID-19 disease without using proper personal protective equipment; OR (4) Other situations as indicated by local risk assessments (for confirmed asymptomatic cases, the period of contact is measured as the 2 days before through the 14 days after the date on which the sample was taken which led to confirmation).*

[2]*A delay of more than one day to get confirmatory results of SARS-CoV-2 infection*

O	• Diagnostic accuracy (sensitivity, specificity)	Additional clinical outcomes of interest: • Length of ED stay • Mortality • Need for and length of hospital stay • Need for and length of ICU stay • Need for and length of respiratory support • Complications of imaging Additional health systems outcomes of interest: • COVID-19 transmission to health workers • Availability of care • Access to care • Quality of care

KQ 2: In symptomatic patients with suspected COVID-19,[3] and in contexts where laboratory testing (RT-PCR) is not available/results are delayed[4]/results are initially negative, should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used for the diagnostic workup of COVID-19?

	Inclusion Criteria	Comments
Setting	• Need to rule in/rule out COVID19 infection in a symptomatic patient, where laboratory testing (RT PCR) is not available/results are delayed/results are initially negative	Need for early recognition
P	• Symptomatic patients with suspected COVID-19	
I	• Chest imaging (including CXR, CT scan, LUS)	
C	• Alternative chest imaging (including CXR, CT scan, LUS) • No chest imaging	
O	• Diagnostic accuracy (sensitivity, specificity)	Additional clinical outcomes of interest: • Length of ED stay • Mortality • Treatment failure • Need for and length of hospital stay • Need for and length of ICU stay • Need for and length of respiratory support • Complications of imaging Additional health systems outcomes of interest: • COVID-19 transmission to health workers • Availability of care • Access to care • Quality of care

[3]*Suspect case is (A) a patient with acute respiratory illness (fever and at least one sign/symptom of respiratory disease, e.g., cough, shortness of breath), AND a history of travel to or residence in a location reporting community transmission of COVID-19 disease during the 14 days prior to symptom onset, or (B) A patient with any acute respiratory illness AND having been in contact with a confirmed or probable COVID-19 case in the last 14 days prior to symptom onset; or (C) A patient with severe acute respiratory illness (fever and at least one sign/symptom of respiratory disease, e.g., cough, shortness of breath; AND requiring hospitalization) AND in the absence of an alternative diagnosis that fully explains the clinical presentation.*

[4]*A delay of more than one day to get confirmatory results of SARS-CoV-2 infection*

KQ 3: In patients with suspected or confirmed COVID-19, not currently hospitalized and with mild symptoms[5], should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used to support decision on hospital admission versus home discharge?

	Inclusion Criteria	Comments
Setting	Patient with suspected or confirmed COVID-19 and mild symptoms presenting to the healthcare system (e.g. emergency department); context of a decision on hospital admission versus home discharge	
P	• Patients with confirmed[6] COVID-19 and mild to moderate symptoms not currently hospitalized	
I	• Chest imaging (including CXR, CT scan, LUS) to guide decision for hospital admission vs. home discharge	
C	• Alternative chest imaging (including CXR, CT scan, LUS) • No chest imaging	
O	Clinical outcomes of interest: • Mortality • Need for and length of hospital stay • Need for and length of ICU stay • Need for and length of respiratory support • Complications of imaging Health systems outcomes of interest: • COVID-19 transmission to health workers • Availability of care • Access to care • Quality of care Measures of association between imaging findings and health outcomes described above (secondary outcome)	

[5] *Mild symptoms: respiratory symptoms with no signs of pneumonia*

[6] *A confirmed case is a person with laboratory confirmation of COVID-19 infection, irrespective of clinical signs and symptoms*.

KQ 4: In patients with suspected or confirmed COVID-19, not currently hospitalized and with moderate to severe symptoms,[7] should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used to support decision on regular ward admission versus intensive care unit admission?

	Inclusion Criteria	Comments
Setting	Context of a decision to choose between admission to regular ward vs. ICU	Could apply to patients already hospitalized; decision whether to keep on the regular ward or transfer to the intensive care unit
P	• Patients with suspected or confirmed COVID-19 and moderate to severe symptoms	
I	• Chest imaging (including CXR, CT scan, LUS) to guide decision for admission to regular ward versus intensive care unit	
C	• Alternative chest imaging (including CXR, CT scan, LUS) • No chest imaging	
O	Clinical outcomes of interest: • Mortality • Need for and length of hospital stay • Need for and length of ICU stay • Need for and length of respiratory support • Complications of imaging Health systems outcomes of interest: • COVID-19 transmission to health workers • Availability of care • Access to care • Quality of care Measures of association between imaging findings and health outcomes described above (secondary outcome)	

[7] *Moderate symptoms: signs of pneumonia or in children fast breathing and chest indrawing but no need of oxygen or no signs of severe pneumonia present. Severe symptoms: signs of pneumonia with any of; respiratory distress, hypoxaemia, high respiratory rate for age, and in children presence of danger signs.*

KQ 5: In patients with suspected or confirmed COVID-19, currently hospitalized and with moderate or severe symptoms[8], should chest imaging (including CXR, CT scan, LUS) vs. no chest imaging be used to modify the therapeutic management?

	Inclusion Criteria	Comments
Setting	ICU admission; context of a decision to choose whether to escalate respiratory support	
P	• Patients with suspected or confirmed COVID-19 and severe symptoms	
I	• Chest imaging (including CXR, CT scan, LUS)to guide decision for escalating respiratory support vs. not	
C	• Alternative chest imaging (including CXR, CT scan, LUS) • No chest imaging	
O	Clinical outcomes of interest: • Mortality • Need for and length of hospital stay • Need for and length of ICU stay • Need for and length of respiratory support • Complications of imaging Health systems outcomes of interest: • COVID-19 transmission to health workers • Availability of care • Access to care • Quality of care Measures of association between imaging findings and health outcomes described above (secondary outcome)	

[8] *Severe symptoms: signs of pneumonia with any of—respiratory distress, hypoxaemia, high respiratory rate for age, and in children presence of danger signs. Critical illness: acute respiratory distress syndrome (ARDS), sepsis.*

KQ 6: In patients with suspected or confirmed COVID-19 and clinical deterioration[9] and/or suspicion of pulmonary embolism, should imaging (including CT pulmonary angiography, lower extremity venous ultrasound) vs. no imaging be used to diagnose pulmonary embolism?

	Inclusion Criteria	Comments
Setting	Need to rule in or rule out diagnosis of pulmonary embolism	
P	• Patients with suspected or confirmed COVID-19 and clinical deterioration and suspicion of pulmonary embolism	
I	• Imaging (including CT pulmonary angiography, low extremity venous ultrasound)	With/without d- Dimer testing
C	• Alternative imaging (including CT scan and lower extremity venous ultrasound) • Clinical assessment	
O	• Diagnostic accuracy (sensitivity, specificity) • Prevalence (secondary outcome)	Additional clinical outcomes of interest: • Mortality • Need for and length of hospital stay • Need for and length of ICU stay • Need for and length of respiratory support • Complications of imaging Additional health systems outcomes of interest: • COVID-19 transmission to health workers • availability of care • Access to care • Quality of care

[9] *Abrupt worsening of hyoxia, edema or erythema of an extremity, unexplained shortness of breath out of proportion to oxygen saturation, increased tachycardia, or for mechanically ventilated patients:increased dead space fraction out of proportion to change in lung compliance.*

KQ 7: In patients with COVID-19 whose symptoms resolved, should chest imaging (including CXR, CT scan, LUS) be added to vs. not added to laboratory criteria to support decision on discharge home vs. no discharge home?

	Inclusion Criteria	Comments
Setting	Isolation setting; context of a decision to choose between discharge home vs. no discharge home	In some contexts laboratory testing (RT PCR) may not be available
P	• Patients with COVID-19 whose symptoms resolved	
I	• Chest imaging (different modalities) to guide decision for discharge home vs. no discharge home added	Standard of care is: Laboratory criteria (2 negative RT PCR tests) along with clinical criteria (symptoms resolution for 14 days)
C	• Chest imaging not added	
O	Clinical outcomes of interest: • Mortality • Need for and length of hospital stay • Need for and length of ICU stay • Need for and length of respiratory support • Complications of imaging Health systems outcomes of interest: • COVID-19 transmission to health workers • Availability of care • Access to care • Quality of care Measures of association between imaging findings and health outcomes described above (secondary outcome)	

Appendix 3

3.1 Included Studies by Key Question

3.1.1 Key Question 1
None

3.1.2 Key Question 2
Chest CT – Cohort Studies

1. Ai T, Yang Z, Hou H, et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology. 2020:200642.
2. Caruso D, Zerunian M, Polici M, et al. Chest CT Features of COVID-19 in Rome, Italy. Radiology. 2020:201237.
3. Dangis A, Gieraerts C, Bruecker YD, et al. Accuracy and reproducibility of low-dose submillisievert chest CT for the diagnosis of COVID-19. Radiology: Cardiothoracic Imaging. 2020;2(2):e200196.
4. Prokop M, van Everdingen W, van Rees Vellinga T, et al. CO-RADS - A categorical CT assessment scheme for patients with suspected COVID-19: definition and evaluation. Radiology. 2020:201473.
5. Wen Z, Chi Y, Zhang L, et al. Coronavirus Disease 2019: Initial Detection on Chest CT in a Retrospective Multicenter Study of 103 Chinese Subjects. Radiology: Cardiothoracic Imaging. 2020;2(2):e200092.
6. Yang X, Wang Z, Liu X, et al. Screening for 274 suspected cases of novel coronavirus pneumonia. 2020.

Chest CT – Case-Control Studies

1. Bai HX, Hsieh B, Xiong Z, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. Radiology. 2020:200823.
2. Bai HX, Wang R, Xiong Z, et al. AI Augmentation of Radiologist Performance in Distinguishing COVID-19 from Pneumonia of Other Etiology on Chest CT. Radiology. 2020:201491.
3. Chen X, Tang Y, Mo Y, et al. A diagnostic model for coronavirus disease 2019 (COVID-19) based on radiological semantic and clinical features: a multi-center study. European radiology. 2020.
4. Jin C, Chen W, Cao Y, et al. Development and Evaluation of an AI System for COVID-19 Diagnosis. medRxiv. 2020:2020.2003.2020.20039834.
5. Li L, Qin L, Xu Z, et al. Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT. Radiology. 2020:200905.
6. Wang S, Kang B, Ma J, et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). medRxiv. 2020:2020.2002.2014.20023028.
7. Ying S, Zheng S, Li L, et al. Deep learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) with CT images. medRxiv. 2020:2020.2002.2023.20026930.

Chest CT – Case Series

1. Ai J, Gong J, Xing L, et al. Analysis of factors associated early diagnosis in coronavirus disease 2019 (COVID-19). medRxiv. 2020:2020.2004.2009.20059352.
2. Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. Radiology. 2020:200432.
3. Li Y, Xia L. Coronavirus Disease 2019 (COVID-19): Role of Chest CT in Diagnosis and Management. AJR American journal of roentgenology. 2020:1-7.
4. Long C, Xu H, Shen Q, et al. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? European journal of radiology. 2020;126:108961.

5. Xu YH, Dong JH, An WM, et al. Clinical and computed tomographic imaging features of novel coronavirus pneumonia caused by SARS-CoV-2. The Journal of infection. 2020;80(4):394-400.
6. Xu ZF, Wu WX, Jin YB, Pan AZ. Key Points of Clinical and CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV) Imported Pneumonia Based On 21 Cases Analysis. medRxiv. 2020:2020.2003.2003.20030775.

Chest X-Ray

1. Castiglioni I, Ippolito D, Interlenghi M, et al. Artificial intelligence applied on chest X-ray can aid in the diagnosis of COVID-19 infection: a first experience from Lombardy, Italy. medRxiv. 2020:2020.2004.2008.20040907.
2. Weinstock MB, Echenique A, Russell JW, et al. Chest x-ray findings in 636 ambulatory patients with COVID-19 presenting to an urgent care center: a normal chest x-ray is no guarantee. J Urgent Care Med. 2020;14(7):13-18.
3. Wong HYF, Lam HYS, Fong AH, et al. Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients. Radiology. 2019:201160.

3.1.3 Key Question

3 None

3.1.4 Key Question

4 None

3.1.5 Key Question 5

1. Chen J, Qi T, Liu L, et al. Clinical progression of patients with COVID-19 in Shanghai, China. The Journal of infection. 2020.
2. Colombi D, Bodini FC, Petrini M, et al. Well-aerated Lung on Admitting Chest CT to Predict Adverse Outcome in COVID-19 Pneumonia. Radiology. 2020:201433.
3. Yuan M, Yin W, Tao Z, Tan W, Hu Y. Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China. PloS one. 2020;15(3):e0230548.

3.1.6 Key Question 6

1. Grillet F, Behr J, Calame P, Aubry S, Delabrousse E. Acute Pulmonary Embolism Associated with COVID-19 Pneumonia Detected by Pulmonary CT Angiography. Radiology. 2020:201544.
2. Leonard-Lorant I, Delabranche X, Severac F, et al. Acute Pulmonary Embolism in COVID-19 Patients on CT Angiography and Relationship to D-Dimer Levels. Radiology. 2020:201561.

3.1.7 Key Question

7 None

Appendix 4

4.1 Evidence tables

4.1.1 Characteristics of diagnostic accuracy studies of chest computed tomography to diagnose COVID-19

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size
Ai, et al., 2020[15]	Retrospective cohort January 6 to February 6, 2020	Suspected of COVID-19; underwent both chest CT imaging and SARS-CoV-2 RT-PCR; time interval between CT and RT-PCR <=7 days.	China (Wuhan)	Age (mean, years): 51 Female: 54% Interval between CT and RT-PCR (median, days): 1 (range: 0 to 7)	Total: 1014 COVID-19: 601 Non-COVID-19: 413
Caruso, et al., 2020[19]	Prospective cohort March 4 to 19, 2020	Suspected COVID-19 patients with fever and respiratory symptoms such as cough, and dyspnea; patients with mild respiratory symptoms and close contact with a confirmed COVID-19 patient; or patients with a previously positive test result. Patients who underwent chest CT with contrast for vascular indication were excluded.	Italy (Rome)	Age (mean, years): 57 (range 18 to 89) Female: 47% Fever >37.5°C: 61% Cough: 56% Dyspnea: 33% WBC count elevated: NR Lymphocyte count decreased: 60% CRP increased: 88% Lactic acid dehydrogenase increased: 81% Cardiovascular disease: NR DM: NR COPD: NR HIV: NR Tuberculosis: NR Days from symptom onset (mean): NR Patients requiring hospitalization vs. home-isolation: 39% vs. 61%	Total: 158 COVID-19: 62 Non-COVID-19: 96
Dangis, et al., 2020[24]	Retrospective cohort March 14 to 24, 2020	Possible COVID-19 infection and both SARS-CoV-2 RT-PCR and low-dose chest CT at presentation.	Belgium (Bonheiden)	COVID-19 vs. non-COVID-19 Age (mean, years): 67.4 vs. 57.5 Female: 50.6% vs. 41.2% BMI (kg/m^2): 29.0 vs. 28.8 Time since symptom onset (mean, days): 7 vs. 7 Fever: 68.7% vs. 45.9% Cough: 73.5% vs. 67.0% Dyspnea: 53.0% vs. 41.3% Chest pain: 10.8% vs. 21.1%	Total: 192 COVID-19: 83 Non-COVID-19: 109

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size
Prokop, et al., 2020[32]	Retrospective cohort March 14 to 25, 2020	Presenting to the emergency department with suspected COVID-19 based on lower respiratory tract infection symptoms including cough and clinically relevant dyspnea requiring hospital admission with or without fever >38 degrees C; CT performed and SARS-CoV-2 RT-PCR within 5 days of CT .	the Netherlands	Age (mean, years): 62 Female: 42% DM: 14% Lung disease: 39% Cancer: 21% Immune deficiency: 16% Cardiovascular conditions: 44% Duration of symptoms (days, IQR): 6.0 (2.0 to 10.0)	Total: 105 COVID-19: 53 Non-COVID-19: 41
Wen, et al., 2020[35]	Retrospective cohort January 21 to February 14, 2020	Under investigation for COVID-19; excluded persons with fever more than 14 days but no acute respiratory infection signs or symptoms or exposure history; acute respiratory infection signs or symptoms >14 days but no exposure history; and acute respiratory infection symptoms in the last 14 days but no exposure history, laboratory tests, or other examination sufficient to exclude COVID-19. All patients were hospitalized >=2 weeks.	China (Henan Province)	Female: 53% Age (mean, years): 46	Total: 103 COVID-19: 88 Non-COVID-19: 15
Yang, et al., 2020[39]	Retrospective cohort January 23 to February 9, 2020	Evaluated for possible COVID-19 with RT-PCR for SARS-CoV-2 and CT.	China (Nanchang)	Age (median, years): 42 Female: 38.0% Typical symptoms: 93.8%	Total: 274 COVID-19: 53 Non-COVID-19: 221

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size
Bai, et al., 2020[16]	Case-control January 6 to February 20, 2020 (cases) and 2017 to 2019 (controls)	Cases (COVID-19): Positive SARS-CoV-2 PCR and abnormal chest CT findings. Controls (possible viral pneumonia): Positive results on Respiratory Pathogen Panel and chest CT final impression with the word "pneumonia" within 7 days.	Cases: China (Hunan province) Controls: United States (Rhode Island)	Cases vs. controls Age (mean, years): 44.8 vs. 64.7 Female: 46% vs. 50% Fever: 65% vs. 56% WBC count elevated: 29% vs. 56% Lymphocyte count decreased: 84% vs. 56% Cardiovascular disease: 6% vs. 29% COPD: 4% vs. 25% HIV: 0% vs. 2% Time from onset (mean, days): 4.9 (cases) Severity (cases): 3% mild, 87% medium, 6% severe, 3% very severe	Total: 424 Cases: 219 Controls: 205
Chen, et al., 2020[22]	Case-control January 1 to February 8, 2020	Cases: Positive for SARS-CoV-2 RT-PCR. Controls: Pneumonia without COVID-19.	China (Guangdong)	Case vs. controls (training and validation sets) Age (mean, years): 42.9 vs. 46.7 Female: 41.4% vs. 34.8% Dry cough: 68.6% vs. 84.8% Fatigue: 31.4% vs. 12.1% Sore throat:12.9% vs. 9.1% Stuffy: 2.9% vs. 6.1% Runny nose: 4.3% vs. 4.6% WBC count high: 7.1% vs. 59.1% Lymphocyte count low: 45.7% vs. 36.4% Neutrophil count high: 4.3% vs. 60.6% CRP (mg/L): 26.4 vs. 69.3	Cases: 70 (overall) Controls: 66 (overall) Validation set: 38 (number of cases and controls not reported)

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size
Case series					
Ai, et al., 2020[14]	Case series February 9 to March 20, 2020	Positive SARS-CoV-2 RT-PCR and hospitalized.	China (Xiangyang)	Age (mean, years): 50.3 Female: 50% CK elevated: 10% WBC elevated: 1.9% Neutrophils elevated: 1.9% Lymphocytes decreased: 55% Monocytes increased: 14% CRP elevated: 67%	COVID-19: 108
Fang, et al., 2020[25]	Case series January 19 to February 4, 2020	Positive SARS-CoV-2 RT-PCR with 1) travel or residential history in Wuhan/local endemic area or contact with individuals from these areas within 14 days and 2) fever or acute respiratory symptoms of unknown case. Chest CT and RT-PCR testing within 3 days or less.	China (Taizhou)	Age (median years): 45 (IQR 39-55) Female: 43% Time from disease onset to CT (mean, days): 3 Time from disease onset to RT-PCR (mean, days): 3 Number of RT-PCR tests (days) to COVID-19 diagnosis 1 (Initial): 70.5% 2 (1 to 2 days): 23.5% 3 (2 to 5 days): 4% 4 (7 days): 2%	COVID-19: 51

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size
Li and Xia., 2020[30]	Case series January 23 to 29, 2020	Positive RT-PCR for COVID-19.	China (Wuhan)	Age (mean, years): 58 Female: 45% Fever: 90% Fatigue and poor appetite: 5.9% Cough: 2.0% No symptoms: 2.0%	COVID-19: 51
Long, et al., 2020[31]	Case series January 20 to February 8, 2020	Fever >38°C and COVID-19 pneumonia suspicion, underwent thin-section CT of the chest and RT-PCR examinations.	China (Yichang)	Age (mean, years): 44.8 Female: 44.4% Duration of fever (mean, days): 2.6 Leukocyte count normal or decreased: 91.7% Lymphocyte count decreased: 63.8% Fasting glucose increased: 47.2%	COVID-19: 36

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size
Xu, et al., 2020[37]	Case series January and February 2020	Positive RT-PCR for COVID-19 and CT scan available.	China (Baoding)	COVID-19 severity of mild vs. moderate vs. severe vs. critically severe Female: 100% vs. 30% vs. 46% vs. 22% Age >50: 66% vs. 33% vs 35% vs. 0% Fever: >38.1 degrees C: 67% vs. 50% vs. 39% vs. 33% Cough: 67% vs. 40% vs. 39% vs. 33% Expectoration: 33% vs. 10% vs. 14% vs. 11% Fatigue: 33% vs. 50% vs. 7.1% vs. 0% Headache: 33% vs. 10% vs. 11% vs. 0% Dyspnea: 33% vs. 30% Vs. 0% vs. 0% Normal or slightly reduced leukocyte count: 100% vs. 100% vs. 100% vs. 89% Decreased counts of lymphocytes: 67% vs. 40% vs. 21% vs. 22% Increased C-reactive protein: 100% vs. 50% vs. 57% vs. 22%	Total COVID-19: 50 Critically severe: 3 Severe: 10 Moderate: 28 Mild: 9
Xu, et al, 2020[38]	Case series January 20 to February 6, 2020	Confirmed SARS-CoV-2 infection.	China (Foshan)	COVID-19 severity of severe vs. mild Age: 59 vs. 35 Female: 43% vs. 57% Fever: 100% vs. 71% Cough: 100% vs. 71% Shortness of breath: 86% vs. 21% Weak: 67% vs. 33% Diarrhea: 29% vs 0% Muscle Pain: 0% vs. 21% Exposure history: 100% vs. 100%	Total COVID-19: 21 (4 patients without lung lesions excluded from analysis of correlates of severity) Severe: 7 Mild: 14 (4 without lung lesions)

Abbreviations: BMI=body mass index; CK=creatinine kinase; COPD=chronic obstructive pulmonary disease; COVID-19=coronavirus disease 2019; CRP=c-reactive protein; CT=computed tomography; DM=diabetes mellitus; HIV=human immunodeficiency virus; NR=not reported; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2; WBC=white blood cell

4.1.2 Characteristics of diagnostic accuracy studies of chest computed tomography to diagnose COVID-19, continued

Author, Year	Imaging	Reference Standard	Definition of Positive Test	Imaging Reader
Ai, et al., 2020[15]	CT Reconstructed slice thickness: 0.625 to 1.25 mm Automatic current tube modulation	SARS-CoV-2 PCR within 7 days from CT	Imaging read as positive for COVID-19	2 radiologists who came to consensus
Caruso, et al., 2020[19]	CT Reconstruction slice thickness: 1.25 mm Automatic current tube modulation	COVID-19 positive: SARS-CoV-2 RT-PCR COVID-19 negative: Negative SARS-CoV-2 RT-PCR followed by >=1 negative PCR for non-cases	CT positive for viral pneumonia using clinically available dedicated application (Thoracic VCAR v13.1, GE)	2 radiologists in consensus evaluated images using a clinically available dedication application for diagnosis of viral pneumonia
Dangis, et al., 2020[24]	CT Reconstructed slice thickness: 1 mm and 0.7 mm increment with standard lung-tissue kernel and 3 mm and 3 mm increment with standard soft tissue kernel Low-dose chest CT protocol applied Dose-length product (mGy-cm): 41.4 vs. 38.7 Effective dose (mSv): 0.58 vs. 0.54	SARS-CoV-2 RT-PCR, initial negative RT-PCR underwent repeat RT-PCR the following day	Imaging classified as positive for COVID-19 (scored based on the presence of findings as presented by Ng et al and Shi et al)	2 radiologists with 8 and 7 years of experience
Prokop, et al., 2020[32]	CT Slice thickness: NR Dose length product (mGy-cm): 39.4	1: SARS-CoV-2 RT-PCR 2: SARS-CoV-2 RT-PCR or clinical diagnosis with negative RT-PCR	Based on categorization using the COVID-19 Reporting and Data System, threshold not utilized (only AUROC reported)	Average of 8 radiologists (4 had <5 years experience, the remainder had 5 to 27 years experience)
Wen, et al., 2020[35]	CT Slice thickness: 2 to 3 mm without interslice gap Automatic current tube modulation CT dose index (mGy): 9.34 4.13 Dose-length product (mGy-cm): 314.03	SARS-CoV-2 RT-PCR First RT-PCR positive: 42% Second RT-PCR: 33% Third RT-PCR: 16% Fourth RT-PCR: 9%	CT read as positive for COVID-19; Fleischner Society lexicon used	3 radiologists with 8 to 15 years of experience; disagreements resolved through discussion and consensus
Yang, et al., 2020[39]	CT Slice thickness: NR	SARS-CoV-2 RT-PCR	A: Imaging read as positive B: Imaging total score>=2 C: Imaging read as positive and score >=2 D: Imaging read as positive or score >=2	2 radiologists jointly reviewed CT images

Author, Year	Imaging	Reference Standard	Definition of Positive Test	Imaging Reader
Bai, et al., 2020[16]	CT Reconstructed slice thickness: Varied (0.6 to 10 mm, different machines) Automatic current tube modulation	Cases: SARS-CoV-2 RT-PCR Controls: Positive Respiratory Pathogen Panel and chest CT final impression with the word "pneumonia" within 7 days	Imaging read as positive for COVID-19	A: Radiologist 1 (China) B: Radiologist 2 (China) C: Radiologist 3 (China) D: Radiologist 4 (United States) E: Radiologist 5 (United States) F: Radiologist 6 (United States) G: Radiologist 7 (United States)
Chen, et al., 2020[22]	CT Reconstructed slice thickness: 1 mm slices with a slice gap of 0.8 mm	Cases: SARS-CoV-2 RT-PCR positive Controls: Consecutive negative SARS-CoV-2 RT-PCR	Classified as positive for COVID-19 by multivariate logistic regression model based on radiological features	2 senior radiologists
Ai, et al., 2020[14]	CT Slice thickness: NR	SARS-CoV-2 RT-PCR First RT-PCR positive: 68% Second RT-PCR positive: 16% Third RT-PCR positive: 8.3% Fourth or fifth RT-PCR positive: 8.3%	"Relatively obvious imaging features of COVID-19"	Not described
Fang, et al., 2020[25]	CT Slice thickness: 5 mm	SARS-CoV-2 RT-PCR First RT-PCR positive: 71%	CT compatible with viral pneumonia (including typical and atypical CT manifestations)	Not described
Li and Xia., 2020[30]	CT Slice thickness: 1.25 mm	SARS-CoV-2 RT-PCR	CT read as viral pneumonia	Read by 1 radiologist and checked by another
Long, et al., 2020[31]	CT Reconstructed slice thickness: 2 mm	SARS-CoV-2 RT-PCR Among cases, first RT-PCR positive: 30 Second RT-PCR positive: 3 Third RT-PCR positive: 3	Abnormal CT consistent with COVID-19	2 radiologists; disagreement resolved by consensus
Xu, et al., 2020[37]	CT Reconstruction slice thickness: 0.625 mm Automatic current tube modulation	SARS-CoV-2 RT-PCR	CT positive based on lesion distribution, lesion location, lesion density (ground glass opacity, consolidation, or mixed), thickness of interlobular and intralobular septa, enlarged lymph nodes within the mediastinum and pleural effusion	3 experienced radiologists; disagreements resolved by consensus

Author, Year	Imaging	Reference Standard	Definition of Positive Test	Imaging Reader
Xu, et al, 2020[38]	CT Slice thickness: 1 mm	"Laboratory testing of respiratory secretions"	CT read as positive for COVID-19	2 radiologists with 8 to 10 years experience, disagreements resolved by 3rd radiologist

Abbreviations: AUROC=area under the receiver operating characteristic; COVID-19=coronavirus disease 2019; CT=computed tomography; NR=not reported; PCR=polymerase chain reaction; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2

4.1.3 Results of diagnostic accuracy studies of chest computed tomography to diagnose COVID-19

Author, Year	True Positives (n)	False Positives (n)	False Negatives (n)	True Negatives (n)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	AUROC	Risk of bias Other limitations
Ai, et al., 2020[15]	Overall: 580 <60 years: 362 >=60 years: 218 Female: 308 Male: 272	Overall: 308 <60 years: 225 >=60 years: 83 Female: 160 Male: 148	Overall: 21 <60 years:15 >=60 years: 6 Female: 9 Male: 12	Overall: 105 <60 years: 81 >=60 years: 24 Female: 70 Male: 35	Overall: 0.96 (0.95 to 0.98) <60 years: 0.96 (0.94 to 0.98) >=60: 0.97 (0.94 to 0.99) Female: 0.97 (0.95 to 0.99) Male: 0.96 (0.93 to 0.98)	Overall: 0.25 (0.22 to 0.30) <60 years: 0.26 (0.22 to 0.32) >=60: 0.22 (0.16 to 0.31) Female: 0.30 (0.25 to 0.37) Male: 0.19 (0.14 to 0.25)	Overall: 0.65 (0.62 to 0.68) <60 years: 0.62 (0.58 to 0.66) >=60 years: 0.72 (0.67 to 0.77) Female: 0.66 (0.60 to 0.69) Male: 0.65 (0.60 to 0.69)	Overall: 0.83 (0.76 to 0.89) <60 years: 0.84 (0.76 to 0.90) >=60 years: 0.80 (0.63 to 0.91) Female: 0.89 (0.80 to 0.94) Male: 0.74 (0.61 to 0.85)	NR	Moderate
Caruso, et al., 2020[19]	60	42	2	54	0.97 (0.88 to 0.99)	0.56 (0.45 to 0.66)	0.59 (0.53 to 0.64)	0.96 (0.87 to 0.99)	NR	Moderate
Dangis, et al., 2020[24]	1 (all patients): 72 2 (clinical symptoms >48 hours): 65	1: 7 2: 6	1: 11 2: 3	1: 102 2: 82	1: 0.87 (0.80 to 0.98) 2: 0.96 (0.91 to 0.999)	1: 0.94 (0.89 to 0.982) 2: 0.93 (0.88 to 0.98)	1: 0.91 (0.85 to 0.97) 2: 0.92 (0.85 to 0.98)	1: 0.90 (0.85 to 0.96) 2: 0.96 (0.92 to 0.999)	NR	Moderate
Prokop, et al., 2020[32]	NR	NR	NR	NR	NR	NR	NR	NR	1: 0.91 (0.85 to 0.97) 2: 0.95 (0.91 to 0.99)	Moderate
Wen, et al., 2020[35]	82	7	6	8	0.93 (0.86 to 0.97)	0.53 (0.27 to 0.79)	0.92 (0.87 to 0.95)	0.57 (0.35 to 0.77)	NR	Moderate NPV appears to be an error, calculated as 0.57

Author, Year	True Positives (n)	False Positives (n)	False Negatives (n)	True Negatives (n)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	AUROC	Risk of bias Other limitations
Yang, et al., 2020[39]	A: 48 B: 47 C: 42 D: 53	A: 70 B: 151 C: 52 D: 169	A: 5 B: 6 C: 11 D: 0	A: 151 B: 70 C: 52 D: 52	A: 0.91 (0.79 to 0.97) B: 0.89 (0.77 to 0.96) C: 0.79 (0.66 to 0.89) D: 1.0 (0.93 to 1.0)	A: 0.68 (0.62 to 0.74) B: 0.32 (0.26 to 0.38) C: 0.50 (0.40 to 0.60) D: 0.24 (0.18 to 0.30)	A: 0.41 (0.36 to 0.46) B: 0.24 (0.21 to 0.26) C: 0.45 (0.39 to 0.51) D: 0.24 (0.23 to 0.25)	A: 0.97 (0.93 to 0.99) B: 0.92 (0.84 to 0.96) C: 0.83 (0.73 to 0.89) D: 1.0	A: 0.79 (0.86 to 0.73) B: 0.60 (0.52 to 0.68) C: 0.78 (0.85 to 0.71) D: 0.62 (0.69 to 0.54)	Moderate
Bai, et al., 2020[16]	A: 158 B: 157 C: 206 Age-matched sample (n=58) A: 24 B: 20 C: 29 D: 28 E: 25 F: 22 G: 21	A: 13 B: 24 C: 156 Age-matched sample (n=58) A: 0 B: 2 C: 26 D: 0 E: 2 F: 2 G: 0	A: 61 B: 62 C: 13 Age-matched sample (n=58) A: 6 B: 10 C: 1 D: 2 E: 5 F: 8 G: 9	A: 192 B: 181 C: 49 Age-matched sample (n=58) A: 28 B: 26 C: 2 D: 28 E: 26 F: 26 G: 28	A: 0.72 (0.66 to 0.78) B: 0.72 (0.65 to 0.78) C: 0.94 (0.90 to 0.97) Age to matched sample (n=58) A: 0.80 (0.61 to 0.92) B: 0.67 (0.47 to 0.83) C: 0.97 (0.83 to 1.00) D: 0.93 (0.78 to 0.99) E: 0.83 (0.65 to 0.94) F: 0.73 (0.54 to 0.88) G: 0.70 (0.51 to 0.85)	A: 0.94 (0.89 to 0.97) B: 0.88 (0.83 to 0.92) C: 0.24 (0.13 to 0.30) Age to matched sample (n=58) A: 1.00 (0.88 to 1.00) B: 0.93 (0.76 to 0.99) C: 0.07 (0.01 to 0.24) D: 1.00 (0.88 to 1.00) E: 0.93 (0.76 to 0.99) F: 0.93 (0.76 to 0.99) G: 1.00 (0.88 to 1.00)	A: 0.92 (0.87 to 0.96) B: 0.87 (0.81 to 0.91) C: 0.57 (0.52 to 0.62) Age to matched sample (n=58) A: 1.00 (0.86 to 1.00) B: 0.91 (0.71 to 0.99) C: 0.53 (0.39 to 0.66) D: 1.00 (0.88 to 1.00) E: 0.93 (0.76 to 0.99) F: 0.92 (0.73 to 0.99) G: 1.00 (0.84 to 1.00)	A: 0.76 (0.70 to 0.81) B: 0.74 (0.69 to 0.80) C: 0.79 (0.67 to 0.88) Age to matched sample (n=58) A: 0.82 (0.65 to 0.93) B: 0.72 (0.55 to 0.86) C: 0.67 (0.09 to 0.99) D: 0.93 (0.78 to 0.99) E: 0.84 (0.66 to 0.95) F: 0.76 (0.59 to 0.89) G: 0.76 (0.59 to 0.88)	Not reported	High Cases and controls from different countries, cases with negative CT excluded; marked variability in accuracy by Chinese radiologists

Author, Year	True Positives (n)	False Positives (n)	False Negatives (n)	True Negatives (n)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	AUROC	Risk of bias Other limitations
Chen, et al., 2020[22]	NR	NR	NR	NR	1.0 (CI NR)	0.37 (CI NR)	NR	NR	0.81 (0.67 to 0.95)	Moderate Small validation set; number of cases and controls in validation set not reported; unable to calculate PPV and NPV
Ai, et al., 2020[14]	96	NR	12	NR	0.89 (0.81 to 0.94)	NR	NR	NR	0.74 (0.69 to 0.80)	High Non-peer reviewed
Fang, et al., 2020[25]	50	NR	1	NR	0.95 (0.90 to 1.00)	NR	NR	NR	NR	High
Li and Xia., 2020[30]	49	NR	2	NR	0.96 (0.87 to 0.995)	NR	NR	NR	NR	High Included 2 patients diagnosed with adenovirus (data not abstracted here); RT-PCR not available for all patients
Long, et al., 2020[31]	35	NR	1	NR	0.97 (0.85 to 0.999)	NR	NR	NR	NR	High Small sample size

Author, Year	True Positives (n)	False Positives (n)	False Negatives (n)	True Negatives (n)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	AUROC	Risk of bias Other limitations
Xu, et al., 2020[37]	41	NR	9	NR	0.82 (0.69 to 0.91)	NR	NR	NR	NR	High
Xu, et al, 2020[38]	Total: 17 Severe: 7 Mild: 10	NR	Total: 4 Severe: 0 Mild: 4	NR	Total: 0.81 (0.58 to 0.95) Severe: 1.0 Mild: 0.71 (0.42 to 0.92)	NR	NR	NR	NR	High

Abbreviations: AUROC=area under the receiver operating characteristic; CI=confidence interval; CT=computed tomography; NPV=negative predictive value; NR=not reported; PPV=positive predictive value; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2

4.1.4 Characteristics of diagnostic accuracy studies of artificial intelligence systems to diagnose COVID-19

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size
Bai, et al., 2020[17]	Case-control Cases: January to April 2020 Controls: 2017 to 2019	Cases: Positive SARS-C0V-2 PCR and abnormal chest CT. Controls: Non-COVID-19 pneumonia and definite evidence of pneumonia on chest CT.	China (Hunan province), United States (Rhode Island and Pennsylvania)	Cases vs. controls (training, validation, and test sets) Age (mean, years): 46 vs. 62 Female: 48% vs. 42% Fever: 58% vs. 54% WBC elevated: 2% vs. 51% Lymphocyte count deceased: 36% vs. 55% Cardiovascular disease: 3% vs. 35% HTN: 12% vs. 39% COPD: 4% vs. 24% DM: 6% vs. 17% HIV: 0% vs. 2% Time from onset to presentation <10 days in COVID-19 patients: 60% 10 to 29 days: 19% >=30 days: 12% COVID-19 mild severity: 7% Medium: 78% Severe: 10% Critical: 5%	Cases: 521 (all patients) Controls: 665 (all patients) Test set for analysis: 119 (number of cases and controls unclear)
Jin, et al., 2020[27]	Case-control Dates not reported	Cases: COVID-19 cases from 3 hospitals in Wuhan, criteria for diagnosis not reported. Controls: Healthy controls from 3 hospitals in Wuhan and non-COVID-19 controls (clinical status not reported) from 2 international databases.	Cases: China (Wuhan) Controls: China (Wuhan) and international databases	External test cohort (Wuhan) (n=299) Age <20 years: 0.3% Age 20-39 years: 48.4% Age 40-59 years: 29.4% Age >=60 years: 21.7% Female: 66.2% COVID-19 positive: 61.2% COVID-19 negative: 38.9%	External cohort Total: 1,255 (combines Wuhan patients and international databases Cases: NR Controls: NR
Li, et al., 2020[29]	Case-control Cases: December 31, 2019 to February 17, 2020 Controls: August 16, 2016 to February 17, 2020	Cases: Positive for SARS-CoV-2 RT-PCR . Controls: Community acquired pneumonia or non-pneumonia patients.	China	Testing set, cases vs. controls (community acquired pneumonia) vs. controls (non-pneumonia) Age (mean, years): 52 vs. 53 vs. 41 Female: 48% vs. 38% vs. 48% COVID-19 cases, all training and testing sets Fever: 81% Cough: 66% Time from symptom onset to CT (median, days): 7 (range 0 to 20)	Testing set Total: 353 (434 CT scans) Cases: 68 (131 CT scans) Controls: 285 (155 community acquired pneumonia [175 CT scans], 130 non-pneumonia [132 CT scans])

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size
Wang, et al., 2020[33]	Case-control Dates not reported	Cases: Positive SARS-CoV-2 RT-PCR. Controls: Non-COVID-19.	Not reported	Not reported	Cases: 70 Controls: 220
Ying, et al., 2020[40]	Case-control Dates not reported	Cases: COVID-19 cases from 2 hospitals in Wuhan, diagnostic criteria not reported. Controls: Patients with bacterial pneumonia from 1 hospital.	China (Wuhan, Guangzhou)	Not reported	Test cohort Cases: 27 Controls: 24 (healthy controls) and 30 (bacterial pneumonia)

Abbreviations: COPD=chronic obstructive pulmonary disease; COVID-19=coronavirus disease 2019; CT=computed tomography; DM=diabetes mellitus; HTN=hypertension; NR=not reported; PCR=polymerase chain reaction; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2; WBC=white blood cell

4.1.5 Characteristics of diagnostic accuracy studies of artificial intelligence systems to diagnose COVID-19, continued

Author, Year	Imaging	Reference Standard	Definition of Positive Test	Imaging Reader
Bai, et al., 2020[17]	CT Reconstructed slice thickness: Varied (0.6 to 10 mm, different machines) Automatic tube current modulation	SARS-CoV-2 RT-PCR	Imaging read as positive for COVID-19	1: Average of 6 radiologists 2: Artificial intelligence model 3: Radiologists with AI assistance
Jin, et al., 2020[27]	CT Slice thickness: NR	Not reported	Positive classification for COVID-19 by AI system	AI (based on deep convolutional neural network)
Li, et al., 2020[29]	CT Slick thickness: 0.5 to 3 mm	Cases: SARS-CoV-2 RT-PCR Controls: Bacterial cultures for community acquired pneumonia	Positive classification for COVID-19 by AI system	AI (COVID-19 detection neural network [COVnet])
Wang, et al., 2020[33]	NR	SARS-CoV-2 RT-PCR	Positive for COVID-19 by deep learning algorithm	A: AI (deep learning algorithm) B: Radiologist 1 C: Radiologist 2
Ying, et al., 2020[40]	CT (spiral) Reconstructed thickness: 0.625 mm Automatic current tube modulation	SARS-CoV-2 RT-PCR	Positive classification for COVID-19 vs.: A: Healthy controls B: Bacterial pneumonia	AI: (details relation extraction neural network)

Abbreviations: AI=artificial intelligence; COVID-19=coronavirus disease 2019; CT=computed tomography; NR=not reported; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2

4.1.6 Results of diagnostic accuracy studies of artificial intelligence systems to diagnose COVID-19

Author, Year	True Positives (n)	False Positives (n)	False Negatives (n)	True Negatives (n)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	AUROC (95% CI)	Risk of bias Other limitations
Bai, et al., 2020[17]	NR	NR	NR	NR	1: 0.79 (0.64 to 0.89) 2: 0.95 (0.83 to 1.0) 3: 0.88 (0.74 to 0.95)	1: 0.88 (0.78 to 0.94) 2: 0.96 (0.88 to 0.99) 3: 0.91 (0.82 to 0.96)	NR	NR	NR	High Cases and controls assembled from different countries; cases with negative CT excluded; unable to calculate PPV and NPV
Jin, et al., 2020[27]	NR	NR	NR	NR	0.95 (0.95 to 0.96)	0.94 (0.93 to 0.95)	0.92 (0.91 to 0.93)	0.97 (0.96 to 0.97)	0.98 (0.98 to 0.98)	High Non-peer reviewed Cases and controls from different geographic areas, no information about clinical presentation, no information about reference standard
Li, et al., 2020[29]	114	13	13	294	0.90 (0.83 to 0.94)	0.96 (0.93 to 0.98)	0.09 (0.84 to 0.94)	0.96 (0.93 to 0.97)	0.96 (0.94 to 0.99)	High
Wang, et al., 2020[33]	NR	NR	NR	NR	A: 0.75 (CI NR) B: 0.71 (CI NR) C: 0.73 (CI NR)	A: 0.86 (CI NR) B: 0.51 (CI NR) C: 0.50 (CI NR)	A: 0.69 (CI NR) B: 0.29 (CI NR) C: 0.29 (CI NR)	A: 0.89 (CI NR) B: 0.86 (CI NR) C: 0.86 (CI NR)	A: 0.81 (0.71 to 0.84) B: NR C: NR	High Non-peer reviewed
Ying, et al., 2020[40]	A: 25 B: 26	A: 1 B: 7	A: 2 B: 1	A: 23 B: 23	A: 0.93 (0.76 to 0.99) B: 0.96 (0.81 to 0.999)	A: 0.96 (0.79 to 0.999) B: 0.77 (0.58 to 0.90)	A: 0.96 (0.79 to 0.99) B: 0.79 (0.66 to 0.88)	A: 0.92 (0.75 to 0.98) B: 0.96 (0.77 to 0.99)	A: 0.99 (CI NR) B: 0.95 (CI NR)	High Non-peer reviewed

Abbreviations: AUROC=area under the receiver operator curve; CI=confidence interval; NPV=negative predictive value; NR=not reported; PPV=positive predictive value

4.1.7 Characteristics of diagnostic accuracy studies of chest x-ray to diagnose COVID-19

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size	Imaging	Reference Standard	Definition of Positive Test	Imaging Reader
Castiglioni, et al., 2020[20]	Retrospective cohort March 14 to 20, 2020	Suspected of COVID-19 and admitted to 1 hospital.	Italy (Monza)	Age, sex, and clinical characteristics not reported	Total (external cohort): 110 COVID-19: 74 Non-COVID-19: 36	CXR (digital bedside)	COVID-19 positive: SARS-CoV-2 RT-PCR positive COVID-10 negative: Negative PCR followed by at least 1 negative PCR	Imaging classified as positive for COVID-19	1: Artificial intelligence (ten convolutional neural networks) 2: Radiologist 1 3: Radiologist 2
Weinstock, et al., 2020[34]	Case series March 9 to 24, 2020	Presented to urgent care center and positive SARS-CoV-2 RT-PCR.	United States (New York City and New Jersey)	Female: 43% Age 30 to 70 years: 78%	COVID-19: 636	CXR	SARS-CoV-2 RT-PCR	A: CXR read as mildly, moderately, or severely abnormal B: CXR read as moderately or severely abnormal	Each CXR read by 1 of 11 radiologists

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size	Imaging	Reference Standard	Definition of Positive Test	Imaging Reader
Wong, et al., 2020[36]	Case series January 1 to March 5, 2020	Positive SARS-CoV-2 RT-PCR in patients from 4 regional and tertiary hospitals.	Hong Kong	Age (mean, years): 56 Female: 59% Mild fever (37-38°C): 38% High fever (≥38°C): 22% Cough: 41% Sputum: 20% Hemoptysis: 0% Sore throat: 14% Diarrhea: 5% Chest discomfort: 9% Dyspnea: 6% Asymptomatic: 14% Diabetes: 13% Hypertension: 20% COPD: 0%	COVID-19: 64	CXR	RT-PCR positive for SARS-CoV-2 First RT-PCR positive: 91% Subsequent RT-PCR positive: 9%	Abnormality on CXR, otherwise not defined; severity score calculated using adapted and simplified Radiographic Assessment of Lung Edema score	2 radiologists reviewed by consensus, disagreements were resolved by a 3rd radiologist if needed

Abbreviations: COPD=chronic obstructive pulmonary disease; COVID-19=coronavirus disease 2019; CXR=chest x-ray; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2

4.1.8 Results of diagnostic accuracy studies of chest x-ray to diagnose COVID-19

Author, Year	True Positives (n)	False Positives (n)	False Negatives (n)	True Negatives (n)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	AUROC	Risk of bias Other limitations
Castiglioni, et al., 2020[20]	1: 59 2: 47 3: 47	1: 7 2: 8 3: 5	1: 15 2: 27 3: 27	1: 29 2: 28 3: 31	1: 0.80 (0.72 to 0.86) 2: 0.64 (0.52 to 0.74) 3: 0.64 (0.52 to 0.74)	1: 0.81 (0.73 to 0.87) 2: 0.78 (0.61 to 0.90) 3: 0.86 (0.71 to 0.95)	1: 0.89 (0.82 to 0.94) 2: 0.85 (0.76 to 0.92) 3: 0.90 (0.80 to 0.96)	1: 0.66 (0.57 to 0.75) 2: 0.51 (0.42 to 0.59) 3: 0.53 (0.61 to 0.79)	NR	High Non-peer reviewed
Weinstock, et al., 2020[34]	A: 265 B: 70	NR	A: 371 B: 566	NR	A: 0.42 (0.38 to 0.46) B: 0.11 (0.08 to 0.14)	NR	NR	NR	NR	High
Wong, et al., 2020[36]	44	NR	20	NR	0.69 (0.56 to 0.80)	NR	NR	NR	NR	High

Abbreviations: AUROC=area under the receiver operator curve; CI=confidence interval; NR=not reported

4.1.9 Characteristics of diagnostic accuracy study of lung ultrasound to diagnose COVID-19

Author, Year	Study Design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size	Imaging	Reference Standard	Definition of Positive Test	Imaging Reader
Benchoufi, et al., 2020[18]	Case series March 19 to April 1, 2020	Suspected or diagnosed COVID-19 and underwent chest CT and lung US.	France (Paris)	Age (mean, years): 61 Female: 36% Smoker: 17% HTN: 29% DM: 16% Oxygen saturation <95%: 50%	Total: 107	CT (parameters not described) US: Various types of probes, curved or linear, 2 to 12 MHz	1: CT classification of abnormal (vs. normal) 2: CT classification of moderate or severe (>=10% of lung parenchyma) (vs. mild or normal)	US score >=1 (for abnormal) or >=6 (for moderate or severe), based on sum of severity scores (0=up to 3 observed B-lines to 3=consolidation foci) at 8 points of the chest wall	CT: Not described US: 1 physician for US (expert read 68% of US and trainee 32%)

Abbreviations: COVID-19=coronavirus disease 2019; CT=computed tomography; DM=diabetes mellitus; HTN=hypertension; US=ultrasound

4.1.10 Results of diagnostic accuracy study of lung ultrasound to diagnose COVID-19

Author, Year	True Positives (n)	False Positives (n)	False Negatives (n)	True Negatives (n)	Sensitivity	Specificity	Positive predictive value	Negative predictive value	AUROC	Risk of bias Other limitations
Benchoufi, et al., 2020[18]	NR	NR	NR	NR	1: 0.95 (CI NR) 2: 0.92 (CI NR)	1: 0.83 (CI NR) 2: 0.74 (CI NR)	NR	NR	1: 0.93 (0.95 in multivariate logistic regression) 2: 0.89 (0.90 in multivariate logistic regression)	High Non-peer reviewed Some ultrasounds performed by trainee; 17 patients excluded due to missing CT severity score; reference standard was CT; imaging reader for CT not described

Abbreviations: AUROC=area under the receiver operator curve; CI=confidence interval; CT=computed tomography; NR=not reported

4.2.1 Characteristics of studies of prediction of outcomes after diagnosis of COVID-19

Author, Year	Study design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size	Imaging	Definition of Positive Test	Imaging reader
Chen, et al., 2020[21]	Retrospective cohort January 20 to February 6, 2020	Diagnosed with COVID-19.	China (Shanghai)	Age (median, years): 51 Female: 49% Time from onset of symptoms (mean, days): 4 Fever: 87% Cough: 36% Fatigue: 16% Dizziness and headache: 11% Shortness of breath: 7.6% Rhinorrhea: 6.8% Sore throat: 6.4% Diarrhea: 3.2% Loss of appetite: 3.2% Asymptomatic: 2.8% Cardiovascular and cerebrovascular diseases: 22% Respiratory system diseases: 2.0% WBC count (x 109/L): 4.71 Lymphocyte count (x 109/L): 1.12 CRP (mg/L): 12	Total: 249 ICU admission: 22 No ICU admission: 227	CT Parameters not reported	Radiological lesion (not otherwise defined)	Not reported

Author, Year	Study design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size	Imaging	Definition of Positive Test	Imaging reader
Colombi, et al., 2020[23]	Retrospective cohort February 17 to March 10, 2020	SARS-CoV-2 RT-PCR positive, with imaging findings on chest CT.	Italy (Piacenza)	ICU admission or death vs. no ICU admission or death Age (mean, years): 73 vs. 62 Female: 26% vs. 24% Smoking (current or former): 18% vs. 10% CV comorbidities: 71% vs. 39% Pulmonary comorbidities: 20% vs. 14% Chronic kidney failure: 11% vs. 2% Diabetes: 20% vs. 11% Fever: 99% vs. 96% Cough: 62% vs. 60% Dyspnea: 43% vs. 28% Asthenia: 12% vs. 12% Other: 22% vs. 18% Time since symptom onset: 5 vs. 6 Temperature at admission (degrees C): 37.8 vs. 37.5 SpO2 (%): 91% vs. 94% WBC count (x 109/L): 6.8 vs. 5.2 Lymphocyte count (x 109/L): 0.87 vs. 1.1 CRP (mg/dL): 13.3 vs. 5.1	Total: 236 ICU admission or death: 108 No ICU admission or death: 128	CT Reconstruction slice thickness: 1 to 2 mm Low-dose CT acquisition performed	1: Clinical model 2: Model with % lung well-aerated assessed visually and clinical parameters; threshold not pre-specified 3: Model with % lung well-aerated assessed with software and clinical parameters; threshold not pre-specified 4: Model with clinical parameters, well-aerated lung volume <2.9 L and adipose tissue are >262 cm2; threshold not pre-specified	1: Not applicable 2: 2 radiologists with 5 and 14 years of experience 3: Software to calculate CT parameters 4: Software to calculate CT parameters

Author, Year	Study design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size	Imaging	Definition of Positive Test	Imaging reader
Yuan, et al., 2020[41]	Retrospective cohort January 1 to 25, 2020	Diagnosed with COVID-19 (SARS-CoV-2 RT-PCR positive) and discharged with recovered symptoms or died in hospital.	China (Wuhan)	Mortality vs. survival Age (median, years): 68 vs. 55 Female: 60% vs. 53% Time since symptom onset (median, days): 8 HTN: 50% vs. 0% DM: 60% vs. 0% Cardiac disease: 30% vs. 0% Fever: 60% vs. 88% Cough: 50% vs. 65% Myalgia: 10% vs. 12% Dyspnea: 100% vs. 6%	Total: 27 Mortality: 10 Survival: 17	CT Slice thickness: 5 mm	CT score >24.5; sum of radiologic score (1 =normal attenuation, 2=ground glass, 3=consolidation) times lung parenchyma distribution score (1=<25% abnormality, 2=25-50%, 3=50-75%, 4 >75%) for 6 lung zones (range 0 to 72)	2 radiologists, discrepancies resolved by consensus

Abbreviations: COVID-19=coronavirus disease 2019; CRP=c-reactive protein; CT=computed tomography; CV=cardiovascular; DM=diabetes mellitus; HTN=hypertension; ICU=intensive care unit; RT-PCR=reverse transcription polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2; WBC=white blood cell

4.2.2 Results of studies predicting outcome after COVID-19

Author, Year	Outcome	Results	Risk of bias and other limitations
Chen, et al., 2020[21]	ICU admission	Risk of ICU admission, radiological lesion vs. no radiological lesion: Unadjusted OR 4.46 (95% CI 0.62 to 31.9); not included in multivariate model	Severity of symptoms at baseline unclear; "radiological finding" not defined
Colombi, et al., 2020[23]	ICU admission or death	Sensitivity (95% CI) 1: 0.75 (0.66 to 0.82) 2: 0.72 (0.63 to 0.80) 3: 0.75 (0.66 to 0.83) 4: 0.75 (0.66 to 0.83) Specificity (95% CI) 1: 0.73 (0.65 to 0.81) 2: 0.81 (0.73 to 0.88) 3: 0.80 (0.72 to 0.86) 4: 0.81 (0.73 to 0.88) Positive predictive value (95% CI) 1: 0.70 (0.61 to 0.78) 2: 0.76 (0.68 to 0.82) 3: 0.75 (0.68 to 0.81) 4: 0.77 (0.69 to 0.83) Negative predictive value (95% CI) 1: 0.78 (0.72 to 0.83) 2: 0.78 (0.73 to 0.83) 3: 0.80 (0.73 to 0.85) 4: 0.79 (0.74 to 0.84) AUROC (95% CI) 1: 0.83 (0.78 to 0.88) 2: 0.86 (0.81 to 0.90) 3: 0.86 (0.80 to 0.90) 4: 0.86 (0.81 to 0.90)	No control for confounders; severity of symptoms at baseline unclear; thresholds for CT findings not pre-specified
Yuan, et al., 2020[41]	Mortality	Sensitivity: 0.96 (CI NR) Specificity: 0.84 (CI NR) AUROC: 0.90 (95% CI 0.87 to 0.93)	No control for confounders; 95% CI for sensitivity and specificity, PPV, and NPV not calculable; threshold for CT score not pre-specified; small sample; severity of symptoms at baseline unclear

Abbreviations: AUROC=area under the receiver operator curve; CI=confidence interval; CT=computed tomography; ICU=intensive care unit; NPV=negative predictive value; NR=not reported; OR=odds ratio; PPV=positive predictive value

4.2.3 Characteristics of studies of pulmonary embolism outcomes after diagnosis of COVID-19

Author, Year	Study design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size	Imaging	Definition of Positive Test	Imaging reader
Grillet, et al., 2020[26]	Imaging series March 15 to April 14, 2020	SARS-CoV-2 RT-PCR positive (n=97) or positive CT and negative RT-PCR (n=9), underwent contrast CT (performed when clinical features of severe disease were present).	France (Besancon)	Pulmonary embolus vs. no pulmonary embolus Age (mean, years): 67 vs. 66 Female: 9% vs. 36% Cardiovascular disease: 43% vs. 38% Chronic respiratory insufficiency: 17% vs. 13% DM, type 2: 23% vs. 18% Malignancy: 23% vs. 21% Critical care: 74% vs. 29% Time from symptom onset to CT scan (days): 12 vs. 8 Mechanical ventilation: 65% vs. 25%	100	CT Contrast: 60 mL iodinated contrast agent at flow rate of 4 mL/s, triggered on the main pulmonary artery Slice thickness: Not reported	CT with contrast read as positive for pulmonary embolus	2 radiologists with 6 and 11 years of experience

Author, Year	Study design Dates	Eligibility Criteria	Country	Population Characteristics	Sample Size	Imaging	Definition of Positive Test	Imaging reader
Leonard-Lorant, et al., 2020[28]	Imaging series March 1 to 31, 2020	Underwent CT pulmonary angiography examination for suspicion or follow-up of SARS-CoV-2 infection.	France (Strasbourg)	Pulmonary embolus vs. no pulmonary embolus Age (median, years): 64 vs. 63, p=0.59 Female: 22% vs. 43%, p=0.04 BMI (kg/m2): 27 vs. 29, p=0.10 ICU hospitalization: 75% vs. 32%, p=0.001 SAPS II, median IQR (median points, if ICU): 46 vs. 42, p=0.37 Worst PaO2/FiO2 ratio (median): 116 vs. 168, p=0.06 Clinical suspicion for pulmonary embolus: 53% vs. 68% Thromboembolic prophylaxis before CT pulmonary angiography: 78% vs. 23%, p=0.001 Anticoagulation before CT pulmonary angiography: 6% vs. 7%, p=1.0 Time from symptom onset to CT scan (days): 14 vs. 10, p=0.001 D-dimer >=500 mcg/L: 12% vs. 32% Pulmonary embolus in main pulmonary artery: 22% Lobar artery: 34% Segmental artery: 28% Sub segmental artery: 16%	106	CT Contrast: 50 to 75 mL high concentration iodine contrast media, with bolus-tracking technique and threshold of 160 to 250 HU in the main pulmonary artery Slice thickness: 1 mm	CT with contrast read as positive for pulmonary embolus	1 radiologist

Abbreviations: BMI=body mass index; COVID-19=coronavirus disease 2019; CT=computed tomography; DM=diabetes mellitus; ICU=intensive care unit; RT-PCR=reverse transcription polymerase chain reaction; SAPS II=Simplified Acute Physiology Score; SARS-CoV-2=severe acute respiratory syndrome coronavirus 2

4.2.4 Results of studies of pulmonary embolism outcomes after diagnosis of COVID-19

Author, Year	Outcome	Results	Risk of bias and other limitations
Grillet, et al., 2020[26]	Pulmonary embolus	23.0% (23/100)	Selection of patients for CT angiography unclear; included patients with diagnosed and suspected COVID-19; clinical information limited
Leonard-Lorant, et al., 2020[28]	Pulmonary embolus	30.2% (32/106)	Selection of patients for CT angiography unclear; clinical information limited

Abbreviations: COVID-19=coronavirus disease 2019; CT=computed tomography

Appendix 5

5.1 Quality assessment table

5.1.1 Quality assessment of diagnostic accuracy studies

Author, year	Was a consecutive or random sample of patients enrolled?	Was a case-control or case series design avoided?	For case-control studies, were cases and controls matched?	Were clinical characteristics adequately described?	Were the index test results interpreted without knowledge of the results of the reference standard?	Were explicit criteria for positive imaging findings defined?	Is the reference standard likely to correctly classify the target condition?	Did patients receive the same reference standard?	Were all patents included in the analysis?	Risk of bias
Ai, et al., 2020[14]	Yes	No	No	Partial	Unclear	No	Some misclassification likely	Yes	Yes	High
Ai, et al., 2020[15]	Unclear	Yes	Not applicable	No	Yes	No	Some misclassification likely	Yes	Unclear	Moderate
Bai, et al., 2020[16]	Unclear	No	No	No	Yes	No	Some misclassification likely	Yes	Unclear	High
Bai, et al., 2020[17]	Unclear	No	No	Partial	Yes	Unclear	Yes	Yes	Unclear	High
Benchoufi, et al., 2020[18]	Unclear	Yes	Not applicable	No	Unclear	Yes	Some misclassification likely	Yes	Unclear	High
Caruso, et al., 2020[19]	Yes	Yes	Not applicable	Partial	Unclear	No	Yes	Yes	Yes	Moderate
Castiglioni, et al., 2020[20]	Yes	Yes	Not applicable	No	Unclear	Unclear	Yes	Yes	Yes	High
Chen, et al., 2020[22]	Yes	No	No	Partial	Yes	Yes	Yes	Yes	Yes	Moderate
Dangis, et al., 2020[24]	Yes	Yes	Not applicable	Partial	Yes	No	Yes	Yes	Yes	Moderate
Fang, et al., 2020[25]	Yes	No	Not applicable	Partial	Unclear	No	Yes	Yes	Yes	High
Jin, et al., 2020[27]	Unclear	No	No	No	Yes	Yes (AI)	Unclear	Unclear	Yes	High

Author, year	Was a consecutive or random sample of patients enrolled?	Was a case-control or case series design avoided?	For case-control studies, were cases and controls matched?	Were clinical characteristics adequately described?	Were the index test results interpreted without knowledge of the results of the reference standard?	Were explicit criteria for positive imaging findings defined?	Is the reference standard likely to correctly classify the target condition?	Did patients receive the same reference standard?	Were all patents included in the analysis?	Risk of bias
Li, et al., 2020[29]	Yes	No	No	No	Yes	Yes (AI)	Some misclassification likely	Yes	Yes	High
Li and Xia, 2020[30]	No	No	Not applicable	Partial	Unclear	No	Some misclassification likely	Yes	No	High
Long, et al., 2020[31]	No	No	Not applicable	Partial	Unclear	No	Yes	No	No	High
Prokop, et al., 2020[32]	Yes	Yes	Not applicable	No	Yes	Yes	Yes	Yes	Yes	Moderate
Wang, et al., 2020[33]	Unclear	No	No	No	Yes	Yes (AI)	Unclear	Yes	Yes	High
Weinstock, et al., 2020[34]	Yes	No	Not applicable	No	No	No	Unclear	Yes	Unclear	High
Wen, et al., 2020[35]	Unclear	Yes	Not applicable	No	Yes	No	Yes	Yes	Yes	Moderate
Wong, et al., 2020[36]	Unclear	No	Not applicable	Partial	Unclear	No	Yes	Yes	Yes	High
Xu, et al., 2020[37]	Unclear	No	Not applicable	Partial	Unclear	No	Some misclassification likely	Yes	Yes	High
Xu, et al., 2020[38]	Unclear	No	Not applicable	Partial	Unclear	No	Unclear	Unclear	Yes	High
Yang, et al., 2020[39]	Unclear	Yes	Not applicable	No	Unclear	Yes	Some misclassification likely	Yes	Yes	Moderate
Ying, et al., 2020[40]	Unclear	No	No	No	Yes	Yes (AI)	Some misclassification likely	Yes	Yes	High

Abbreviations: AI=artificial intelligence