Thoracic imaging tests for the diagnosis of COVID-19

Islam , N; Salameh , JP; Leeflang, Mariska M G; Hooft, Lotty; McGrath , TA; van der Pol , CB; Frank , RA; Kazi , S; Prager , R; Hare , SS; Dennie , C; Spijker , R; Deeks, Jon; Dinnes, Jacqueline; Jenniskens, K; Korevaar , DA; Cohen , JF; Van den Bruel, A; Takwoingi, Yemisi; van de Wijgert, J

DOI:
10.1002/14651858.CD013639.pub3

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Islam , N, Salameh , JP, Leeflang, MMG, Hooft, L, McGrath , TA, van der Pol , CB, Frank , RA, Kazi , S, Prager , R, Hare , SS, Dennie , C, Spijker , R, Deeks, J, Dinnes, J, Jenniskens, K, Korevaar , DA, Cohen , JF, Van den Bruel, A, Takwoingi, Y, van de Wijgert, J, Wang, J, McInnes, MDF, Cochrane COVID-19 Diagnostic Test Accuracy Group, & Agarwal, R 2020, 'Thoracic imaging tests for the diagnosis of COVID-19', *Cochrane Database of Systematic Reviews*, vol. 2020, no. 9, CD013639.
https://doi.org/10.1002/14651858.CD013639.pub3

Cochrane Library

Cochrane Database of Systematic Reviews

Thoracic imaging tests for the diagnosis of COVID-19 (Review)

Islam N, Salameh JP, Leeflang MMG, Hooft L, McGrath TA, van der Pol CB, Frank RA, Kazi S, Prager R, Hare SS, Dennie C, Spijker R, Deeks JJ, Dinnes J, Jenniskens K, Korevaar DA, Cohen JF, Van den Bruel A, Takwoingi Y, van de Wijgert J, Wang J, McInnes MDF, Cochrane COVID-19 Diagnostic Test Accuracy Group

Islam N, Salameh J-P, Leeflang MMG, Hooft L, McGrath TA, van der Pol CB, Frank RA, Kazi S, Prager R, Hare SS, Dennie C, Spijker R, Deeks JJ, Dinnes J, Jenniskens K, Korevaar DA, Cohen JF, Van den Bruel A, Takwoingi Y, van de Wijgert J, Wang J, McInnes MDF.
Thoracic imaging tests for the diagnosis of COVID-19.
Cochrane Database of Systematic Reviews 2020, Issue 11. Art. No.: CD013639.
DOI: 10.1002/14651858.CD013639.pub3.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Thoracic imaging tests for the diagnosis of COVID-19

Nayaar Islam[1,2], Jean-Paul Salameh[1], Mariska MG Leeflang[3], Lotty Hooft[4], Trevor A McGrath[1], Christian B van der Pol[5], Robert A Frank[1], Sakib Kazi[1], Ross Prager[6], Samanjit S Hare[7], Carole Dennie[1,8], René Spijker[4,9], Jonathan J Deeks[10,11], Jacqueline Dinnes[10,11], Kevin Jenniskens[4], Daniël A Korevaar[12], Jérémie F Cohen[13], Ann Van den Bruel[14], Yemisi Takwoingi[10,11], Janneke van de Wijgert[15,16], Junfeng Wang[17], Matthew DF McInnes[1,2], Cochrane COVID-19 Diagnostic Test Accuracy Group[10]

[1]Department of Radiology, University of Ottawa, Ottawa, Canada. [2]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada. [3]Epidemiology and Data Science, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands. [4]Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands. [5]Department of Radiology, McMaster University, Hamilton, Canada. [6]Department of Medicine, University of Ottawa, Ottawa, Canada. [7]Department of Radiology, Royal Free London NHS Trust, London, UK. [8]Department of Medical Imaging, The Ottawa Hospital, Ottawa, Canada. [9]Medical Library, Amsterdam UMC, University of Amsterdam, Amsterdam Public Health, Amsterdam, Netherlands. [10]NIHR Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK. [11]Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, Birmingham, UK. [12]Department of Respiratory Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands. [13]Obstetrical, Perinatal and Pediatric Epidemiology Research Team (EPOPé), Centre de Recherche Épidémiologie et Statistique Sorbonne Paris Cité (CRESS), Inserm UMR1153, Université de Paris, Paris, France. [14]Academic of Primary Care, KU Leuven, Leuven, Belgium. [15]Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands. [16]Institute of Infection, Veterinary, and Ecological Sciences, University of Liverpool, Liverpool, UK. [17]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands

Contact address: Matthew DF McInnes, mmcinnes@toh.ca.

Editorial group: Cochrane Infectious Diseases Group.
Publication status and date: New search for studies and content updated (conclusions changed), published in Issue 11, 2020.

Citation: Islam N, Salameh J-P, Leeflang MMG, Hooft L, McGrath TA, van der Pol CB, Frank RA, Kazi S, Prager R, Hare SS, Dennie C, Spijker R, Deeks JJ, Dinnes J, Jenniskens K, Korevaar DA, Cohen JF, Van den Bruel A, Takwoingi Y, van de Wijgert J, Wang J, McInnes MDF. Thoracic imaging tests for the diagnosis of COVID-19. *Cochrane Database of Systematic Reviews* 2020, Issue 11. Art. No.: CD013639. DOI: 10.1002/14651858.CD013639.pub3.

ABSTRACT

Background

The respiratory illness caused by SARS-CoV-2 infection continues to present diagnostic challenges. Early research showed thoracic (chest) imaging to be sensitive but not specific in the diagnosis of coronavirus disease 2019 (COVID-19). However, this is a rapidly developing field and these findings need to be re-evaluated in the light of new research. This is the first update of this 'living systematic review'. This update focuses on people suspected of having COVID-19 and excludes studies with only confirmed COVID-19 participants.

Objectives

To evaluate the diagnostic accuracy of thoracic imaging (computed tomography (CT), X-ray and ultrasound) in people with suspected COVID-19.

Search methods

We searched the COVID-19 Living Evidence Database from the University of Bern, the Cochrane COVID-19 Study Register, The Stephen B. Thacker CDC Library, and repositories of COVID-19 publications through to 22 June 2020. We did not apply any language restrictions.

Selection criteria

We included studies of all designs that recruited participants of any age group suspected to have COVID-19, and which reported estimates of test accuracy, or provided data from which estimates could be computed. When studies used a variety of reference standards, we retained the classification of participants as COVID-19 positive or negative as used in the study.

Data collection and analysis

We screened studies, extracted data, and assessed the risk of bias and applicability concerns using the QUADAS-2 domain-list independently, in duplicate. We categorised included studies into three groups based on classification of index test results: studies that reported specific criteria for index test positivity (group 1); studies that did not report specific criteria, but had the test reader(s) explicitly classify the imaging test result as either COVID-19 positive or negative (group 2); and studies that reported an overview of index test findings, without explicitly classifying the imaging test as either COVID-19 positive or negative (group 3). We presented the results of estimated sensitivity and specificity using paired forest plots, and summarised in tables. We used a bivariate meta-analysis model where appropriate. We presented uncertainty of the accuracy estimates using 95% confidence intervals (CIs).

Main results

We included 34 studies: 30 were cross-sectional studies with 8491 participants suspected of COVID-19, of which 4575 (54%) had a final diagnosis of COVID-19; four were case-control studies with 848 cases and controls in total, of which 464 (55%) had a final diagnosis of COVID-19. Chest CT was evaluated in 31 studies (8014 participants, 4224 (53%) cases), chest X-ray in three studies (1243 participants, 784 (63%) cases), and ultrasound of the lungs in one study (100 participants, 31 (31%) cases).

Twenty-six per cent (9/34) of all studies were available only as preprints. Nineteen studies were conducted in Asia, 10 in Europe, four in North America and one in Australia. Sixteen studies included only adults, 15 studies included both adults and children and one included only children. Two studies did not report the ages of participants. Twenty-four studies included inpatients, four studies included outpatients, while the remaining six studies were conducted in unclear settings. The majority of included studies had a high or unclear risk of bias with respect to participant selection, index test, reference standard, and participant flow.

For chest CT in suspected COVID-19 participants (31 studies, 8014 participants, 4224 (53%) cases) the sensitivity ranged from 57.4% to 100%, and specificity ranged from 0% to 96.0%. The pooled sensitivity of chest CT in suspected COVID-19 participants was 89.9% (95% CI 85.7 to 92.9) and the pooled specificity was 61.1% (95% CI 42.3 to 77.1).

Sensitivity analyses showed that when the studies from China were excluded, the studies from other countries demonstrated higher specificity compared to the overall included studies. When studies that did not classify index tests as positive or negative for COVID-19 (group 3) were excluded, the remaining studies (groups 1 and 2) demonstrated higher specificity compared to the overall included studies. Sensitivity analyses limited to cross-sectional studies, or studies where at least two reverse transcriptase polymerase chain reaction (RT-PCR) tests were conducted if the first was negative, did not substantively alter the accuracy estimates. We did not identify publication status as a source of heterogeneity.

For chest X-ray in suspected COVID-19 participants (3 studies, 1243 participants, 784 (63%) cases) the sensitivity ranged from 56.9% to 89.0% and specificity from 11.1% to 88.9%. The sensitivity and specificity of ultrasound of the lungs in suspected COVID-19 participants (1 study, 100 participants, 31 (31%) cases) were 96.8% and 62.3%, respectively. We could not perform a meta-analysis for chest X-ray or ultrasound due to the limited number of included studies.

Authors' conclusions

Our findings indicate that chest CT is sensitive and moderately specific for the diagnosis of COVID-19 in suspected patients, meaning that CT may have limited capability in differentiating SARS-CoV-2 infection from other causes of respiratory illness. However, we are limited in our confidence in these results due to the poor study quality and the heterogeneity of included studies. Because of limited data, accuracy estimates of chest X-ray and ultrasound of the lungs for the diagnosis of suspected COVID-19 cases should be carefully interpreted.

Future diagnostic accuracy studies should pre-define positive imaging findings, include direct comparisons of the various modalities of interest on the same participant population, and implement improved reporting practices. Planned updates of this review will aim to: increase precision around the accuracy estimates for chest CT (ideally with low risk of bias studies); obtain further data to inform accuracy of chest X-rays and ultrasound; and obtain data to further fulfil secondary objectives (e.g. 'threshold' effects, comparing accuracy estimates across different imaging modalities) to inform the utility of imaging along different diagnostic pathways.

PLAIN LANGUAGE SUMMARY

How accurate is chest imaging for diagnosing COVID-19?

Why is this question important?

People with suspected COVID-19 need to know quickly whether they are infected, so they can receive appropriate treatment, self-isolate, and inform close contacts.

Currently, formal diagnosis of COVID-19 requires a laboratory test (RT-PCR) of nose and throat samples. RT-PCR requires specialist equipment and takes at least 24 hours to produce a result. It is not completely accurate, and may require a second RT-PCR or a different test to confirm diagnosis.

COVID-19 is a respiratory disease. Clinicians may use chest imaging to diagnose people who have COVID-19 symptoms, while awaiting RT-PCR results or when RT-PCR results are negative, and the person has COVID-19 symptoms.

What did we want to find out?

We wanted to know whether chest imaging is accurate enough to diagnose COVID-19 in people with suspected infection. This is the first update of this review; in it we included studies in people with suspected COVID-19 only; we excluded studies in people with confirmed COVID-19.

The evidence is up to date to 22 June 2020.

What are chest imaging tests?

X-rays or scans produce an image of the organs and structures in the chest.

- X-rays (radiography) use radiation to produce a 2-D image. Usually done in hospitals, using fixed equipment by a radiographer, they can also be done on portable machines.

- Computed tomography (CT) scans use a computer to merge 2-D X-ray images and convert them to a 3-D image. They require highly specialised equipment and are done in hospital by a specialist radiographer.

- Ultrasound scans use high-frequency sound waves to produce an image. They can be done in hospital or other healthcare settings, such as a doctor's office.

What did we do?

We searched for studies that assessed the accuracy of chest imaging to diagnose COVID-19 in people with suspected COVID-19. Studies could be of any design and take place anywhere.

What did we find?

We found 34 studies with 9339 people. All the studies confirmed SARS-CoV-2 infection using RT-PCR alone or RT-PCR with another test.

Most studies (31 studies; 8014 participants) evaluated chest CT; three evaluated chest X-rays (1243 participants) and one evaluated lung ultrasound (100 participants). Nineteen studies took place in Asia, 10 in Europe, four in North America and one in Australia. Participants were hospital inpatients (24 studies), and outpatients (4 studies); the setting was unclear in six studies.

Where four or more studies evaluated a particular type of chest imaging, we pooled their results and analysed them together.

Chest CT

Pooled results showed that chest CT correctly diagnosed COVID-19 in 89.9% of people who had COVID-19. However, it incorrectly identified COVID-19 in 38% of people who did not have COVID-19.

Chest X-ray

Correct diagnosis of COVID-19 with chest X-rays ranged from 57% to 89%. However, incorrect diagnosis of COVID-19 in people who did not have COVID-19 ranged from 11% to 89%.

Lung ultrasound

Lung ultrasound correctly diagnosed COVID-19 in 96% of people with COVID-19. However, it incorrectly diagnosed COVID-19 in 38% of people who did not have COVID-19.

How reliable are the results?

The studies differed from each other and used different methods to report their results. About a quarter of the studies were published as preprints, which do not undergo the same rigorous checks as published studies. We cannot draw confident conclusions based on results from studies in this review.

What does this mean?

The evidence suggests that chest CT is better at ruling out COVID-19 infection than distinguishing it from other respiratory problems. So, its usefulness may be limited to excluding COVID-19 infection rather than distinguishing it from other causes of lung infection.

Chest CT accuracy has improved since our first review, perhaps because radiologists now use better definitions of a positive diagnosis. The stage of the pandemic may also have an effect – with later studies building on knowledge and experience gained earlier.

We plan to update this review as more evidence becomes available. Future studies should predefine what a positive test is, and compare different types of imaging tests on similar groups of people.

SUMMARY OF FINDINGS

Summary of findings 1. 'Summary of findings' table

Question	What is the diagnostic accuracy of chest imaging (computed tomography (CT), chest X-ray and ultrasound) in the evaluation of people suspected to have COVID-19?
Population	Children or adults suspected to have COVID-19
Index test	Chest imaging tests used for the diagnosis of COVID-19, including: • Chest CT • Chest X-rays • Ultrasound of the lungs
Target condition	Detection of current SARS-CoV-2 infection
Reference standard	A positive diagnosis for COVID-19 by one or a combination of the following. • A positive RT-PCR test for SARS-CoV-2 infection, from any manufacturer in any country, from any source, including nasopharyngeal swabs or aspirates, oropharyngeal swabs, bronchoalveolar lavage fluid (BALF), sputum, saliva, serum, urine, rectal or faecal samples. • Positive on WHO criteria for COVID-19 which includes some testing RT-PCR negative. • Positive on China CDC criteria for COVID-19 which includes some testing RT-PCR negative. • Positive serology in addition to consistent symptomatology. • Positive on study-specific list of criteria for COVID-19 which includes some testing RT-PCR negative. • Other criteria (symptoms, imaging findings, other tests, infected contacts). A negative diagnosis for COVID-19 by one or a combination of the following: • COVID suspects with negative RT-PCR test results, whether tested once or more than once. • Pre-pandemic controls (healthy or with another disease). • Current healthy or with another disease (no RT-PCR test).
Limitations in the evidence	
Risk of bias	Participant selection: high in 10 (29%) studies and unclear in 17 (50%) studies Application of index tests – chest CT: high in 7/31 (23%) studies and unclear in 17/31 (55%) studies Application of index tests – chest X-ray: unclear in 3/3 (100%) studies Application of index tests – ultrasound of the lungs: unclear in 1/1 study Flow and timing: high in 7 (21%) studies and unclear in 18 (53%) studies
Concerns about applicability of the evidence	Participants: high in 3 (9%) studies and unclear in 2 (6%) studies Index test – chest CT: high in 3/31 (10%) studies and unclear in 1/31 (3%) study Index test – chest X-ray: low in 3/3 (100%) studies Index test – ultrasound of the lungs: unclear in 1/1 (100%) study Reference standard: unclear in 2 (6%) studies
Findings	

- We included 34 studies (9339 participants total), which consisted of 30 cross-sectional studies with 8491 participants suspected of COVID-19 (4575 (54%) cases), and 4 case-control studies with 848 cases and controls in total (464 (55%) cases).
- Of our 34 included studies, four studies with 1349 participants (595 (44%) cases) were categorised as group 1 (studies that report specific criteria for index test positivity), 22 studies with 7075 participants (3942 (56%) cases) were categorised as group 2 (studies that do not report specific criteria for index test positivity, but have the test reader(s) explicitly classify the imaging test as either COVID-19 positive or negative), and eight studies with 915 participants (486 (53%) cases) were categorised as group 3 (studies that report an overview of index test findings in participants with and without the target condition, without explicitly classifying the imaging test as either COVID-19 positive or negative).
- Most studies (n = 31) evaluated the accuracy of chest CT scans.
- Chest CT was sensitive and moderately specific in the diagnosis of COVID-19 in suspected cases.
- Sensitivity analyses showed that studies conducted in countries other than China, as well as studies categorised into groups 1 and 2 demonstrated higher specificity compared to the overall included studies, while cross-sectional studies, as well as studies that implemented RT-PCR testing at least twice for participants with initial negative results had a minimal effect on our findings.
- Publication status was not identified as sources of heterogeneity.
- The low number of studies, the lack of transparent reporting, and the concerns of bias and applicability prevented comparisons between different imaging modalities.
- Given various prevalence settings, predicted outcomes for the number of individuals receiving a false positive result or a false negative (missed) result per 1000 people undergoing chest CT are outlined as follows.

Predicted outcomes per 1000 people undergoing chest CT				
Prevalence of COVID-19	Positive CT result, n (95% CI)	False positive CT result, n (95% CI)	Negative CT result, n (95% CI)	False negative CT result, n (95% CI)
50%	644 (579 to 717)	195 (116 to 289)	356 (283 to 421)	51 (36 to 72)
20%	491 (369 to 633)	311 (183 to 462)	509 (367 to 631)	20 (14 to 29)
5%	416 (264 to 591)	370 (218 to 548)	585 (409 to 736)	5 (4 to 7)
Quantity of evidence for participants suspected of having COVID-19				
Imaging modality		Sensitivity (95% CI)	Specificity (95% CI)	Number of participants (cases)
Chest CT		89.9% (85.7 to 92.9)	61.1% (42.3 to 77.1)	8014 (4224)
Chest X-ray*		-	-	1243 (784)
Ultrasound of the lungs†		-	-	100 (31)

*The three studies that evaluated chest X-ray demonstrated ranges of sensitivity and specificity of 56.9% to 89.0% and 11.1% to 88.9%, respectively. Pooling was not feasible due to lack of available data.
†The one study that evaluated ultrasound of the lungs demonstrated a sensitivity of 96.8% and a specificity of 62.3%.

BACKGROUND

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection and resulting coronavirus disease 2019 (COVID-19) pandemic continue to present diagnostic evaluation challenges. While the World Health Organization (WHO) reports laboratory confirmation of COVID-19 infection, such as a positive reverse transcriptase polymerase chain reaction (RT-PCR) result as the standard for diagnosing COVID-19, the value of imaging tests in the diagnostic pathway remain undefined (WHO 2020). Research on the role of imaging in COVID-19 patients is evolving and more refined assessment methods for imaging tests, such the COVID-19 Reporting and Data System (CO-RADS), are being investigated (Prokop 2020).

Decisions about patient and isolation pathways for COVID-19 vary according to health services and settings, available resources, and outbreaks in different settings. They will change over time, as accurate tests, effective treatments, and vaccines are identified. The decision points between these pathways vary, but all include points at which knowledge of the accuracy of diagnostic information is needed to inform medical decisions.

Therefore, it is essential to understand the accuracy of tests and diagnostic features to develop effective diagnostic and management pathways for different settings. This supports strategies aiming to identify those who are infected, and consequently the management of patients either through isolation precautions, contact tracing, quarantine, hospital admission or admission to a specialised facility, admission to the intensive care unit, or initiation of specific therapies, and implementation of mitigation strategies to limit the spread of the disease. This review from the suite of Cochrane 'living systematic reviews' summarises evidence on the accuracy of different imaging tests and diagnostic features in participants regardless of their symptoms, grouped according to the research questions and settings that we are aware of. Estimates of accuracy from this review will help inform diagnostic, screening, isolation, and patient management decisions. We have included an explanation of terminology and acronyms in Appendix 1.

Target condition being diagnosed

The target condition being evaluated is COVID-19 disease, the disease caused by infection with SARS-CoV-2. People infected with SARS-CoV-2 can be asymptomatic; these people are not considered to have COVID-19 and thus not within the scope of this review. People with COVID-19 can have a wide variety of symptoms, including fever, cough and aches, as well as lethargy without difficulty breathing at rest, or lethargy with shortness of breath and increased respiratory rate, potentially requiring supplemental oxygen, and in severe cases, requiring mechanical ventilation due to severe hypoxaemic respiratory failure or acute respiratory distress syndrome. Furthermore, in people diagnosed with a pulmonary condition (e.g. pulmonary embolism), symptoms could either be the explanation for the respiratory symptoms, or could be indicative of a condition that is present in addition to COVID-19. In this review, we focused on persons suspected to have COVID-19 who had one or more respiratory symptoms or signs, who had thoracic imaging as part of their evaluation or care.

Index test(s)

Chest computed tomography (CT)

Chest CT refers to the acquisition of images of the chest using computed tomography. Typical imaging protocols would not use intravenous (IV) contrast; however, in this review we considered all variations of imaging protocols with the exception of studies specifically targeted at evaluating the coronary arteries or the heart, which did not include the entire lungs in the field of view. This includes, but is not limited to, non-contrast chest CT, low-dose chest CT (with or without contrast), high-resolution chest CT, and chest CT with IV contrast (routine or pulmonary angiogram).

Chest radiographs/chest X-rays

Chest radiography refers to the evaluation of the lungs using X-rays. This often involves two orthogonal views, posterior-anterior (PA) and lateral, but may be done by a portable machine and only acquire an anterior-posterior (AP) view. In this review, we considered any and all variations of chest radiography protocols that evaluated the lungs. We did not include protocols that did not include the entire thorax and were done for reasons other than for assessment of pulmonary status (e.g. assessment of feeding tube position, which typically only includes the lower thorax, or dedicated evaluation of the ribs).

Ultrasound of the lungs

Ultrasound of the lungs refers to any ultrasound of the thorax done with the intention of evaluating the status of the lungs. This includes, but is not limited to, point-of-care ultrasound (POCUS), done at the bedside by a physician, as well as what is often termed 'consultative' ultrasound, which is done by a technologist and subsequently interpreted by a physician (typically a radiologist). We considered all possible technical parameters (e.g. type of probe, transducer frequency, use of contrast). This did not include ultrasound done with the intended purpose of evaluating only the heart or vessels of the chest.

Clinical pathway

At present, the optimal diagnostic pathway and the role of thoracic imaging for identifying people with COVID-19 is unclear. Compared to RT-PCR testing, a potential major advantage of thoracic imaging is that results are available faster and that it provides a better insight into the status of the lungs. However, chest CT and ultrasound of the lungs are typically only available in secondary and tertiary healthcare settings, and availability varies across these settings.

Role of index test(s)

1. Thoracic imaging may play an integral role in 'ruling out' COVID-19 pneumonia when RT-PCR is unavailable, pending or negative, or when clinical suspicion is 'low' based on other signs, symptoms and routine laboratory tests. Role of test: triage for RT-PCR, to make decisions about performing or not performing RT-PCR or other diagnostic tests.
2. Rapid testing - thoracic imaging is used to rule in or rule out COVID-19 when results from other tests (e.g. RT-PCR) are not available in a timely manner.
3. Concurrent/combination testing with other diagnostic tests (as part of a pair or group of tests) to improve the accuracy of diagnosis. For example, thoracic imaging could be used to

identify false negatives of other tests (e.g. RT-PCR), and to improve the overall accuracy of the testing strategy.

Several diagnostic pathways have been proposed that provide guidance for physicians to identify people with COVID-19. The order and components of these pathways differ with varying dependence on pre-test probability, physical examination, laboratory tests and findings based on RT-PCR results and availability. However, some professional organisations recommend imaging for patients with moderate or severe features of COVID-19 (Rubin 2020). In some hospitals, the results of low-dose chest CT are one of the many parameters (among molecular test results, routine laboratory results and clinical signs and symptoms) used to categorise patients as low risk, moderate to high risk, and proven COVID-19 cases.

Given the rapid progression of COVID-19 and the constantly evolving evidence base, the diagnostic accuracy to inform the utility of thoracic imaging in these pathways is difficult to estimate. This 'living' systematic review aims to identify data regarding the diagnostic accuracy of thoracic imaging in people with suspected COVID-19. This represents our first update of this 'living' systematic review (Salameh 2020a).

Alternative test(s)

Other Cochrane diagnostic test accuracy (DTA) Reviews in the suite of reviews are addressing the following tests.

1. Signs and symptoms, which will be mainly used in primary care, including when presenting at the emergency department (Struyf 2020)
2. Routine laboratory testing, such as for C-reactive protein (CRP) and procalcitonin (PCT) (Stegeman 2020)
3. Antibody tests (Deeks 2020)
4. Laboratory-independent point-of-care and near-patient molecular and antigen tests (Dinnes 2020)
5. Molecular laboratory tests

Summary of previous version of review

In our initial review, studies with confirmed cases only reported high pooled sensitivity for chest CT and X-ray 93.1% (95% CI 90.2 to 95.0) and 82.1% (95% CI 62.5 to 92.7), respectively (Salameh 2020a). Two studies that evaluated ultrasound of the lungs in confirmed cases only both reported zero false negatives. Subgroup analyses of these studies stratified by publication status (preprints versus published studies) showed comparable diagnostic accuracy between estimates of the subgroups – the pooled sensitivity estimates for thoracic CT were 93.0% (95% CI 86.2 to 96.6) for preprints versus 93.0% (95% CI 89.9 to 95.3) for the published studies. Other subgroup analyses were not conducted because of an insufficient number of included studies.

Studies assessing chest CT in suspected participants demonstrated a sensitivity of 86.2% (95% CI 71.9 to 93.8) but a low specificity of 18.1% (95% CI 3.71 to 55.8) in the diagnosis of COVID-19. This indicates a lack of discrimination, as the chances of getting a positive chest CT result are 86% in patients with a SARS-CoV-2 infection and 82% in patients without. Furthermore, a sensitivity of < 90% may not be appropriate for the evaluation of patients with suspected COVID-19 given the risk associated with false negative diagnosis as individuals with these results may relax their measures to limit transmission of SARS-CoV-2 within their environment. We did not assess accuracy estimates for ultrasound of the lungs or chest X-ray in suspected participants as these data were not available.

Compared to the previous version of this review, this update focuses on people suspected of having COVID-19 and excludes studies evaluating only confirmed cases of COVID-19.

Changes in the evidence base since the previous version

Evolving research on imaging tests in COVID-19 patients includes the use of formal scoring systems to evaluate imaging tests, such as CO-RADS, which offers the potential for improved specificity (Prokop 2020). Previous studies either did not specify what criteria they used for index test positivity, or used 'any abnormality' to define index test positive. The value of formal scoring systems will be explored in this update, as well as in future updates of this review.

OBJECTIVES

The objective is to evaluate the diagnostic accuracy of thoracic imaging (computed tomography (CT), chest X-ray and ultrasound) in the evaluation of people suspected to have COVID-19.

METHODS

Criteria for considering studies for this review

Types of studies

The eligibility criteria were kept broad to be able to include all patient groups and all variations of a test.

We included studies of all designs that produced estimates of test accuracy or provided data from which estimates could be computed, for the primary objective. In this review, we categorised our included two types of study designs.

1. Cross-sectional studies including participants suspected to have the target condition
2. Case-control studies including two independently recruited groups of cases with the target condition and controls who are currently healthy or have another disease

This update of the review only included studies focusing on patients with suspected COVID-19 (i.e. both sensitivity and specificity were estimated). This represents a modification from the study protocol and the initial version of this review; this change was made with approval by the Cochrane COVID-19 Diagnostic Test Accuracy Group, as well as all of the study authors.

We carefully considered the limitations of different study designs in the quality assessment, the analysis, and the interpretation of findings.

Inclusion criteria

We included studies if the following criteria were met.

1. They included patients suspected of COVID-19 as outlined in the 'Target conditions' section. There were no age or gender restrictions.

2. The index test was chest CT, X-ray, or ultrasound, meeting the criteria described in the 'Index tests' section.
3. The index test was interpreted by humans, and not an algorithm (machine learning/artificial intelligence (AI)).
4. A reference standard for a positive and negative classification of target condition status was applied as outlined in the 'Reference standards' section.
5. Data were available to extract 2x2 data (true positive (TP), true negative (TN), false positive (FP), false negative (FN)). If data were not available, we contacted study authors for additional data if the study met the primary objective only (2x2 data).
6. They included 10 or more patients who underwent the index test and reference standard.

Participants

Our focus was on studies that recruited participants suspected to have COVID-19. We included all age groups.

Index tests

Chest CT, or chest X-ray, or ultrasound of the lungs. The roles of the test can be a replacement of polymerase chain reaction (RT-PCR), add-on test, triage test, rapid testing, or used concurrently with other diagnostic tests.

Definitions of imaging test positivity

Since COVID-19 is such a new disease, and the imaging findings were unknown until recently, there is considerable heterogeneity and change in the definitions used for positivity. Some groups have used constellations of specific findings (such as multiple peripheral ground-glass opacities on CT), some have used an approach in which they consider the combined effect of specific findings (a 'gestalt' approach), and some have used formal classification systems, such as COVID-19 Reporting and Data System (CO-RADS) (a 5-point scale ranging from 1 (i.e. very low suspicion for pulmonary involvement of COVID-19) to 5 (i.e. very high suspicion for pulmonary involvement of COVID-19)) (Prokop 2020). As such, we did not limit ourselves to a predefined threshold for, or definition of positivity. Instead, we extracted the definition for positivity used in each study, and the constellation of imaging features used to inform this definition. This offers an opportunity to determine if the definition of positivity contributes to variability in accuracy.

Target conditions

As explained above, our target condition is COVID-19. However, we included all studies reporting data on COVID-19 or COVID-19 pneumonia that might provide data relevant to our objective.

Reference standards

A positive diagnosis for COVID-19 by one or a combination of the following:

1. a positive RT-PCR test for SARS-CoV-2 infection, from any manufacturer in any country, and from any sample type, including nasopharyngeal swabs or aspirates, oropharyngeal swabs, bronchoalveolar lavage fluid (BALF), sputum, saliva, serum, urine, rectal or faecal samples;
2. positive on WHO criteria for COVID-19;
3. positive on China CDC criteria for COVID-19;
4. positive serology for SARS-CoV-2 antibodies in addition to consistent symptomatology;
5. positive on study-specific list of criteria for COVID-19 which includes:
   a. other criteria (symptoms, imaging findings, other tests, infected contacts).

A negative diagnosis for COVID-19 by one or a combination of the following:

1. COVID-19 suspects with negative RT-PCR test results, whether tested once or more than once;
2. pre-pandemic controls (healthy or diseased);
3. current healthy or with another disease (no RT-PCR test).

When studies used a variety of reference standards, we included all of them. In the assessment of methodological quality, we judged how likely each reference standard definition is to correctly classify individuals. All reference standards are likely to be imperfect in some way; details of reference standard evaluation are provided in the 'Risk of bias' tool (Appendix 2). We used a consensus process to agree on the classification of the reference standard as to what we regarded as good, moderate and poor. 'Good' reference standards need to have very little chance of misclassification; 'moderate', a small but acceptable risk; and 'poor', a larger and probably unacceptable risk.

Search methods for identification of studies

Electronic searches

We used three different sources for our electronic searches through 22 June 2020, which were devised with the help of an experienced Cochrane Information Specialist with DTA expertise (RSp). These searches aimed to identify all articles related to COVID-19 and SARS-CoV-2 and were not restricted to those evaluating imaging tests. Thus, the searches used no terms that specifically focused on an index test, diagnostic accuracy or study methodology.

Due to the increased volume of published and preprint articles, we used artificial intelligence text analysis from 25 May 2020 and onwards to conduct an initial classification of documents, based on their title and abstract information, for relevant and irrelevant documents. Appendix 3.

1. Living search from the University of Bern

We used the COVID-19 living search results of the Institute of Social and Preventive Medicine (ISPM) at the University of Bern. This search includes PubMed, Embase and preprints indexed in bioRxiv and medRxiv databases. The strategies as described on the ISPM website (ispmbern.github.io/covid-19), are shown in Appendix 4.

2. Cochrane COVID-19 Study Register searches

We also included searches undertaken by Cochrane to develop the Cochrane COVID-19 Study Register. These include searches of trials registers at ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform (WHO ICTRP), as well as PubMed (see Appendix 4 for details). Search strategies were designed for maximum sensitivity, to retrieve all human studies on COVID-19. We did not apply any language limits.

3. The Stephen B. Thacker CDC Library, COVID-19 Research Articles Downloadable Database

We included Embase records within the CDC library on COVID-19 research articles database (see Appendix 4 for details) and deduplicated these against the Cochrane COVID-19 Study Register.

Searching other resources

We checked repositories of COVID-19 publications against these search results including the following.

1. EPPI centre eppi.ioe.ac.uk/COVID19_MAP/covid_map_v4.html.
2. The Norwegian Institute of Public Health 'NIPH systematic and living map on COVID-19 evidence' www.nornesk.no/forskningskart/NIPH_diagnosisMap.html.
3. From these websites we searched company and product websites for studies about test accuracy.
4. We contacted companies to ask for further information about studies.
5. We also contacted research groups that we were made aware of who are completing test evaluations (e.g. UK Public Health England-funded studies, Foundation for Innovative New Diagnostics (FIND) studies).

Data collection and analysis

Selection of studies

The review authors screened studies independently, in duplicate. A third, experienced review author resolved disagreements about initial title and abstract screening. We resolved disagreements about eligibility assessments through discussion between three review authors.

Data extraction and management

The review authors performed data extraction independently, in duplicate. Three review authors discussed any disagreements to resolve them.

For each study, we extracted 2x2 contingency tables of the number of true positives, false positives, false negatives and true negatives. If a study reported accuracy data for more than one index test reader, we took the average of the data from all readers to compute the average 2x2 contingency table (McGrath 2017). If a study reported accuracy data for multiple thresholds of index test positivity, we extracted the 2x2 contingency table corresponding to the threshold producing the highest Youden's Index (YI) (YI = sensitivity + specificity – 1). If a study reported accuracy data for various CT findings or combinations of CT findings, we extracted the 2x2 contingency table corresponding to the CT finding, or the combination of findings producing the highest YI. If a study reported accuracy data for both an AI algorithm and one or more radiologists, we extracted only the 2x2 contingency table corresponding to the radiologist accuracy data. If a study used multiple reference standards, but 2x2 contingency tables including RT-PCR as the only reference standard could be determined, we extracted and analysed these data. If a study graphically displayed accuracy data and did not report the raw data values, we first contacted the authors, but if no response was received, we extracted the 2x2 contingency table by estimating accuracy data from the graphs.

For studies that used the 5-point CO-RADS classification scale and did not specify a threshold for disease positivity, we extracted 2x2 contingency tables for CO-RADS thresholds 4 and 5 as defining test positivity. When these studies were included in meta-analyses, the 2x2 contingency table corresponding to the highest Youden's index was used. When reporting ranges of accuracy estimates for CO-RADS studies, common thresholds (i.e. thresholds 4 and 5) were combined.

In addition, we extracted the following items.

1. Study setting (including country), age of study participants, study dates, disease prevalence at the time of acquisition (as reported in the study), number of participants, participant symptoms, number of imaging studies (and if more than one study was done per participant), participant outcomes and other relevant participant demographic parameters.
2. Study design.
3. Imaging timing relative to disease course.
4. CT, chest X-ray and ultrasound findings.
5. Criteria for 'positive' diagnosis of COVID-19 on imaging.
6. Index test technical parameters.
7. Reference standard results and details. If RT-PCR was performed, timing of test, number of tests and method of acquisition (or similar details regarding other reference standards used).
8. Details regarding interpretation of the index test (level of training, number of readers, the inter-observer variability).
9. The number of true positives, false positives, false negatives and true negatives or summary statistics from which they can be computed.

Categorisation of included studies

We categorised included studies into three groups, based on study design with respect to classification of index test results.

1. Group 1: studies that report specific criteria for index test positivity (e.g. CO-RADS threshold).
2. Group 2: studies that do not report specific criteria for index test positivity, but have the test reader(s) explicitly classify the imaging test as either COVID-19 positive or negative.
3. Group 3: studies that report an overview of index test findings (e.g. ground glass, consolidation, pleural effusion) in participants with and without the target condition based on reference standard results, without explicitly classifying the imaging test as either COVID-19 positive or negative.

This categorisation also allowed us to differentiate studies with a clear intent to diagnose COVID-19 on imaging tests (i.e. groups 1 and 2) from studies reporting findings of imaging tests without making a diagnosis (i.e. group 3). In this review, we conducted sensitivity analysis excluding group 3 studies.

Assessment of methodological quality

The review authors assessed the risk of bias and applicability concerns independently, in duplicate, using the QUADAS-2 domain-list. Three review authors resolved any disagreements through discussion. See Appendix 2 for an explanation of the operationalisation of the four QUADAS-2 domains – participant selection, index test(s), reference standard(s), flow and timing.

Statistical analysis and data synthesis

We presented estimates of sensitivity and specificity using paired forest plots, and summarised results in tables, as appropriate. We analysed the data on a participant level, not a lesion or lung segment level, since this is what determines care.

We used a bivariate model for meta-analyses, taking into account the within- and between-study variance, and the correlation between sensitivity and specificity across studies (Chu 2006; Reitsma 2005). We also performed sensitivity analyses by limiting inclusion in the meta-analysis to: studies conducted in countries other than China, cross-sectional studies, and studies that completed RT-PCR testing at least twice for participants with initial negative results. We undertook meta-analyses using metandi and meta-regression using meqrlogit in STATA (Harbord 2009; StataCorp 2019).

We did not undertake comparisons of test accuracy across different imaging modalities due to limited data, as four or more studies for a given modality were required to perform the meta-analysis and only the group of chest CT studies met this threshold. However, in future updates, as more data become available, we will perform test comparisons using hierarchical meta-regression. We will consider using all available data regardless of whether or not studies have compared imaging modalities head-to-head in the same study population (i.e. indirect comparison), as well as restricting test comparisons to only comparative studies (i.e. direct comparisons).

Ranges of sensitivities and specificities were estimated for studies that used a common threshold for test positivity (i.e. CO-RADS thresholds 4 and 5).

Investigations of heterogeneity

We investigated heterogeneity by visual inspection of paired forest plots and SROC plots. We evaluated the impact of publication status (preprint versus published) on accuracy estimates using meta-regression for the variable separately by adding the covariate term to a bivariate model. Subgroup analyses were limited to variables of interest which consisted of subgroups with five or more studies, as this threshold was required to ensure stability of the bivariate model.

Assessment of reporting bias

For this review, we did not undertake tests for publication bias and made no formal assessment of reporting bias.

Summary of findings

We provided a summary of the key findings of this review in a 'Summary of findings 1' table indicating the strength of evidence for each finding and emphasising the main gaps in our current level of available evidence.

Updating

The prior version of this review contained studies up to 5 May 2020. This updated review contains the results of an updated search performed on 22 June 2020. With the substantial number of studies published since 22 June 2020, we plan to update this review shortly and have already performed searches and completed abstract screening for the next update up until 30 September 2020.

RESULTS

Results of the search

We screened a total of 668 unique references (published or preprint studies) for inclusion; this is inclusive of the 561 references we screened in our initial review. Of the 273 records selected for full-text assessment, we included 34 studies in this review (13 of these 34 included studies were previously included in our initial review). Refer to Figure 1 for the PRISMA flow diagram of search and inclusion results (Salameh 2020b; Moher 2009). Exclusions were mainly due to ineligible study outcomes (n = 8), ineligible study design (n = 6), or ineligible patient populations (n = 6); see Figure 1.

Figure 1. Study flow diagram.

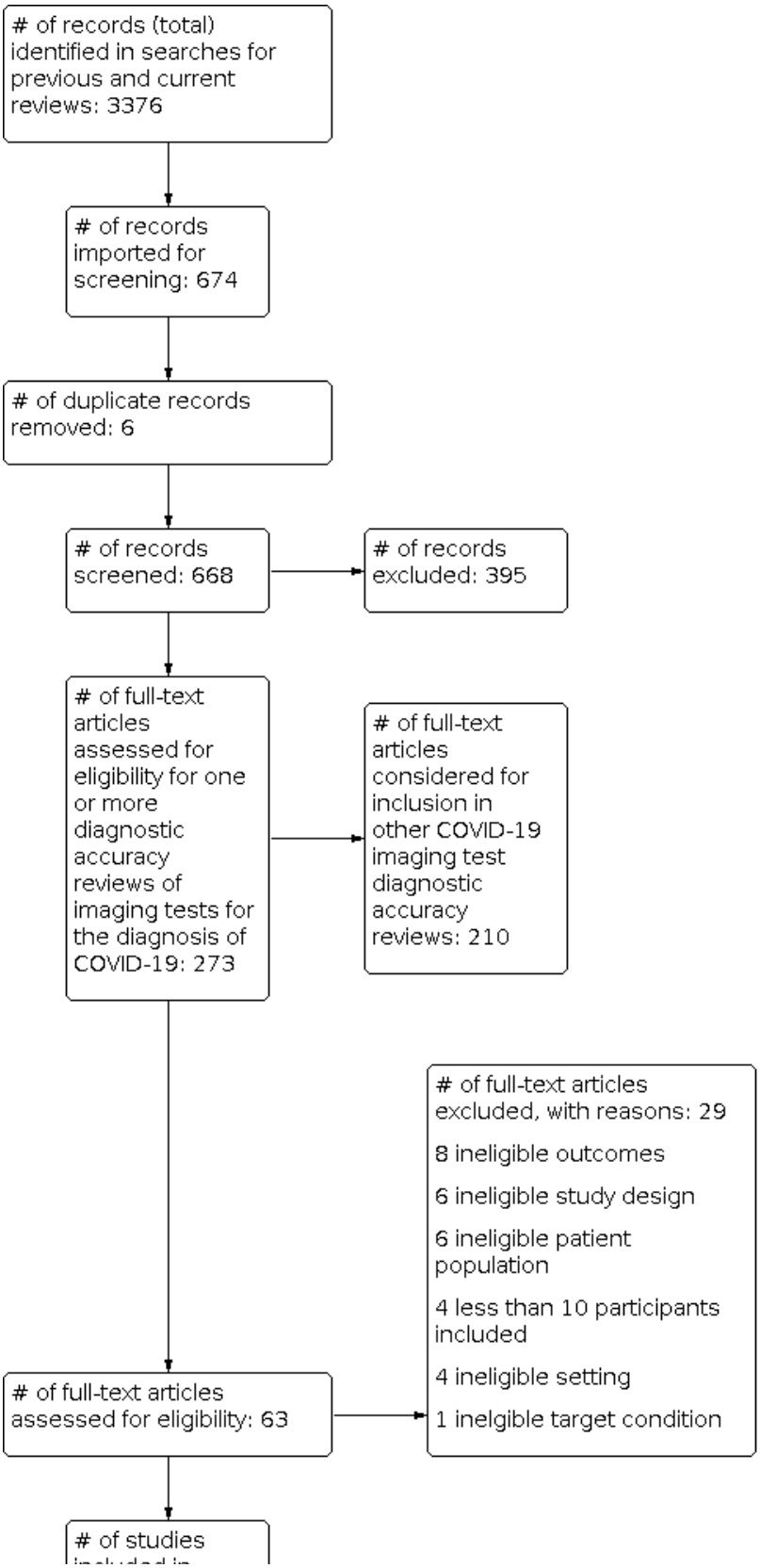
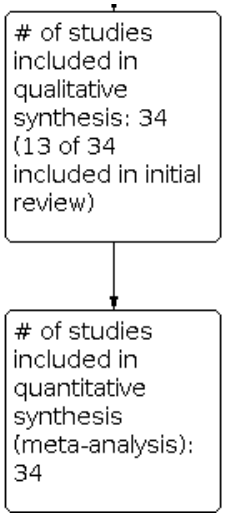


Figure 1. (Continued)



Description of included studies

We categorised the 34 included studies into two study designs. In the first category, we included 30 cross-sectional studies (26 CT, two X-ray, one both CT and X-ray, and one ultrasound) with 8491 participants suspected of having COVID-19, of which 4575 (54%) had a final diagnosis of COVID-19. In the second category, we included four case-control studies (all CT) with 848 cases and controls in total, of which 464 (55%) had a final diagnosis of COVID-19, with cases being patients with confirmed COVID-19 by methods other than thoracic imaging and controls being patients confirmed to not have COVID-19 by methods other than thoracic imaging.

We also categorised the 34 included studies into the following groups based on study design, with respect to classification index test results. Group 1 (studies that report specific criteria for index test positivity, such as a CO-RADS threshold) included four studies (all CT) with 1349 participants, of which 595 (44%) had a final diagnosis of COVID-19. Group 2 (studies that do not report specific criteria for index test positivity, but have the test reader(s) explicitly classify the imaging test as either COVID-19 positive or negative) included 22 studies (19 CT, two X-ray, and one ultrasound) with 7075 participants, of which 3942 (56%) had a final diagnosis of COVID-19. Group 3 (studies that report an overview of index test findings (e.g. ground-glass, consolidation, pleural effusion) in participants with and without the target condition based on reference standard results, without explicitly classifying the imaging test as either COVID-19 positive or negative) included eight studies (seven CT and one both CT and X-ray) with 915 participants, of which 486 (53%) had a final diagnosis of COVID-19.

The median sample size was 160.5 (interquartile range (IQR) 83.5 to 315.8). Nineteen studies were conducted in Asia (China (n = 18) and Japan (n = 1)), 10 in Europe (Italy (n = 3), Belgium (n = 2), the Netherlands (n = 2), France (n = 2) and Turkey (n =1)), and the remaining studies were conducted in North America (USA; n = 4) and in Australia (n = 1). The level of training of readers was not clearly reported in 10/34 studies (29%), while 23/34 studies (68%) reported that radiologists performed the reading, and 1/33 studies (3%) was completed by radiology residents. Technical parameters regarding the protocol of chest CT used were not clearly reported in 20/31 (65%) studies. Non-contrast CT was used in 4/31 (13%) studies, high-resolution chest CT was used in 4/31 (13%) studies, low-dose CT with or without contrast was used in 2/31 (6%) studies and CT with intravenous (IV) contrast was used in 1/31 (3%) study. Manuscripts of 9/34 (26%) of the studies were published as preprints at the time of the search. We updated the publication status of all the preprint studies previously included in our initial review (n = 6) as of 1 October 2020, and while one of these studies was published since then, there were no changes to the data between the preprint and published versions. Characteristics of the included studies are summarised in Table 1, and outlined in detail in the Characteristics of included studies.

Participant characteristics

Sixteen studies included only adult participants (16 years old and over), one study included only children, 15 studies included both children and adults (although in most cases, only a minority of included patients were children), and the remaining two studies did not clearly report the age range of participants. RT-PCR was used as the reference standard for the diagnosis of COVID-19 in all studies, with 29 studies using only RT-PCR as the reference standard and five studies using a combination of RT-PCR and other criteria (clinical symptoms and infected household contact (n = 2) clinical symptoms and imaging tests (n = 1), clinical symptoms (n = 1), and laboratory tests (n = 1)) as the reference standard. With respect to RT-PCR testing, three studies tested each participant once, six studies tested some participants more than once, eight studies tested twice or more, and 17 studies did not report on the frequency of testing per participant. Twenty-four studies included inpatients, four studies included outpatients, while the remaining six studies were conducted in unclear settings. Eleven studies (32%) described the co-morbidities of the study population, which commonly included hypertension, cardiovascular disease, and diabetes; however, the overall presence of co-morbidities in the participant groups of these studies was unclear.

Index tests

Thirty-three studies evaluated a single imaging modality and one study evaluated two imaging modalities. In total, the 34 studies reported a total of 35 imaging modality evaluations.

Chest CT was evaluated in 31 studies, chest X-ray was evaluated in three studies, and one study examined the diagnostic performance of ultrasound of the lungs.

Methodological quality of included studies

Figure 2 provides a summary of the overall methodological quality assessment using the QUADAS-2 tool for all 34 included studies. Refer to Figure 3 for study-level quality assessment.

Figure 2. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies

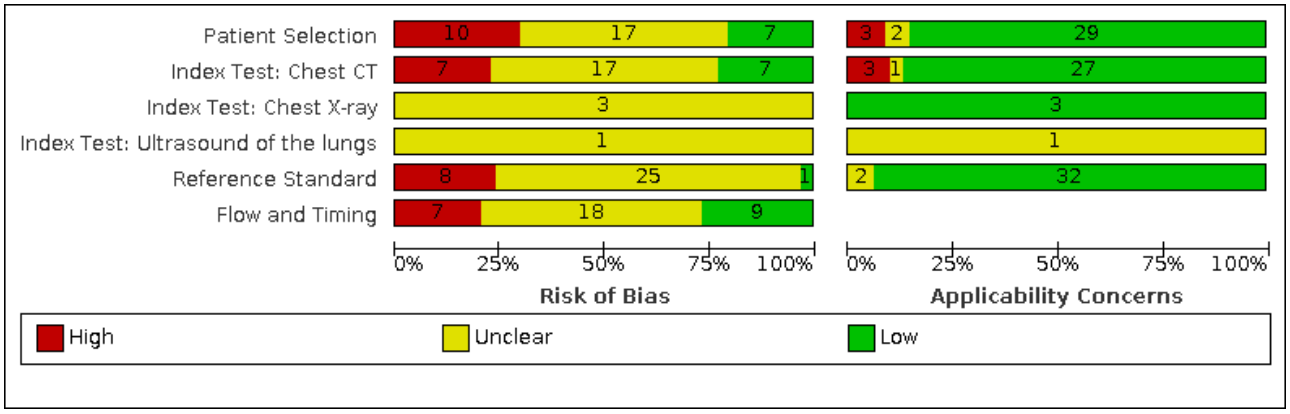
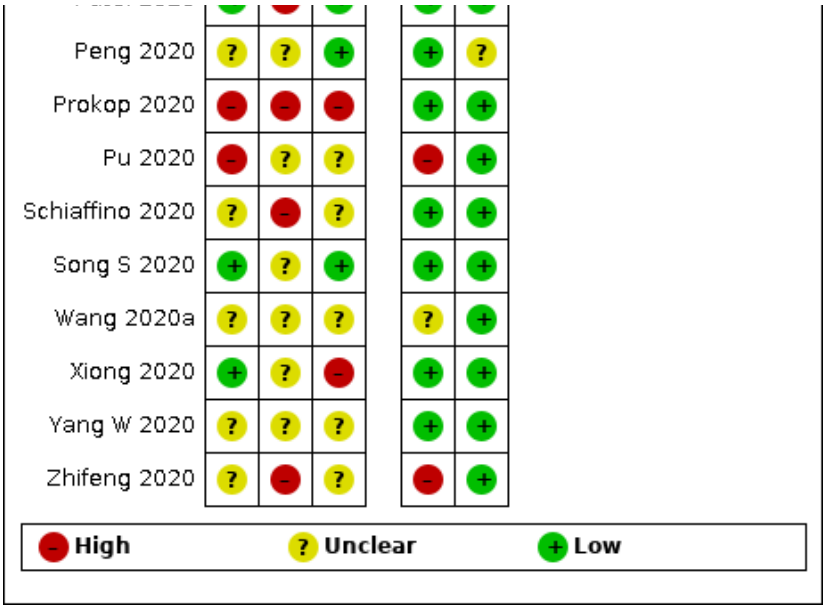


Figure 3. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study

Figure 3. (Continued)



Overall, risk of bias based on concerns about the selection of participants was found to be high and unclear in 10 (29%) and 17 (50%) studies, respectively. Risk of bias because of concerns regarding application of chest CT was high and unclear in 7/31 (23%) and 17/31 (55%) studies, respectively; risk of bias because of concerns regarding application of chest X-ray was unclear in 3/3 studies, and unclear in 1/1 study because of concerns about the application of ultrasound of the lungs. Risk of bias based on concerns about the reference standard was high and unclear in 8 (24%) and 25 (74%) studies, respectively; risk of bias based on concerns related to participant flow and timing was high and unclear in 7 (21%) and 18 (53%) studies, respectively. Concerns about the applicability of the evidence to participants were high and unclear in 3 (9%) and 2 (6%) studies, respectively. Concerns about the applicability of the evidence to the index test were high and unclear in 3/31 (10%) and 1/31 (3%) studies of chest CT, respectively, low in 3/3 studies of chest X-ray, and unclear in 1/1 (100%) study of ultrasound of the lungs. Furthermore, concerns about the applicability of the evidence to the reference standard were unclear in two (6%) studies. Additional details about risk of bias and applicability assessment are presented in Figure 3.

In the patient selection domain, the main concern was either due to inappropriate exclusions (n = 6) or the use of a case-control design involving healthy or other disease controls (n = 4). In the index test domain, the seven CT studies with a high risk of bias did not clearly define the positivity of the imaging tests evaluated. In the reference standard domain, the eight studies with a high risk of bias used an RT-PCR protocol that was not likely to correctly classify the target condition. Finally, in the patient flow domain, the six studies with a high risk of bias did not provide the same reference standard to all participants (n = 4), did not provide all participants with a reference standard (n = 1), or did not have an appropriate time interval between the reference standard and index test (n = 1).

Findings

All studies provided the 2x2 data points (TP/TN/FP/FN) required to derive and pool estimates of sensitivity and specificity. For one study that did not report values from which 2x2 data points could be determined, we extracted 2x2 data points by estimating accuracy data that were reported graphically. When the number of studies evaluating a given modality was less than 4, meta-analysis could not be performed; when the number of studies in a subgroup was less than 5, subgroup analyses could not be performed. In these cases, we summarised the data qualitatively.

Pooled estimates

Figure 4 presents the forest plot of studies that reported 2x2 data for chest CT in suspected cases. The sensitivity of CT in 31 studies (involving 4224 (53%) cases amongst 8014 participants) ranged from 57.4% to 100%, and the specificity ranged from 0% to 96.0%. The pooled sensitivity for chest CT was 89.9% (95% CI 85.7 to 92.9) and the pooled specificity was 61.1% (95% CI 42.3 to 77.1). The scatter of the study points in ROC space on the SROC plot (Figure 5) shows substantial variability in sensitivity and specificity.

Figure 4. Forest plot of chest CT in suspected cases. Sorted by publication status, followed by sensitivity and specificity. Of the 3 studies using the CO-RADS classification system, data displayed for Dofferhoff 2020 and Prokop 2020 correspond to a CO-RADS threshold of 4, and data displayed for De Smet 2020 a threshold of 5, as these thresholds produced the highest Youden's index.

Study	TP	FP	FN	TN	Publication Status	Sensitivity (95% CI)	Specificity (95% CI)
Miao 2020a	31	14	23	62	Pre-print	0.57 [0.43, 0.71]	0.82 [0.71, 0.90]
Peng 2020	25	13	10	20	Pre-print	0.71 [0.54, 0.85]	0.61 [0.42, 0.77]
Debray 2020	120	4	40	77	Pre-print	0.75 [0.68, 0.81]	0.95 [0.88, 0.99]
Patel 2020	125	41	36	115	Pre-print	0.78 [0.70, 0.84]	0.74 [0.66, 0.80]
De Smet 2020	279	33	79	468	Pre-print	0.78 [0.73, 0.82]	0.93 [0.91, 0.95]
Mao 2020	144	3	9	49	Pre-print	0.94 [0.89, 0.97]	0.94 [0.84, 0.99]
Liang 2020	20	67	1	0	Pre-print	0.95 [0.76, 1.00]	0.00 [0.00, 0.05]
Ai T 2020	580	308	21	105	Pre-print	0.97 [0.95, 0.98]	0.25 [0.21, 0.30]
Dong 2020	91	72	3	1	Pre-print	0.97 [0.91, 0.99]	0.01 [0.00, 0.07]
Mei 2020	274	39	145	447	Published	0.65 [0.61, 0.70]	0.92 [0.89, 0.94]
He 2020	26	2	8	46	Published	0.76 [0.59, 0.89]	0.96 [0.86, 0.99]
Bai 2020b	33	9	9	68	Published	0.79 [0.63, 0.90]	0.88 [0.79, 0.95]
Bai 2020a	174	64	45	141	Published	0.79 [0.73, 0.85]	0.69 [0.62, 0.75]
Prokop 2020	43	5	10	47	Published	0.81 [0.68, 0.91]	0.90 [0.79, 0.97]
Hernigou 2020	13	2	3	29	Published	0.81 [0.54, 0.96]	0.94 [0.79, 0.99]
Pu 2020	41	20	9	30	Published	0.82 [0.69, 0.91]	0.60 [0.45, 0.74]
Himoto 2020	5	5	1	10	Published	0.83 [0.36, 1.00]	0.67 [0.38, 0.88]
Miao 2020b	53	76	9	25	Published	0.85 [0.74, 0.93]	0.25 [0.17, 0.34]
Luo L 2020	26	14	4	29	Published	0.87 [0.69, 0.96]	0.67 [0.51, 0.81]
Zhifeng 2020	44	19	6	0	Published	0.88 [0.76, 0.95]	0.00 [0.00, 0.18]
Dofferhoff 2020	136	36	18	122	Published	0.88 [0.82, 0.93]	0.77 [0.70, 0.84]
Luo N 2020	70	7	8	55	Published	0.90 [0.81, 0.95]	0.89 [0.78, 0.95]
Yang W 2020	12	40	1	2	Published	0.92 [0.64, 1.00]	0.05 [0.01, 0.16]
Gezer 2020	92	5	6	119	Published	0.94 [0.87, 0.98]	0.96 [0.91, 0.99]
Xiong 2020	19	8	1	19	Published	0.95 [0.75, 1.00]	0.70 [0.50, 0.86]
Wang 2020a	580	308	21	105	Published	0.97 [0.95, 0.98]	0.25 [0.21, 0.30]
Caruso 2020	60	42	2	54	Published	0.97 [0.89, 1.00]	0.56 [0.46, 0.66]
Song S 2020	108	55	3	45	Published	0.97 [0.92, 0.99]	0.45 [0.35, 0.55]
Deng 2020	423	71	10	83	Published	0.98 [0.96, 0.99]	0.54 [0.46, 0.62]
Pakray 2020	16	2	0	0	Published	1.00 [0.79, 1.00]	0.00 [0.00, 0.84]
Ai J 2020a	20	20	0	13	Published	1.00 [0.83, 1.00]	0.39 [0.23, 0.58]

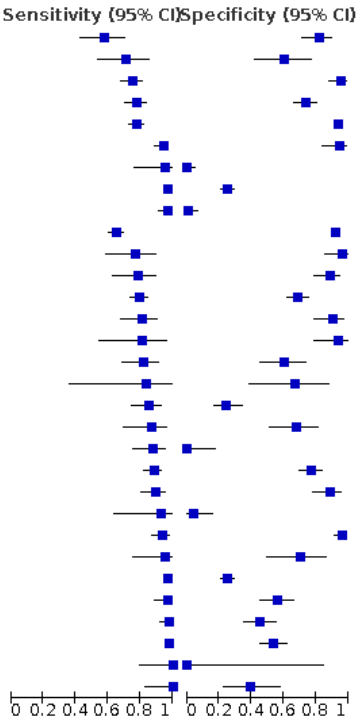
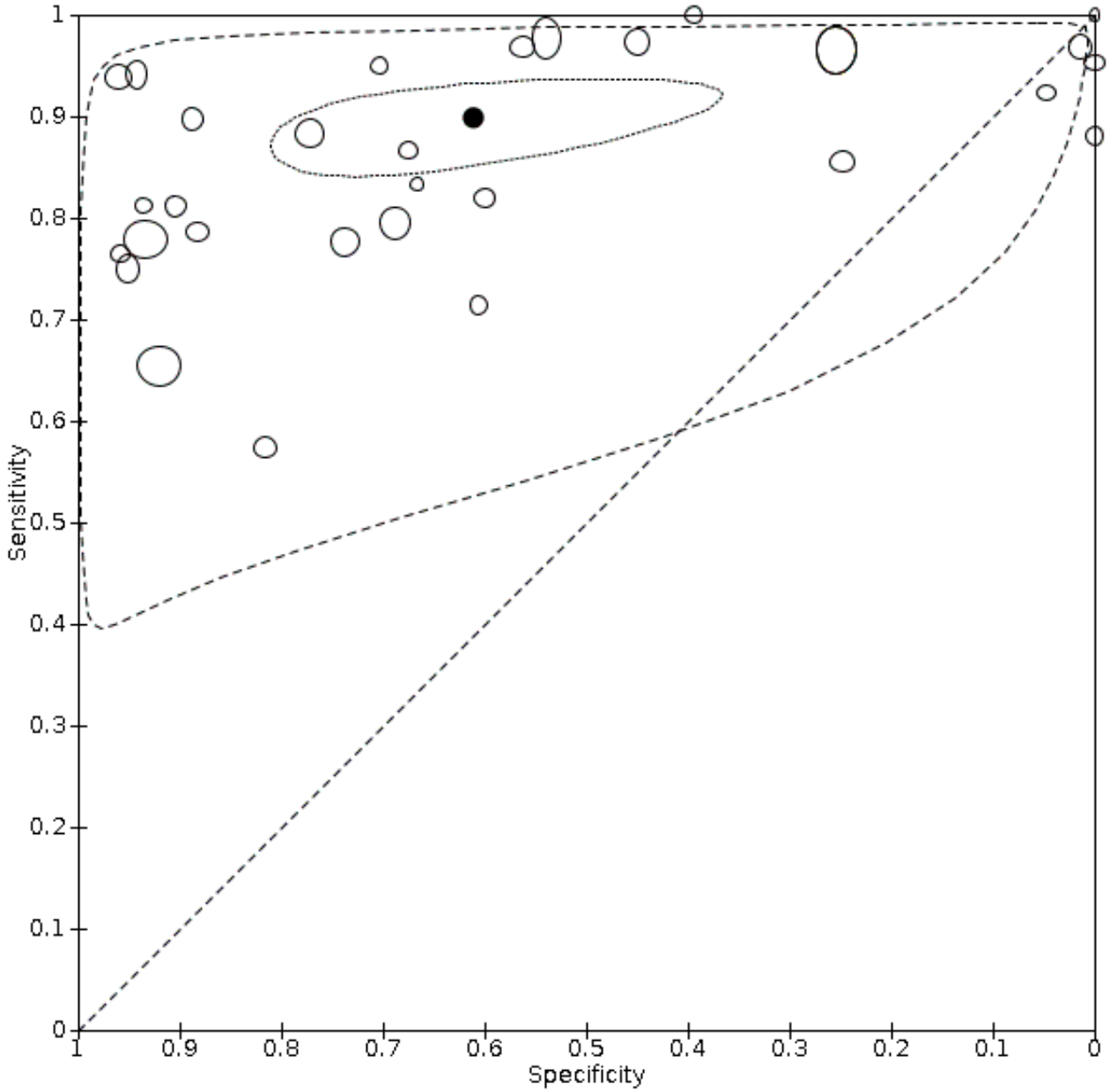
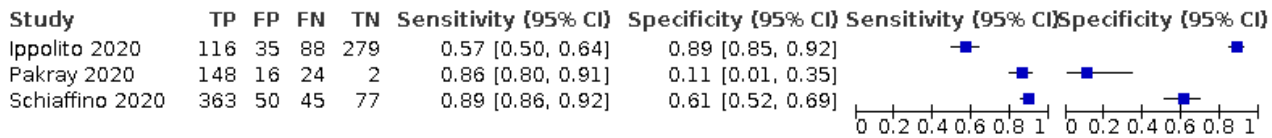


Figure 5. Summary ROC Plot of chest CT in suspected cases.



The sensitivity of X-ray in three studies (including 784 (63%) cases amongst 1243 participants)ranged from 56.9% to 89.0% and the specificity ranged from 11.1% to 88.9% (Figure 6). The sensitivity of ultrasound in one study (including 31 (31%) cases amongst 100 participants) was 96.8% and the specificity was 62.3%. Meta-analyses were not performed for X-ray and ultrasound because of the low number of included studies (< 4).

Figure 6. Forest plot of chest X-ray in suspected cases.



Sensitivity analyses

Sensitivity analyses for CT studies showed that when studies conducted in China were excluded, studies from other countries, demonstrated higher specificity compared to the overall included studies; when studies from group 3 were excluded, studies categorised into groups 1 and 2 demonstrated higher specificity compared to the overall included studies. When the studies from China (n = 17) were excluded, the studies from other countries (n = 14) had a pooled sensitivity of 86.4% (95% CI 79.6 to 91.3) and a pooled specificity of 81.5% (95% CI 67.3 to 90.4). When the studies from group 3 (which report an overview of index test findings in participants with and without the target condition, without explicitly classifying the imaging test as either COVID-19 positive or negative; n = 8) were excluded, studies categorised into either group 1 (which report specific criteria for index test positivity) or group 2 (which do not report specific criteria for index test positivity, but have the test reader(s) explicitly classify the imaging test as either COVID-19 positive or negative) (n = 23 total) together had a pooled sensitivity of 88.5% (95% CI 83.8 to 92.0) and a pooled specificity of 78.4% (95% CI 68.2 to 86.0). Of the eight studies that were categorised into group 3 and excluded for this sensitivity analysis, five studies performing a total of six imaging modality evaluations reported specificity estimates below 5.0%.

Sensitivity analyses for CT studies limiting inclusion to cross-sectional design, as well as implementing RT-PCR testing at least twice for participants with initial negative results, gave accuracy estimates similar to those of the overall included studies. When case-control studies (n = 4) were excluded, studies with a cross-sectional design (n = 24) had a pooled sensitivity of 89.6% (95% CI 84.2 to 93.3) and a pooled specificity of 61.2% (95% CI 40.0 to 78.9). Studies that implemented RT-PCR testing at least twice for participants with initial negative results (n = 6) had a pooled sensitivity of 91.0% (95% CI 74.5 to 97.2) and a pooled specificity of 68.2% (95% CI 48.0 to 83.3). The results of the sensitivity analyses are outlined in Table 2.

The sensitivity of studies that used the CO-RADS scoring system (n = 3) to define index test positivity ranged from 81.1% to 88.3% and the specificity ranged from 77.2% to 90.4%, for a CO-RADS threshold of 4. The sensitivity for a CO-RADS threshold of 5 ranged from 62.3% to 77.9% and the specificity ranged from 83.5% to 94.2%. Meta-analyses for each threshold were not performed because of the small number of included studies (<4).

Investigations of heterogeneity

Investigations of heterogeneity did not identify a statistically significant effect of publication status (preprint versus published) on accuracy estimates. Stratification by publication status for chest CT studies gave pooled sensitivity estimates of 87.8% (95% CI 79.3 to 93.1) for preprint studies versus 90.6% (95% CI 86.1 to 93.8) for published studies (P = 0.82), and pooled specificity estimates of 61.1% (95% CI 42.3 to 77.1) for preprint studies versus 49.6% (95% CI 41.7 to 57.5) for published studies (P = 0.41). These results are outlined in Table 3.

The sensitivity of chest CT studies that defined index test positivity based on radiologist impression (n = 14) ranged from 57.4% to 100% and the specificity ranged from 0% to 95.1%. The sensitivity of chest CT studies that used a formal scoring system to define index test positivity (n = 4; the threshold demonstrating the highest Youden's index in each study was used), ranged from 77.9% to 88.3% and the specificity ranged from 67.4% to 93.4%. Meta-regression was not performed due to the small number of studies included in the formal scoring system subgroup (< 5).

Subgroup analyses for chest X-ray studies were not feasible because of the low number of included studies.

DISCUSSION

This is the first update of a Cochrane living review evaluating the diagnostic accuracy of thoracic imaging (computed tomography (CT), chest X-ray and ultrasound) in the evaluation of people suspected to have COVID-19. This version of the review is based on preprints and published studies up until 22 June 2020.

Summary of main results

Chest CT (31 studies, 8014 participants, 4224 (53%) cases) demonstrated a sensitivity of 89.9% (95% CI 85.7 to 92.9), and a specificity of 61.1% (95% CI 42.3 to 77.1) in the diagnosis of COVID-19 in suspected participants. Compared with the findings of our initial review, in which chest CT was determined to have a sensitivity of 86.2% (95% CI 71.9 to 93.8) and specificity of 18.1% (95% CI 3.71 to 55.8) in suspected participants, our current update demonstrates similar sensitivity estimates and higher specificity estimates. Possible explanations for this improved specificity could include better-developed definitions for index test positivity used by index test readers (such as CO-RADS) in the studies added in this update. The stage of the pandemic during which included studies were conducted could also have influenced the differing specificity estimates, with studies from the early stage of the pandemic included in our initial review and studies from a later stage added in this update. As might be expected, studies conducted later in the pandemic would benefit from knowledge gained in prior work.

Sensitivity analyses for chest CT studies showed that studies conducted in countries other than China demonstrated higher specificity. While it appears that country of origin had an effect on our findings, the effect is more likely associated with the time at which studies were published with respect to the phase of the pandemic; the majority of the studies from China were conducted early in the pandemic, when knowledge about COVID-19 and its presentation on imaging tests was not well developed in comparison to later stages of the pandemic. Study design with respect to classification index test results also appears to have an effect on our findings, as studies that either report specific criteria for index test positivity (group 1) or do not report specific criteria for index test positivity, but have the test reader(s) explicitly classify the imaging test as either COVID-19 positive or negative (group 2) demonstrated higher specificity compared to the overall included studies. This can be explained by the study design of group 3: these studies report an overview of index test findings or classify participants as having "any abnormality" versus "no abnormality", without explicitly diagnosing participants as COVID-19 positive based on the index test. As the studies in group 3 are not intended to be diagnostic test accuracy studies, the specificity estimates they produce are expected to be very low.

Sensitivity analyses limiting inclusion to cross-sectional studies, as well as to studies that implemented RT-PCR testing at least twice for participants with initial negative results, gave accuracy estimates similar to those of the overall included studies. Thus, study design and reference standard conduct had a minimal effect

on our findings. Publication status did not appear to contribute to heterogeneity.

Chest X-ray (3 studies, 1243 participants, 784 (63%) cases) had ranges of sensitivity and specificity of 56.9% to 89.0% and 11.1% to 88.9% in the diagnosis of COVID-19 in suspected participants, respectively. The sensitivity and specificity of ultrasound of the lungs (1 study, 100 participants, 31 (31%) cases) were 96.8% and 62.3%, respectively. As the initial review did not include any studies that evaluated chest X-ray or ultrasound in the diagnosis of suspected COVID-19 participants, comparisons between our current and previous findings are not possible.

Strengths and weaknesses of the review

Our search strategy was broad and allowed for identification of a wide range of articles about COVID-19 diagnosis. Record screening, data extraction, and methodological assessment were performed independently and in duplicate by the review authors. Though we are relatively confident in the accuracy and completeness of our findings, please inform us at mmcinnes@toh.ca should errors be found so that we can address them in a future update. Furthermore, compared to our initial review, this current update includes a greater number of studies that evaluated accuracy estimates of imaging tests in the diagnosis of suspected COVID-19 participants. In future updates, these studies will remain included in pooled accuracy estimates along with newly included studies.

We did not identify publication status as a statistically significant source for variability of accuracy estimates of chest CT. These findings may suggest that the variable we investigated did not significantly contribute to variability. Alternatively, there may be confounding variables within our analyses that are obscuring the contribution to variability of the investigated variable. These findings could also be attributed to our sample size, in that our sample size may be underpowered to detect small differences, and for this reason we were unable to determine the influences of the investigated variables. Furthermore, we were unable to evaluate additional variables due to a limited number of included studies. For example, we hypothesised that the use of a formal scoring system to define index test positivity confers higher specificity estimates, compared to index test positivity determined by radiologist impression. However, as only four studies in this analysis used a formal scoring system, we were unable to perform meta-regression to investigate the effect of this variable at this stage.

Due to the lack of available data for chest X-ray and ultrasound of the lungs, we were unable to derive pooled sensitivity and specificity estimates for these modalities. For this same reason, direct comparisons of various imaging modalities were not possible at this stage.

We were not able to evaluate accuracy estimates based on specific findings on imaging tests (e.g. ground glass, consolidation, pleural effusion) or combinations of such findings because of the lack of data granularity reported in included studies; however, this will be considered in future updates of the review.

In this update, we began exploring our secondary objective of evaluating 'threshold' effects of imaging findings of COVID-19 and accuracy measures, particularly that of the CO-RADS classification system. Studies using CO-RADS (all of which evaluated chest CT) tended to show higher specificity estimates for thresholds of 4 and 5, compared with the pooled specificity of included chest CT studies. However, we were unable to formally evaluate the varying thresholds due to the limited number of included studies that used the CO-RADS system.

We were not able to evaluate several planned additional secondary objectives due to insufficient data. Important questions concerning possible associations between findings on thoracic imaging for patients with COVID-19 and number of days after symptom onset or symptom severity remain. We hope that future updates of this review will be able to evaluate these associations as research on the role of imaging tests in the diagnosis of COVID-19 evolves.

The quality of reporting and weaknesses in the primary studies included in this review continue to impact the overall robustness of our study as it did in our previous review. Several studies failed to describe their participants (e.g. recruitment setting), the details of reference standard conduct used for identifying COVID-19 cases, and the definition used for positivity of the imaging tests. Furthermore, of the studies that did describe the implemented reference standard conduct, two used a composite reference standard including index test findings, which creates the risk of incorporation bias. While the lack of rigour and quality in most of the published studies could be due to the observational nature of the initial studies published during the emergence of the COVID-19 pandemic, future studies need to prioritise scientific rigour and completeness of reporting and we encourage investigators to refer to the STARD 2015 checklist (Bossuyt 2015; Hong 2018).

We recommend that the accuracy estimates reported in this review are interpreted with caution because of the use of RT-PCR as the reference standard. The results of RT-PCR are not always sensitive, and it is possible that chest CT may be more sensitive than the reference standard in some patients. However, the results of our sensitivity analysis evaluating chest CT studies that used at least two RT-PCR results to define disease negative status in suspected COVID-19 participants did not appear to be different compared to the pooled accuracy estimates of all included chest CT studies: the former had pooled sensitivity and specificity of 91.0% (95% CI 74.5 to 97.2) and 68.2% (95% CI 48.0 to 83.3), respectively, the latter had pooled sensitivity and specificity of 89.9% (95% CI 85.7 to 92.9) and 61.1% (95% CI 42.3 to 77.1), respectively. At this stage, RT-PCR remains the best tool for diagnosing COVID-19.

About a quarter of the included studies (9/33) were only available as preprint at the time of the search and had not yet been through the peer-review process. Data extracted from these studies will continue to be updated and included in future versions of our review as these studies become published in peer-reviewed journals.

Applicability of findings to the review question

As the studies in our cohort included suspected COVID-19 participants, our findings are applicable to individuals suspected to have COVID-19. Our search did not identify many studies that evaluated the accuracy of chest CT, ultrasound of the lungs, and chest X-ray for the diagnosis of COVID-19 in paediatric populations. Thus, the diagnostic accuracy of these modalities in children is not as well established. In addition, the lack of data available in the included studies pertaining to signs and symptoms of presenting cases, the severity of the symptoms, as well as timing of symptom

onset adds complexity to the interpretation of the findings in this review.

AUTHORS' CONCLUSIONS

Implications for practice

The uncertainty resulting from high or unclear risk of bias and the heterogeneity of included studies limit our ability to confidently draw conclusions based on our results. Our findings indicate that chest computed tomography (CT) gives a higher proportion of positive results for patients with a SARS-CoV-2 infection as compared to those without: the chances of getting a positive CT result are 89.9% (95% CI 85.7 to 92.9) in patients with a SARS-CoV-2 infection and 38.9% (95% CI 22.9 to 57.7) in patients without. Due to the limited availability of data, accuracy estimates of chest X-ray and ultrasound of the lungs for the diagnosis of COVID-19 in suspected participants should be carefully interpreted.

Implications for research

From our current pool of included reports, we can draw limited conclusions regarding the diagnostic performance of thoracic imaging modalities. Additional studies evaluating the accuracy of COVID-19 in suspected patients are needed to allow for more reliable findings.

In this update, we were unable to assess several secondary objectives due to the lack of available data required to evaluate direct comparisons of different imaging modalities, and the effect of time since onset of symptoms on the diagnostic performance of various index tests. Future studies should ideally pre-define positive imaging findings and include direct comparisons of the various modalities of interest on the same participant population in order to provide robust and reliable data. Furthermore, improved transparency and reporting is necessary for more efficient data extraction in our updated versions of this review. We encourage authors and investigators to refer to the STARD 2015 checklist (Bossuyt 2015; Hong 2018) to ensure that any relevant information is clearly reported in their studies.

We hope that future updates of this review include more informative studies and more studies using formal scoring systems (such as CO-RADS) to allow for additional investigations of variability with improved power and the evaluation of secondary objectives.

ACKNOWLEDGEMENTS

Members of the Cochrane COVID-19 Diagnostic Test Accuracy Review Group include the following.

-The project team (Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MMG, Spijker R, Hooft L, Van den Bruel A, McInnes MDF, Emperador D, Dittrich S);

-The systematic review teams for each review.

- Molecular, antigen, and antibody tests (Adriano A, Beese S, Dretzke J, Ferrante di Ruffano L, Harris I, Price M, Taylor-Phillips S)
- Signs and symptoms (Stuyf T, Domen J, Horn S)
- Routine laboratory markers (Yang B, Langendam M, Ochodo E, Guleid F, Holtman G, Verbakel J, Wang J, Stegeman I)
- Imaging tests (Islam N, Salameh JP, McGrath TA, van der Pol CB, Frank RA, Kazi S, Prager R, Hare SS, Dennie C, Jenniskens K, Korevaar DA, Cohen JF, van de Wijgert J, Wang J);

-The wider team of systematic reviewers from University of Birmingham, UK who assisted with title and abstract screening across the entire suite of reviews for the diagnosis of COVID-19 (Agarwal R, Baldwin S, Berhane S, Herd C, Kristunas C, Quinn L, Scholefield B).

We thank Dr Jane Cunningham (World Health Organization) for participation in technical discussions and comments on the manuscript.

The Cochrane Central Editorial Service collaborated with Cochrane Infectious Diseases Group (CIDG) on the management of the editorial process. We thank Helen Wakeford and Clare Dooley (Editorial and Methods Department, Cochrane) and Deirdre Walshe (CIDG) for editorial feedback; Kanekal Gautham (Contact Editor), Sophie Beese and Bella Harris (Managing Editors) from the Cochrane Diagnostic Test Accuracy Reviews Editorial Team; Robin Featherstone for search peer review; peer reviewers Fergus Macbeth, Ali Gholamrezanezhad, Paul Garner and Robert Walton; consumer referee Shirley Hall; and Heather Maxwell who copy-edited the review.

The CIDG editorial base is funded by UK aid from the UK government for the benefit of low- and middle-income countries (project number 300342-104). The views expressed do not necessarily reflect the UK government's official policies.

Jonathan Deeks is a United Kingdom National Institute for Health Research (NIHR) Senior Investigator Emeritus. Yemisi Takwoingi is supported by a NIHR Postdoctoral Fellowship. Jonathan Deeks, Jacqueline Dinnes, and Yemisi Takwoingi are supported by the NIHR Birmingham Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

REFERENCES

References to studies included in this review

Ai J 2020a *{published data only}*

Ai J, Gong J, Xing L, He R, Tian F, Wang J, et al. Analysis of factors associated early diagnosis in coronavirus disease 2019 (COVID-19). *medRxiv* 2020.

Ai T 2020 *{published data only}*

Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing for Coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology* 2020;**296**(2):E32-E40.

Bai 2020a *{published data only}*

Bai HX, Hsieh B, Xiong Z, Halsey K, Choi JW, Tran TM, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. *Radiology* 2020;**296**(2):E46-E54.

Bai 2020b *{published data only}*

Bai HX, Wang R, Xiong Z, Hsieh B, Chang K, Halsey K, et al. Artificial intelligence augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other origin at chest CT. *Radiology* 2020;**296**(3):E156-65.

Bar 2020 *{published data only}*

Bar S, Lecourtois A, Diouf M, Goldberg E, Bourbon C, Arnaud E, et al. The association of lung ultrasound images with COVID-19 infection in an emergency room cohort. *Anaesthesia* 2020;**75**(12):1620-25.

Caruso 2020 *{published data only}*

Caruso D, Zerunian M, Polici M, Pucciarelli F, Polidori T, Rucci C, et al. Chest CT features of COVID-19 in Rome, Italy. *Radiology* 2020;**2962**(2):E79-E85.

Debray 2020 *{published data only}*

Debray M-P, Tarabay H, Males L, Mahdjoub E, Pavlovsky T. Observer agreement and clinical significance of chest CT reporting in patients suspected of COVID-19. *medRxiv* 2020.

Deng 2020 *{published data only}*

Deng Z, Zhang X, Li Y, Xu H, Gang Y, Wang H, et al. Value of chest CT screening in the early COVID-19 outbreak. *Chinese Journal of Radiology* 2020;**5**(54):430-4.

De Smet 2020 *{published data only}*

Smet KD, Smet DD, Demedts I, Bouckaert B, Ryckaert T, Laridon E, et al. Diagnostic power of chest CT for COVID-19: to screen or not to screen. *medRxiv* 2020.

Dofferhoff 2020 *{published data only}*

Dofferhoff AS, Swinkels A, Sprong T, Berk Y, Spanbroek M, Nabuurs-Franssen MH, et al. Diagnostic algorithm for COVID-19 at the ER. *Nederlands Tijdschrift voor Geneeskunde* 2020;**164**:D5042.

Dong 2020 *{published data only}*

Dong J, Wu L, Jin Q, Chen J, He J. Chest CT Scan of hospitalized patients with COVID-19: a case-control study. *medRxiv* 2020.

Gezer 2020 *{published data only}*

Gezer NS, Ergan B, Barış MM, Appak Ö, Sayıner AA, Balcı P, et al. COVID-19 S: A new proposal for diagnosis and structured reporting of COVID-19 on computed tomography imaging. *Diagnostic and Intervional Radiology* 2020;**26**(4):315–22.

He 2020 *{published data only}*

He J-L, Luo L, Luo Z-D, Lyu J-X, Ng M-Y, Shen X-P, et al. Diagnostic performance between CT and initial real-time RT-PCR for clinically suspected 2019 coronavirus disease (COVID-19) patients outside Wuhan, China. *Respiratory Medicine* 2020;**168**:105980.

Hernigou 2020 *{published data only}*

Hernigou J, Cornil F, Poignard A, El Bouchaibi S, Mani J, Naouri JF, et al. Thoracic computerised tomography scans in one hundred eighteen orthopaedic patients during the COVID-19 pandemic: identification of chest lesions; added values; help in managing patients; burden on the computerised tomography scan department. *International Orthopaedics (SICOT)* 2020;**44**(8):1571–80.

Himoto 2020 *{published data only}*

Himoto Y, Sakata A, Kirita M, Hiroi T, Kobayashi K I, Kubo K, et al. Diagnostic performance of chest CT to differentiate COVID-19 pneumonia in non-high-epidemic area in Japan. *Japanese Journal of Radiology* 2020;**38**(5):400-6.

Ippolito 2020 *{published data only}*

Ippolito D, Pecorelli A, Maino C, Capodaglio C, Mariani I, Giandola T, et al. Diagnostic impact of bedside chest X-ray features of 2019 novel coronavirus in the routine admission at the emergency department: case series from Lombardy region. *European Journal of Radiology* 2020;**129**:109092.

Liang 2020 *{published data only}*

Liang Y, Liang J, Zhou Q, Li X, Lin F, Deng Z, et al. Prevalence and clinical features of 2019 novel coronavirus disease (COVID-19) in the Fever Clinic of a teaching hospital in Beijing: a single-center, retrospective study. *medRxiv* 2020.

Luo L 2020 *{published data only}*

Luo L, Luo Z, Jia Y, Zhou C, He J, Lyu J, et al. CT differential diagnosis of COVID-19 and non-COVID-19 in symptomatic suspects: a practical scoring method. *BMC Pulmonary Medicine* 2020;**20**(1):129.

Luo N 2020 *{published data only}*

Luo N, Zhang H, Zhou Y, Kong Z, Sun W, Huang N, et al. Utility of chest CT in diagnosis of COVID-19 pneumonia. *Diagnostic and Interventional Radiology* 2020;**26**(5):437–42.

Mao 2020 *{published data only}*

Mao X, Liu X-P, Xiong M, Yang X, Jin X, Li Z, et al. Development and validation of chest CT-based imaging biomarkers for early stage COVID-19 screening. *medRxiv* 2020.

Mei 2020 *{published data only}*

Mei X, Lee H-C, Diao K, Huang M, Lin B, Liu C, et al. Artificial intelligence-enabled rapid diagnosis of COVID-19 patients. medRxiv 2020.

Miao 2020a *{published data only}*

Miao C, Zhuang J, Jin M, Xiong H, Huang P, Zhao Q, et al. A comparative multi-centre study on the clinical and imaging features of confirmed and unconfirmed patients with COVID-19. *medRxiv* 2020.

Miao 2020b *{published data only}*

Miao C, Jin M, Miao L, Yang X, Huang P, Xiong H, et al. Early chest computed tomography to diagnose COVID-19 from suspected patients: a multicenter retrospective study. *American Journal of Emergency Medicine* 2020;**S0735-6757**(20):30281-3.

Pakray 2020 *{published data only}*

Pakray A, Walker D, Figacz A, Kilanowski S, Rhodes C, Doshi S, et al. Imaging evaluation of COVID-19 in the emergency department. *Emergency Radiology* 2020;**27**:579-88.

Patel 2020 *{published data only}*

Patel M, Chowdhury J, Zheng M, Abramian O, Verga S, Zhao H, et al. High Resolution CHEST CT(HRCT) evaluation in patients hospitalized with COVID-19 infection. *medRxiv* 2020.

Peng 2020 *{published data only}*

Peng D, Zhang J, Xu Y, Liu Z, Wu P. Clinical analysis and early differential diagnosis of suspected pediatric patients with 2019 novel coronavirus infection. *medRxiv* 2020.

Prokop 2020 *{published data only}*

Prokop M, van Everdingen W, van Rees Vellinga T, Quarles van Ufford H, Stöger L, Beenen L, et al. CO-RADS: a Categorical CT assessment scheme for patients suspected of having COVID-19—definition and evaluation. *Radiology* 2020;**296**(2):E97–E104.

Pu 2020 *{published data only}*

Pu J, Leader J, Bandos A, Shi J, Du P, Yu J, et al. Any unique image biomarkers associated with COVID-19? *European Radiology* 2020;**30**(NA):6221–7.

Schiaffino 2020 *{published data only}*

Schiaffino S, Tritella S, Cozzi A, Carriero S, Blandi L, Ferraris L, et al. Diagnostic performance of chest X-Ray for COVID-19 pneumonia during the SARS-CoV-2 pandemic in Lombardy, Italy. *Journal of Thoracic Imaging* 2020;**35**(4):W105–6.

Song S 2020 *{published data only}*

Song S, Wu F, Liu Y, Jiang H, Xiong F, Guo X, et al. Correlation between chest CT findings and clinical features of 211 COVID-19 suspected patients in Wuhan, China. *Open Forum Infectious Diseases* 2020;**7**(6):ofaa171.

Wang 2020a *{published data only}*

Wang Y, Hou H, Wang W, Wang W. Combination of CT and RT-PCR in the screening or diagnosis of COVID-19. *Journal of Global Health* 2020;**10**(1):010347.

Xiong 2020 *{published data only}*

Xiong Z, Fu L, Zhou H, Liu JK, Wang AM, Huang Y, et al. Construction and evaluation of a novel diagnosis pathway for 2019-Corona virus disease. *Zhonghua Yi Xue Za Zhi* 2020;**100**(16):1223-9.

Yang W 2020 *{published data only}*

Yang W, Cao Q, Qin L, Wang X, Cheng Z, Pan A, et al. Clinical characteristics and imaging manifestations of the 2019 novel coronavirus disease (COVID-19):A multi-center study in Wenzhou city, Zhejiang, China. *Journal of Infection* 2020;**80**(4):388-93.

Zhifeng 2020 *{published data only}*

Zhifeng J, Feng A, Li T. Consistency analysis of COVID-19 nucleic acid tests and the changes of lung CT. *Journal of Clinical Virology* 2020;**127**:104359.

References to studies excluded from this review

Ai J 2020b *{published data only}*

Ai J, Chen J, Wang Y, Liu X, Fan W, Qu G, et al. The cross-sectional study of hospitalized coronavirus disease 2019 patients in Xiangyang, Hubei province. *medRxiv* 2020. [DOI: 10.1101/2020.02.19.20025023]

Arentz 2020 *{published data only}*

Arentz M, Yim E, Klaff L, Lokhandwala S, Riedo FX, Chong M, et al. Characteristics and outcomes of 21 critically ill patients with COVID-19 in Washington State. *JAMA* 2020;**323**(16):1612-4.

Chang 2020 *{published data only}*

Chang D, Lin M, Wei L, Xie L, Zhu G, Dela Cruz CS, et al. Epidemiologic and clinical characteristics of novel Coronavirus infections involving 13 patients outside Wuhan, China. *JAMA* 2020;**323**(11):1092-3. [DOI: 10.1001/jama.2020.1623 10.1001/jama.2020.1623.]

Cheng 2020 *{published data only}*

Cheng Z, Lu Y, Cao Q, Qin L, Pan Z, Yan F, et al. Clinical features and chest CT manifestations of Coronavirus disease 2019 (COVID-19) in a single-centersStudy in Shanghai, China. *American Journal of Roentgenology* 2020;**215**(1):121-6. [DOI: 10.2214/ajr.20.22959 10.2214/AJR.20.22959.]

Chen S 2020 *{published data only}*

Chen S, Wu JJ, Li MZ, Xu DZ, Zhu YZ, Wang HC, et al. Clinical features of 109 cases of novel coronavirus pneumonia. *Chinese Journal of Infectious Diseases* 2020;**38**(12):E015. [DOI: 10.3760 / cma.j.issn.1000-6680.2020.0015.]

Chen X 2020 *{published data only}*

Chen X, Tang Y, Mo Y, Li S, Lin D, Yang Z, 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;**30**(9):4893-902. [DOI: 10.1007/s00330-020-06829-2]

Chen Z 2020 *{published data only}*

Chen Z, Fan H, Cai J, Li Y, Wu B, Hou Y, et al. High-resolution computed tomography manifestations of COVID-19 infections in patients of different ages. *European Journal of Radiology* 2020;**126**:108972.

Çinkooğlu 2020 *{published data only}*

Çinkooğlu A, Bayraktaroğlu S, Savaş R. Lung changes on chest CT during 2019 novel Coronavirus (COVID-19) pneumonia. *European Journal of Breast Health* 2020;**16**(2):89-90. [DOI: 10.5152/ejbh.2020.010420]

Colombi 2020 *{published data only}*

Colombi D, Bodin FC, Petrini M, Maffi G, Morelli N, Milanese G, et al. Well-aerated lung on admitting chest CT to predict adverse outcome in COVID-19 pneumonia. *Radiology* 2020;**296**(2):E86-E96. [DOI: 10.1148/radiol.2020201433]

Dai 2020 *{published data only}*

Dai H, Zhang X, Xia J, Zhang T, Shang Y, Huang R, et al. High-resolution chest CT features and clinical characteristics of patients infected with COVID-19 in Jiangsu, China. *International Journal of Infectious Diseases* 2020;**95**:106-12. [DOI: S1201-9712(20)30218-6]

Ding 2020 *{published data only}*

Ding X, Xu J, Zhou J, Long Q. Chest CT findings of COVID-19 pneumonia by duration of symptoms. *European Journal of Radiology* 2020;**127**:109009. [DOI: 10.1016/j.ejrad.2020.109009]

Guan 2020c *{published data only}*

Guan W, Liu J, Yu C. CT findings of Coronavirus disease (COVID-19) severe pneumonia. *American Journal of Roentgenol0gy* 2020;**5**:W85-W6. [DOI: 10.2214/AJR.20.23035]

Hao 2020 *{published data only}*

Hao W, Li M. Clinical diagnostic value of CT imaging in COVID-19 with multiple negative RT-PCR testing. *Travel Medicine and Infectious Disease* 2020;**34**:101627. [DOI: 10.1016/j.tmaid.2020.101627]

Huang 2020 *{published data only}*

Huang P, Liu T, Huang L, Liu H, Lei M, Xu W, et al. Use of chest CT in combination with negative RT-PCR assay for the 2019 novel Coronavirus but high clinical suspicion. *Radiology* 2020;**295**(1):22-3. [DOI: 10.1148/radiol.2020200330 10.1148/radiol.2020200330. Epub 2020 Feb 12.]

Lu 2020 *{published data only}*

Lu X, Zhang L, Du H, Zhang J, Li Y, Qu J, et al. SARS-CoV-2 infection in children. *New England Journal of Medicine* 2020;**382**(17):1663-5. [DOI: 10.1056/NEJMc2005073]

Poggiali 2020 *{published data only}*

Poggiali E, Dacrema A, Bastoni D, Tinelli V, Demichele E, Mateo Ramos P, et al. Can lung US help critical care clinicians in the early diagnosis of novel Coronavirus (COVID-19) pneumonia? *Radiology* 2020;**295**(E6):200847. [DOI: 10.1148/radiol.2020200847]

Siegel 2020 *{published data only}*

Siegel A, Chang PJ, Jarou ZJ, Paushter DM, Harmath CB, Arevalo JB, et al. Lung base findings of Coronavirus disease (COVID-19) on abdominal CT in patients with predominant gastrointestinal symptoms. *American Journal of Roentgenology* 2020;**215**(3):607-9. [DOI: 10.2214/AJR.20.23232]

Song F 2020 *{published data only}*

Song F, Shi N, Shan F, Zhang Z, Shen J, Lu H, et al. Emerging 2019 novel Coronavirus (2019-nCoV) pneumonia. *Radiology* 2020;**295**(1):210-7. [DOI: 10.1148/radiol.2020200274 10.1148/radiol.2020200274. Epub 2020 Feb 6.]

Tavare 2020 *{published data only}*

Tavare AN, Braddy A, Brill S, Jarvis H, Sivaramakrishnan A, Barnett J, et al. Managing high clinical suspicion COVID-19 inpatients with negative RT-PCR: a pragmatic and limited role for thoracic CT. *Thorax* 2020;**75**(7):537-8. [DOI: thoraxjnl-2020-214916]

Wang 2020b *{published data only}*

Wang Y, Dong C, Hu Y, Li C, Ren Q, Zhang X, et al. Temporal changes of CT findings in 90 patients with COVID-19 pneumonia: a longitudinal study. *Radiology* 2020;**296**(2):E55-E64. [DOI: 10.1148/radiol.2020200843 10.1148/radiol.2020200843.]

Wu J 2020 *{published data only}*

Wu J, Feng CL, Xian XY, Qiang J, Zhang J, Mao QX, et al. Novel coronavirus pneumonia (COVID-19) CT distribution and sign features. *Zhonghua Jie He He Hu Xi Za Zhi* 2020;**43**(4):321-6. [DOI: 10.3760/cma.j.cn112147-20200217-00106 10.3760/cma.j.cn112147-20200217-00106.]

Wu Q 2020 *{published data only}*

Wu Q, Xing Y, Shi L, Li W, Gao Y, Pan S, et al. Epidemiological and clinical characteristics of children with Coronavirus disease 2019. *medRxiv* 2020. [DOI: 10.1101/2020.03.19.20027078]

Wu X 2020a *{published data only}*

Wu X, Sun R, Chen J, Xie Y, Zhang S, Wang, X. Radiological findings and clinical characteristics of pregnant women with COVID-19 pneumonia. *International Journal of Gynaecology and Obstetrics* 2020;**150**(1):58-63. [DOI: 10.1002/ijgo.13165]

Wu X 2020b *{published data only}*

Wu X, Fu B, Chen L, Feng Y. Serological tests facilitate identification of asymptomatic SARS-CoV-2 infection in Wuhan, China. *Journal of Medical Virology* 2020;**10**:1002/jmv.25904. [DOI: 10.1002/jmv.25904]

Xie 2020 *{published data only}*

Xie S, Lei Z, Chen X, Liu W, Wang X, Dong Y, et al. Chest CT-based differential diagnosis of 28 patients with suspected corona virus disease 2019 (COVID-19). *British Journal of Radiology* 2020;**93**(1112):20200243.

Xu 2020a *{published data only}*

Xu X, Yu C, Qu J, Zhang L, Jiang S, Huang D, et al. Imaging and clinical features of patients with 2019 novel coronavirus SARS-CoV-2. *European Journal of Nuclear Medicine and Molecular*

Imaging 2020;**47**(5):1275-80. [DOI: 10.1007/s00259-020-04735-9 10.1007/s00259-020-04735-9.]

Xu 2020b *{published data only}*

Xu X, Yu C, Zhang L, Luo L, Liu J. Imaging features of 2019 novel coronavirus pneumonia. *European Journal of Nuclear Medicine and Molecular Imaging* 2020;**47**(5):1022-23. [DOI: 10.1007/s00259-020-04720-2]

Yang S 2020 *{published data only}*

Yang S, Shi Y, Lu H, Xu J, Li F, Qian Z, et al. Clinical and CT features of early-stage patients with COVID-19: a retrospective analysis of imported cases in Shanghai, China. *European Respiratory Journal* 2020;**55**(4):2000407. [DOI: 10.1183/13993003.00407-2020]

Yuan 2020 *{published data only}*

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. [DOI: 10.1371/journal.pone.0230548 10.1371/journal.pone.0230548. eCollection 2020.]

Additional references

Bossuyt 2015

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *Radiology* 2015 Oct 28;**277**(3):826–32. [DOI: 10.1148/radiol.2015151516]

Chu 2006

Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology* 2006;**59**(12):1331-2; author reply 1332-3. [DOI: 10.1016/j.jclinepi.2006.06.011]

Deeks 2020

Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database of Systematic Reviews* 2020, Issue 6. Art. No: CD013652. [DOI: 10.1002/14651858.CD013652]

Dinnes 2020

Dinnes J, Deeks JJ, Adriano A, Berhane S, Davenport C, Dittrich S, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database of Systematic Reviews* 2020, Issue 8. Art. No: CD013705. [DOI: 10.1002/14651858.CD013705]

Harbord 2009

Harbord RM, Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata Journal* 2009;**9**(2):211-29. [DOI: doi.org/10.1177/1536867X0900900203]

Hong 2018

Hong PJ, Korevaar DA, McGrath TA, Ziai H, Frank R, Alabousi M, et al. Reporting of imaging diagnostic accuracy studies with focus on MRI subgroup: adherence to STARD 2015. *Journal of Magnetic Resonance Imaging* 2018;**47**(2):523-44. [DOI: 10.1002/jmri.25797]

Irwig 1995

Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *Journal of Clinical Epidemiology* 1995;**48**(1):119-30; discussion 131-2. [DOI: 10.1016/0895-4356(94)00099-c]

Kucirka 2020

Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in false-negative rate of reverse transcriptase polymerase chain reaction-based SARS-CoV-2 tests by time since exposure. *Annals of Internal Medicine* 2020;**173**(4):262-7. [DOI: 10.7326/M20-1495]

Li 2020g

Li Y, Yao L, Li J, Chen L, Song Y, Cai Z, et al. Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID-19. *Journal of Medical Virology* 2020;**92**(7):903-8. [DOI: 10.1002/jmv.25786]

Loeffelholz 2020

Loeffelholz MJ, Tang YW. Laboratory diagnosis of emerging human coronavirus infections - the state of the art. *Emerging Microbes and Infections* 2020;**9**(1):747-56. [DOI: 10.1080/22221751.2020.1745095]

McGrath 2017

McGrath TA, McInnes MD, Langer FW, Hong J, Korevaar DA, Bossuyt PM. Treatment of multiple test readers in diagnostic accuracy systematic reviews-meta-analyses of imaging studies. *European Journal of Radiology* 2017 Aug 1;**93**:59–64. [DOI: doi.org/10.1016/j.ejrad.2017.05.032]

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, Group PRISMA . Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology* 2009;**62**(10):1006-12. [DOI: 10.1371/journal.pmed.1000097]

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982-90. [DOI: 10.1016/j.jclinepi.2005.02.022]

Rubin 2020

Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raoof S, et al. The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society. *Radiology* 2020;**296**(1):172–80. [DOI: 10.1148/radiol.2020201365]

Salameh 2020b

Salameh J-P, Bossuyt PM, McGrath TA, Thombs BD, Hyde CJ, Macaskill P, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies

(PRISMA-DTA): explanation, elaboration, and checklist. *BMJ* 2020;**370**:m2632. [DOI: 10.1136/bmj.m2632]

StataCorp 2019 [Computer program]

Stata Statistical Software. StataCorp, Version Release 16. College Station, TX: StataCorp LLC, 2019.

Stegeman 2020

Stegeman I, Ochodo EA, Guleid F, Holtman GA, Yang B, Davenport C, et al. Routine laboratory testing to determine if a patient has COVID-19. *Cochrane Database of Systematic Reviews* 2020, Issue 11. Art. No: CD013787. [DOI: 10.1002/14651858.CD013787]

Struyf 2020

Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MM, et al. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease. *Cochrane Database of Systematic Reviews* 2020, Issue 7. Art. No: CD013665. [DOI: 10.1002/14651858.CD013665]

WHO 2020

World Health Organization (WHO). WHO COVID-19 Case definition. WHO/2019-nCoV/Surveillance_Case_Definition/2020.1 (accessed 17 Oct 2020).

References to other published versions of this review

McInnes 2020

McInnes MD, Leeflang MM, Salameh J-P, McGrath TA, Pol CB, Frank RA, et al. Imaging tests for the diagnosis of COVID-19. *Cochrane Database of Systematic Reviews* 2020, Issue 6. Art. No: CD013639. [DOI: 10.1002/14651858.CD013639]

Salameh 2020a

Salameh J-P, Leeflang MM, Hooft L, Islam N, McGrath TA, Pol CB, et al. Thoracic imaging tests for the diagnosis of COVID-19. *Cochrane Database of Systematic Reviews* 2020, Issue 9. Art. No: CD013639. [DOI: 10.1002/14651858.CD013639.pub2]

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Ai J 2020a

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: unclear Setting: outpatient
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: any abnormality Level of training of readers: unclear Prevalence: 0.4
Target condition and reference standard(s)	Reference standard: RT-PCR twice, if necessary; other (lab tests)
Flow and timing	
Comparative	
Notes	

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Ai J 2020a *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Ai T 2020

Study characteristics

Ai T 2020 *(Continued)*

Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: adults only Setting: inpatient
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.6
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			

Ai T 2020 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Bai 2020a

Study characteristics			
Patient Sampling	Study design: suspected and infected patients (case-control)		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.5		
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			

Bai 2020a *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Bai 2020b

Study characteristics			
Patient Sampling	Study design: case-control		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.4		
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			

Bai 2020b *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?			
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
Were the index test results interpreted without knowledge of the results of the reference standard?			
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Bar 2020

Study characteristics	
Patient Sampling	Study design: suspected patients

Bar 2020 *(Continued)*

Patient characteristics and setting	Age group: adults only Setting: inpatient
Index tests	Index test(s): ultrasound of the lungs (POCUS) Defintion for positive diagnosis on ultrasound: unclear Level of training of readers: unclear Prevalence: 0.3
Target condition and reference standard(s)	Reference standard: RT-PCR twice, if necessary
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			

Bar 2020 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Caruso 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: adults only Setting: outpatient		
Index tests	Index test(s): chest CT; non contrast CT thorax Definition for positive diagnosis on CT: pneumonia Level of training of readers: radiologist Prevalence: 0.4		
Target condition and reference standard(s)	Reference standard: RT-PCR twice, if necessary		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			

Caruso 2020 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Debray 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: adults only Setting: inpatient		
Index tests	Index test(s): Chest CT (non contrast) Defintion for positive diagnosis on CT: quote: “evocative”: multifocal ground-glass opacities, being nodular or not, or crazy-paving with or without consolidations, with a bilateral, peripheral or mixed distribution and involvement of the posterior zones Level of training of readers: radiologist Prevalence: 0.7		
Target condition and reference standard(s)	Reference standard: RT-PCR once; twice in some		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Debray 2020 *(Continued)*

DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Deng 2020

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: children and adults Setting: inpatient
Index tests	Index test(s): chest CT (high resolution) Defintion for positive diagnosis on CT: (1) any one of the following: a) single, multiple, or diffuse ground-glass opacity, with thickened blood vessels and thickened bronchial shadows passing through, with or without localised lobular septal grid thickening; b) single or multiple real shadows, (2) re-examination 3 to 5 days later showed that the original ground-glass opacity or consolidation range increased, the number increased, or accompanied by pleural effusion on one or both sides Level of training of readers: radiologist Prevalence: 0.7
Target condition and reference standard(s)	Reference standard: RT-PCR once

Deng 2020 *(Continued)*

Flow and timing
Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Deng 2020 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

De Smet 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): Chest CT Defintion for positive diagnosis on CT: CO-RADS classification; threshold not pre-specified Level of training of readers: unclear Prevalence: 0.4		
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			

De Smet 2020 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Dofferhoff 2020

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: adults only Setting: inpatient
Index tests	Index test(s): Chest CT (low dose) Defintion for positive diagnosis on CT: CO-RADS classification; threshold not pre-specified Level of training of readers: unclear

Dofferhoff 2020 *(Continued)*

	Prevalence: 0.5		
Target condition and reference standard(s)	Reference standard: RT-PCR once; twice in some		
Flow and timing			
Comparative			
Notes			

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	

Dofferhoff 2020 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Dong 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: any abnormality Level of training of readers: unclear Prevalence: 0.6		
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	

Dong 2020 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Gezer 2020

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: adults only Setting: inpatient

Gezer 2020 *(Continued)*

Index tests	Index test(s): Chest CT (non contrast) Defintion for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.4
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided; other (clinical signs and imaging tests)
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		

Gezer 2020 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

He 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): Chest CT (high resolution) Defintion for positive diagnosis on CT: ground-glass opacity with or without consolidation, crazy paving patten, peripheral and diffuse distribution, and bilateral/multilobular involvement Level of training of readers: radiologist Prevalence: 0.4		
Target condition and reference standard(s)	Reference standard: RT-PCR once; twice in some		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

He 2020 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Hernigou 2020

Study characteristics			

Hernigou 2020 *(Continued)*

Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: adults only Setting: inpatient
Index tests	Index test(s): Chest CT (low dose) Defintion for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.3
Target condition and reference standard(s)	Reference standard: RT-PCR once; twice in some
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			

Hernigou 2020 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Himoto 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: adults only Setting: unclear		
Index tests	Index test(s): chest CT; non contrast CT thorax Definition for positive diagnosis on CT: ground-glass opacity (bilateral) and peripheral predominant lesions without airway abnormalities, nodules, mediastinal lymphadenopathy, pleural effusion Level of training of readers: resident Prevalence: 0.3		
Target condition and reference standard(s)	Reference standard: RT-PCR once; other (clinical signs)		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**

Himoto 2020 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Ippolito 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): Chest radiographs/Chest X-rays Defintion for positive diagnosis on X-ray: reticulations, alveolar opacities or both Level of training of readers: radiologist Prevalence: 0.4		
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
DOMAIN 2: Index Test (Chest X-ray)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Ippolito 2020 *(Continued)*

DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Liang 2020

Study characteristics	
Patient Sampling	Study design: suspected and infected patients
Patient characteristics and setting	Age group: adults only Setting: unclear
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: unclear Level of training of readers: unclear Prevalence: 0.2
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided; other (positive contacts)
Flow and timing	
Comparative	
Notes	
Methodological quality	

Liang 2020 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Luo L 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): Chest CT Defintion for positive diagnosis on CT: scoring system was developed; threshold not pre-specified Level of training of readers: radiologist Prevalence: 0.4		
Target condition and reference standard(s)	Reference standard: RT-PCR twice, if necessary		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Luo L 2020 *(Continued)*

DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Luo N 2020

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: adults only Setting: outpatient
Index tests	Index test(s): Chest CT Defintion for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.6
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided
Flow and timing	
Comparative	
Notes	

Luo N 2020 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Unclear		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Luo N 2020 *(Continued)*

Could the patient flow have introduced bias?	Unclear risk

Mao 2020

Study characteristics			
Patient Sampling	Study design: infected and control patients (case-control)		
Patient characteristics and setting	Age group: adults only Setting: inpatient		
Index tests	Index test(s): Chest CT Defintion for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.7		
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	

Mao 2020 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Mei 2020

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: children and adults Setting: unclear
Index tests	Index test(s): Chest CT Defintion for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.5
Target condition and reference standard(s)	Reference standard: RT-PCR twice, if necessary
Flow and timing	
Comparative	

Mei 2020 *(Continued)*

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		

Mei 2020 *(Continued)*

Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Miao 2020a

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: adults only Setting: inpatient		
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: ground-glass opacity with bilateral pulmonary distribution Level of training of readers: unclear Prevalence: 0.4		
Target condition and reference standard(s)	Reference standard: RT-PCR twice, if necessary		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		

Miao 2020a *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Miao 2020b

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: adults only Setting: outpatient
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.4
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided
Flow and timing	

Miao 2020b *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Miao 2020b *(Continued)*

Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Pakray 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): Chest CT (IV contrast); Chest radiographs/Chest X-rays Defintion for positive diagnosis on CT: assessed for findings including: peripheral ground-glass opacities, consolidations, quote: “crazy paving” pattern, and “reverse halo sign” Defintion for positive diagnosis on X-ray: assessed for: laterality (bilateral or unilateral), lung zone (superior, basal, perihilar, and multifocal), and density (airspace or interstitial) Level of training of readers: radiologist Prevalence: 0.9		
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern

Pakray 2020 *(Continued)*

DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Patel 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): Chest CT (high resolution) Defintion for positive diagnosis on CT: multifocal pneumonia Level of training of readers: radiologist Prevalence: 0.5		
Target condition and reference standard(s)	Reference standard: RT-PCR once; twice in some		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			

Patel 2020 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Peng 2020

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: children only Setting: inpatient
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: ground-glass opacity, consolidations with surrounding halo sign, nodules, residual fiber strips, lymphadenopathy Level of training of readers: radiologist Prevalence: 0.5
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided; other (positive contacts)
Flow and timing	
Comparative	
Notes	
Methodological quality	

Peng 2020 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Prokop 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): Chest CT Defintion for positive diagnosis on CT: CO-RADS classification; threshold not pre-specified Level of training of readers: radiologist Prevalence: 0.5		
Target condition and reference standard(s)	Reference standard: RT-PCR once; twice in some		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Prokop 2020 *(Continued)*

DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Pu 2020

Study characteristics	
Patient Sampling	Study design: infected and control patients (case-control)
Patient characteristics and setting	Age group: children and adults Setting: unclear
Index tests	Index test(s): Chest CT (high resolution) Defintion for positive diagnosis on CT: unclear Level of training of readers: radiologist Prevalence: 0.5
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided
Flow and timing	
Comparative	
Notes	

Pu 2020 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Yes		

Pu 2020 *(Continued)*

Could the patient flow have introduced bias?	Unclear risk

Schiaffino 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: children and adults Setting: inpatient		
Index tests	Index test(s): Chest radiographs/Chest X-rays Defintion for positive diagnosis on X-ray: unclear Level of training of readers: radiologist Prevalence: 0.8		
Target condition and reference standard(s)	Reference standard: RT-PCR twice, if necessary		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
DOMAIN 2: Index Test (Chest X-ray)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		

Schiaffino 2020 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Song S 2020

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: adults only Setting: inpatient
Index tests	Index test(s): Chest CT Defintion for positive diagnosis on CT: diagnosis of viral pneumonia according to: multiple bilateral, ill-defined ground-glass opacities (GGOs) or mixed consolidation with diffuse peripheral distribution or bilateral pulmonary consolidation Prevalence: 0.5
Target condition and reference standard(s)	Reference standard: RT-PCR twice, if necessary
Flow and timing	

Song S 2020 *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Song S 2020 *(Continued)*

Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Wang 2020a

Study characteristics			
Patient Sampling	Study design: unclear		
Patient characteristics and setting	Age group: unclear Setting: unclear		
Index tests	Index test(s): Chest CT Defintion for positive diagnosis on CT: unclear Level of training of readers: unclear Prevalence: 0.6		
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		

Wang 2020a *(Continued)*

If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Xiong 2020

Study characteristics	
Patient Sampling	Study design: suspected patients
Patient characteristics and setting	Age group: children and adults Setting: inpatient
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: subpleural ground-glass opacity without pleural effusion, bronchial changes or lymphadenopathy Level of training of readers: radiologist Prevalence: 0.4

Xiong 2020 *(Continued)*

Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Xiong 2020 *(Continued)*

DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		High risk	

Yang W 2020

Study characteristics			
Patient Sampling	Study design: suspected patients		
Patient characteristics and setting	Age group: adults only Setting: unclear		
Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: ground-glass opacity, patch-like shadows, liver shadow, pleural effusion or pleural thickening Level of training of readers: unclear Prevalence: 0.2		
Target condition and reference standard(s)	Reference standard: RT-PCR, no other details provided		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	

Yang W 2020 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		Unclear risk	

Zhifeng 2020

Study characteristics	
Patient Sampling	Study design: symptomatic infected patients only
Patient characteristics and setting	Age group: adults only Setting: inpatient

Zhifeng 2020 *(Continued)*

Index tests	Index test(s): chest CT Definition for positive diagnosis on CT: unclear Level of training of readers: unclear Prevalence: 0.7
Target condition and reference standard(s)	Reference standard: RT-PCR once
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (Chest CT)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Chest X-ray)			
DOMAIN 2: Index Test (Ultrasound of the lungs)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		

Zhifeng 2020 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

CO-RADS: COVID-19 Reporting and Data System; **CT:** computed tomography; **IV:** intravenous; **POCUS:** point-of-care ultrasound; **RT-PCR:** reverse transcriptase polymerase chain reaction.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Ai J 2020b	Ineligible setting
Arentz 2020	Ineligible population
Chang 2020	< 10 participants
Cheng 2020	Ineligible outcomes
Chen S 2020	Ineligible outcomes
Chen X 2020	Ineligible outcomes
Chen Z 2020	Ineligible study population
Çinkooğlu 2020	Ineligible study design
Colombi 2020	Ineligible outcomes
Dai 2020	Ineligible outcomes
Ding 2020	Ineligible outcomes
Guan 2020c	<10 participants
Hao 2020	< 10 participants
Huang 2020	< 10 participants

Study	Reason for exclusion
Lu 2020	Ineligible patient population
Poggiali 2020	Ineligible outcomes
Siegel 2020	Ineligible study design
Song F 2020	Ineligible outcomes
Tavare 2020	Ineligible study design
Wang 2020b	Ineligible patient population
Wu J 2020	Ineligible setting
Wu Q 2020	Ineligible setting
Wu X 2020a	Ineligible patient population
Wu X 2020b	Ineligible patient population
Xie 2020	Ineligible study design
Xu 2020a	Ineligible outcomes
Xu 2020b	< 10 participants
Yang S 2020	Ineligible setting
Yuan 2020	Ineligible indication

DATA

Presented below are all the data for all of the tests entered into the review.

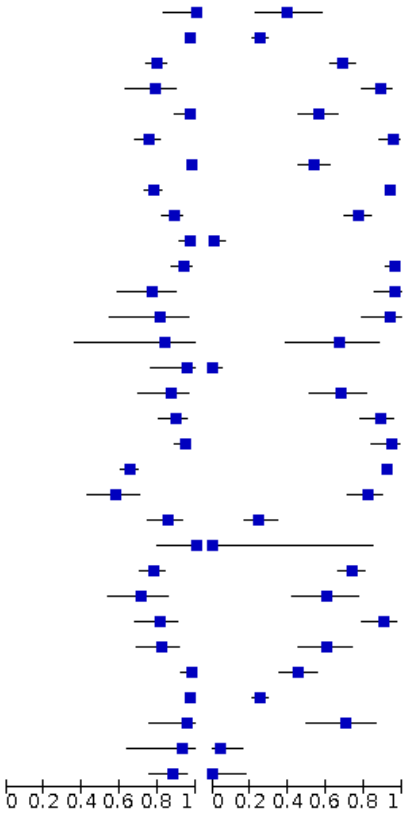
Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 Chest CT in suspected cases	31	8014
2 Chest X-ray in suspected cases	3	1243
3 Ultrasound of the lungs in suspected cases	1	100

Test 1. Chest CT in suspected cases

Chest CT in suspected cases

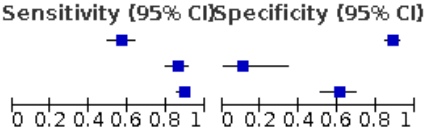
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ai J 2020a	20	20	0	13	1.00 [0.83, 1.00]	0.39 [0.23, 0.58]
Ai T 2020	580	308	21	105	0.97 [0.95, 0.98]	0.25 [0.21, 0.30]
Bai 2020a	174	64	45	141	0.79 [0.73, 0.85]	0.69 [0.62, 0.75]
Bai 2020b	33	9	9	68	0.79 [0.63, 0.90]	0.88 [0.79, 0.95]
Caruso 2020	60	42	2	54	0.97 [0.89, 1.00]	0.56 [0.46, 0.66]
Debray 2020	120	4	40	77	0.75 [0.68, 0.81]	0.95 [0.88, 0.99]
Deng 2020	423	71	10	83	0.98 [0.96, 0.99]	0.54 [0.46, 0.62]
De Smet 2020	279	33	79	468	0.78 [0.73, 0.82]	0.93 [0.91, 0.95]
Dofferhoff 2020	136	36	18	122	0.88 [0.82, 0.93]	0.77 [0.70, 0.84]
Dong 2020	91	72	3	1	0.97 [0.91, 0.99]	0.01 [0.00, 0.07]
Gezer 2020	92	5	6	119	0.94 [0.87, 0.98]	0.96 [0.91, 0.99]
He 2020	26	2	8	46	0.76 [0.59, 0.89]	0.96 [0.86, 0.99]
Hernigou 2020	13	2	3	29	0.81 [0.54, 0.96]	0.94 [0.79, 0.99]
Himoto 2020	5	5	1	10	0.83 [0.36, 1.00]	0.67 [0.38, 0.88]
Liang 2020	20	67	1	0	0.95 [0.76, 1.00]	0.00 [0.00, 0.05]
Luo L 2020	26	14	4	29	0.87 [0.69, 0.96]	0.67 [0.51, 0.81]
Luo N 2020	70	7	8	55	0.90 [0.81, 0.95]	0.89 [0.78, 0.95]
Mao 2020	144	3	9	49	0.94 [0.89, 0.97]	0.94 [0.84, 0.99]
Mei 2020	274	39	145	447	0.65 [0.61, 0.70]	0.92 [0.89, 0.94]
Miao 2020a	31	14	23	62	0.57 [0.43, 0.71]	0.82 [0.71, 0.90]
Miao 2020b	53	76	9	25	0.85 [0.74, 0.93]	0.25 [0.17, 0.34]
Pakray 2020	16	2	0	0	1.00 [0.79, 1.00]	0.00 [0.00, 0.84]
Patel 2020	125	41	36	115	0.78 [0.70, 0.84]	0.74 [0.66, 0.80]
Peng 2020	25	13	10	20	0.71 [0.54, 0.85]	0.61 [0.42, 0.77]
Prokop 2020	43	5	10	47	0.81 [0.68, 0.91]	0.90 [0.79, 0.97]
Pu 2020	41	20	9	30	0.82 [0.69, 0.91]	0.60 [0.45, 0.74]
Song S 2020	108	55	3	45	0.97 [0.92, 0.99]	0.45 [0.35, 0.55]
Wang 2020a	580	308	21	105	0.97 [0.95, 0.98]	0.25 [0.21, 0.30]
Xiong 2020	19	8	1	19	0.95 [0.75, 1.00]	0.70 [0.50, 0.86]
Yang W 2020	12	40	1	2	0.92 [0.64, 1.00]	0.05 [0.01, 0.16]
Zhifeng 2020	44	19	6	0	0.88 [0.76, 0.95]	0.00 [0.00, 0.18]



Test 2. Chest X-ray in suspected cases

Chest X-ray in suspected cases

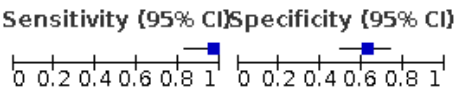
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ippolito 2020	116	35	88	279	0.57 [0.50, 0.64]	0.89 [0.85, 0.92]
Pakray 2020	148	16	24	2	0.86 [0.80, 0.91]	0.11 [0.01, 0.35]
Schiaffino 2020	363	50	45	77	0.89 [0.86, 0.92]	0.61 [0.52, 0.69]



Test 3. Ultrasound of the lungs in suspected cases

Ultrasound of the lungs in suspected cases

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bar 2020	30	26	1	43	0.97 [0.83, 1.00]	0.62 [0.50, 0.74]



ADDITIONAL TABLES

Table 1. Summary of included studies

Study ID	Journal	Country of Corresponding Author	Study design	Age group	Setting	Index test(s)	Definition for index test positivity	Group (categorised by index test positivity)	Level of training of readers	Reference standard	Prevalence
Ai J 2020a	medRxiv	China	Suspected patients	Unclear	Outpatient	Chest CT	Any abnormality	3	Unclear	RT-PCR twice, if necessary; other (lab tests)	0.4
Ai T 2020	Radiology	China	Suspected patients	Adults only	Inpatient	Chest CT	Unclear	2	Radiologist	RT-PCR, no other details provided	0.6
Bai 2020a	Radiology	China	Suspected and infected patients (case-control)	Children and adults	Inpatient	Chest CT	Unclear	2	Radiologist	RT-PCR, no other details provided	0.5
Bai 2020b	Radiology	China	Infected and control patients (case-control)	Children and adults	Inpatient	Chest CT	Unclear	2	Radiologist	RT-PCR, no other details provided	0.4
Bar 2020	Anaesthesia	France	Suspected patients	Adults only	Inpatient	Ultrasound of the lungs (POCUS)	Unclear	2	Unclear	RT-PCR twice, if necessary	0.3
Caruso 2020	Radiology	Italy	Suspected patients	Adults only	Outpatient	Chest CT (non contrast)	Pneumonia	2	Radiologist	RT-PCR twice, if necessary	0.4

Table 1. Summary of included studies *(Continued)*

De Smet 2020	medRxiv	Belgium	Suspected patients	Children and adults	Inpatient	Chest CT	CO-RADS classification; threshold not prespecified	1	Unclear	RT-PCR, no other details provided	0.4
Debray 2020	medRxiv	France	Suspected patients	Adults only	Inpatient	Chest CT (non contrast)	"Evocative": multifocal ground-glass opacities, being nodular or not, or crazy-paving with or without consolidations, with a bilateral, peripheral or mixed distribution and involvement of the posterior zones	2	Radiologist	RT-PCR once; twice in some	0.7
Deng 2020	Chinese Journal of Radiology	China	Suspected patients	Children and adults	Inpatient	Chest CT (high resolution)	(1) Any one of the following: a) Single, multiple, or diffuse ground-glass opacity, with thickened blood vessels and thickened bronchial shadows passing through, with or without localised lobular septal grid thickening; b) Single or multiple real shadows, (2) Re-examination 3 to 5 days later showed that the original ground-glass opacity or consolidation range increased, the number increased, or accompanied by pleural effusion on one or both sides	2	Radiologist	RT-PCR once	0.7
Dofferhoff 2020	Nederlands Tijdschrift voor Geneeskunde	The Netherlands	Suspected patients	Adults only	Inpatient	Chest CT (low dose)	CO-RADS classification; threshold not prespecified	1	Unclear	RT-PCR once; twice in some	0.5
Dong 2020	medRxiv	China	Suspected patients	Children and adults	Inpatient	Chest CT	Any abnormality	3	Unclear	RT-PCR, no other details provided	0.6
Gezer 2020	Diagnostic and Interventional	Turkey	Suspected patients	Adults only	Inpatient	Chest CT (non contrast)	Unclear	2	Radiologist	RT-PCR, no other details provided; other (clini-	0.4

Table 1. Summary of included studies *(Continued)*

	Radiology									cal signs and imaging tests)	
He 2020	Respiratory Medicine	China	Suspected patients	Children and adults	Inpatient	Chest CT (high resolution)	Ground-glass opacity with or without consolidation, crazy paving patten, peripheral and diffuse distribution, and bilateral/multilobular involvement	2	Radiologist	RT-PCR once; twice in some	0.4
Hernigou 2020	International Orthopaedics	Belgium	Suspected patients	Adults only	Inpatient	Chest CT (low dose)	Unclear	2	Radiologist	RT-PCR once; twice in some	0.3
Himoto 2020	Japanese Journal of Radiology	Japan	Suspected patients	Adults only	Unclear	Chest CT (non contrast)	Ground-glass opacity (bilateral) and peripheral predominant lesions without airway abnormalities, nodules, mediastinal lymphadenopathy, pleural effusion	2	Resident	RT-PCR once; other (clinical signs)	0.3
Ippolito 2020	European Journal of Radiology	Italy	Suspected patients	Children and adults	Inpatient	Chest radiographs / Chest X-rays	Reticulations, alveolar opacities or both	2	Radiologist	RT-PCR, no other details provided	0.4
Liang 2020	medRxiv	China	Suspected and infected patients	Adults only	Unclear	Chest CT	Unclear	3	Unclear	RT-PCR, no other details provided; other (positive contacts)	0.2
Luo N 2020	Diagnostic and Interventional Radiology	China	Suspected patients	Adults only	Outpatient	Chest CT	Unclear	2	Radiologist	RT-PCR, no other details provided	0.6
Luo L 2020	BMC Pulmonary Medicine	China	Suspected patients	Children and adults	Inpatient	Chest CT	Scoring system was developed; threshold not prespecified	1	Radiologist	RT-PCR twice, if necessary	0.4

Table 1. Summary of included studies *(Continued)*

Mao 2020	medRxiv	China	Infected and control patients (case-control)	Adults only	Inpatient	Chest CT	Unclear	2	Radiologist	RT-PCR, no other details provided	0.7
Mei 2020	Nature Medicine	USA	Suspected patients	Children and adults	Unclear	Chest CT	Unclear	2	Radiologist	RT-PCR twice, if necessary	0.5
Miao 2020a	medRxiv	China	Suspected patients	Adults only	Inpatient	Chest CT	Ground-glass opacity with bilateral pulmonary distribution	3	Unclear	RT-PCR twice, if necessary	0.4
Miao 2020b	American Journal of Emergency Medicine	China	Suspected patients	Adults only	Outpatient	Chest CT	Unclear	3	Radiologist	RT-PCR, no other details provided	0.4
Pakray 2020	Emergency Radiology	USA	Suspected patients	Children and adults	Inpatient	Chest CT (IV contrast); Chest radiographs / Chest X-rays	CT: peripheral ground-glass opacities, consolidations, “crazy paving” pattern, “reverse halo sign"; X-ray: laterality (bilateral or unilateral), lung zone (superior, basal, perihilar, and multifocal), density (airspace or interstitial)	3	Radiologist	RT-PCR, no other details provided	0.9
Patel 2020	medRxiv	USA	Suspected patients	Children and adults	Inpatient	Chest CT (high resolution)	Multifocal pneumonia	2	Radiologist	RT-PCR once; twice in some	0.5
Peng 2020	medRxiv	China	Suspected patients	Children only	Inpatient	Chest CT	Ground-glass opacity, consolidations with surrounding halo sign, nodules, residual fibre strips, lymphadenopathy	2	Radiologist	RT-PCR, no other details provided; other (positive contacts)	0.5

Table 1. Summary of included studies *(Continued)*

Prokop 2020	Radiology	The Netherlands	Suspected patients	Children and adults	Inpatient	Chest CT	CO-RADS classification; threshold not prespecified	1	Radiologist	RT-PCR once; twice in some	0.5
Pu 2020	European Radiology	USA	Infected and control patients (case-control)	Children and adults	Unclear	Chest CT (high resolution)	Unclear	2	Radiologist	RT-PCR, no other details provided	0.5
Schiaffino 2020	Journal of Thoracic Imaging	Italy	Suspected patients	Children and adults	Inpatient	Chest radiographs / Chest X-rays	Unclear	2	Radiologist	RT-PCR twice, if necessary	0.8
Song S 2020	Open Forum Infectious Diseases	China	Suspected patients	Adults only	Inpatient	Chest CT	Viral pneumonia according to: multiple bilateral, ill-defined ground -glass opacities (GGOs) or mixed consolidation with diffuse peripheral distribution or bilateral pulmonary consolidation	2	Radiologist	RT-PCR twice, if necessary	0.5
Wang 2020a	Journal of Global Health	Australia	Unclear	Unclear	Unclear	Chest CT	Unclear	2	Unclear	RT-PCR, no other details provided	0.6
Xiong 2020	Zhonghua Yi Xue Za Zhi	China	Suspected patients	Children and adults	Inpatient	Chest CT	Subpleural ground-glass opacity without pleural effusion, bronchial changes or lymphadenopathy	2	Radiologist	RT-PCR, no other details provided	0.4
Yang W 2020	Journal of Infection	China	Suspected patients	Adults only	Unclear	Chest CT	Ground-glass opacity, patch-like shadows, fiver shadow, pleural effusion or pleural thickening	3	Unclear	RT-PCR, no other details provided	0.2
Zhifeng 2020	Journal of Clinical Virology	China	Symptomatic infected patients only	Adults only	Inpatient	Chest CT	Unclear	3	Unclear	RT-PCR once	0.7

CI: confidence interval;**CT:** computed tomography; **RT-PCR:** reverse transcriptase polymerase chain reaction

Table 2. Sensitivity analyses for chest CT of suspected cases

Analysis	Studies (n)	Number of participants (cases)	Sensitivity (95% CI)	Specificity (95% CI)
Countries other than China	14	4401 (2188)	86.4% (79.6 to 91.3)	81.5% (67.3 to 90.4)
Categorised into groups 1 and 2	23	7271 (3894)	88.5% (83.8 to 92.0)	78.4% (68.2 to 86.0)
Cross-sectional design	24	5845 (2987)	89.6% (84.2 to 93.3)	61.2% (40.0 to 78.9)
RT-PCR testing at least twice for participants with initial negative results	6	1530 (696)	91.0% (74.5 to 97.2)	68.2% (48.0 to 83.3)

CI: confidence interval;**CT:** computed tomography

Table 3. Subgroup analyses for chest CT of suspected cases

Test, analysis group	Studies (n)	Number of participants (cases)	Sensitivity (95% CI)	Specificity (95% CI)
Publication status [a]				
Preprint	9	2161 (1064)	87.8% (79.3 to 93.1)	61.1% (42.3 to 77.1)
Published	22	5853 (3160)	90.6% (86.1 to 93.8)	49.6% (41.7 to 57.5)
P value			0.82	0.41

CI: confidence interval;**CT:** computed tomography
[a]As of 1 October 2020

APPENDICES

Appendix 1. Glossary

Terminology/acronyms

- **SARS-CoV-2:** severe acute respiratory syndrome coronavirus 2, the name given to the 2019 novel coronavirus.
- **SARS-CoV-2 infection:** people with severe acute respiratory syndrome coronavirus 2, but who may or may not have any clinical manifestations of infection
- **COVID-19:** coronavirus disease 2019, the clinical manifestations/ symptoms caused by infection with SARS-CoV-2, name given to the disease associated with the virus SARS-CoV-2
- **COVID-19 Pneumonia:** COVID-19 that presents as infection-inflammation of the lungs
- **RT-PCR:** Reverse transcription polymerase chain reaction (RT-PCR) is a laboratory technique combining reverse transcription of RNA into DNA and amplification of specific DNA targets using polymerase chain reaction. In this context it is used to detect the presence of SARS-CoV-2 RNA.
- **Target condition:** the disease or condition of interest
- **Index test:** the test that is being assessed (the index test will often be a new test)
- **Reference standard:** the most reliable method for determining if the target condition ispresent or absent, used to verify index test results. This could be a combination of tests.
- **False negative:** the test does not detect a condition in someone when it is present
- **False positive:** the test detects a condition in someone when it is not present
- **True negative:** a correct diagnosis of a condition being absent

- **True positive:** a correct diagnosis of a condition being present
- **Sensitivity:** the proportion of people with the target condition (with disease) that are correctlyidentified by the index test
- **Specificity:** the proportion of people without the target condition (without disease) that arecorrectly identified by the index test
- **Positive predictive value:** the probability that someone who has tested positive for the targetcondition with the index test will actually have it (a true positive)
- **Negative predictive value:** the probability that someone who has tested negative for the targetcondition with the index test will really not have it (a true negative)
- **Secondary care:** medical care that is provided by a specialist or facility upon referral by a primary care physician and that requires more specialized knowledge, skill, or equipment than the primary care physician can provide
- **Tertiary care:** specialized care, usually for inpatients and on referral from a primary or secondary health professional, in a facility that has personnel and facilities for advanced medical investigation and treatment

Appendix 2. QUADAS-2

QUADAS-2	
Index test(s):	Imaging studies of the chest (computed tomography (CT), chest X-ray and ultrasound) for diagnosis of COVID-19
Participants (setting, intended use of index test, presentation, prior testing):	People with suspected COVID-19 All settings, in particular secondary care, emergency care and ICUs In people presenting with suspected COVID-19; suspicion may be based on prior testing, such as general lab testing. Signs and symptoms often used for triage or referral
Reference standard and target condition:	A positive diagnosis for COVID-19 by the following. 1. A positive reverse transcriptase polymerase chain reaction (RT-PCR) test for SARS-CoV-2 infection, from any manufacturer in any country, from any source, including nasopharyngeal swabs or aspirates, oropharyngeal swabs, bronchoalveolar lavage fluid (BALF), sputum, saliva, serum, urine, rectal or faecal samples. 2. Positive on WHO criteria for COVID-19 which includes some testing RT-PCR negative. 3. Positive on China CDC criteria for COVID-19 which includes some testing RT-PCR negative. 4. Positive serology in addition to consistent symptomatology. 5. Positive on study specific list of criteria for COVID-19 which includes some testing RT-PCR negative. 6. Other criteria (symptoms, imaging findings, other tests). A negative diagnosis for COVID-19 by the following. 1. COVID suspects with negative RT-PCR test results, whether tested once or more than once. 2. Pre-pandemic controls (healthy or diseased). 3. Current healthy or with another disease (no RT-PCR test). This list is not exhaustive, as we anticipate that studies will use a variety of reference standards and we plan to include all of them, at least for the earlier versions of the review. Although RT-PCR is considered the best available test, it is suspected of missing a substantial proportion of cases, and thus may not be the ideal reference standard if used as a standalone test (Li 2020g; Loeffelholz 2020). Therefore, we are likely to use alternative reference standards, such as a combination of RT-PCR, and symptoms or imaging findings, or both. We will judge how likely each reference standard definition is to correctly classify individuals in the assessment of methodological quality. All reference standards are likely to be imperfect in some way; details of reference standard evaluation are provided in the 'Risk of bias' tool below. We will use a consensus process to agree the classification of the reference standard as to what we regard

(Continued)

	as good, moderate and poor. 'Good' reference standards need to have very little change of misclassification, 'moderate', a small but acceptable risk, 'poor', a larger and probably unacceptable risk.
Participant selection	
Was a consecutive or random sample of patients enrolled?	YES: if a study explicitly states that all participants within a certain time frame were included; that this was done consecutively; or that a random selection was done. NO: if it is clear that a different selection procedure was employed; e.g. selection based on clinician's preference, or based on institutions (i.e., 'convenience' series) UNCLEAR: if the selection procedure is not clear or not reported at all.
Was a case-control design avoided?	YES: if a study explicitly states that all participants came from the same group of (suspected) patients. NO: if it is clear that a different selection procedure was employed for the participants depending on their COVID-19 status (e.g. proven infected patients in one group and proven non-infected patients in the other group). UNCLEAR: if the selection procedure is not clear or not reported at all.
Did the study avoid inappropriate in- or exclusions?	This needs to be addressed on a case-to-case basis. YES: If all eligible patients were ore or less equally suspected of having COVID-19 and were included and if the numbers in the flow chart show not too many excluded participants (a maximum of 20% of eligible patients excluded without reasons). NO: If over 20% of eligible patients were excluded without providing a reason; if only proven patients were included, or only proven non-patients were included; if in a retrospective study participants without index test or reference standard result were excluded; if exclusion was based on severity assessment post-factum or comorbidities (cardiovascular disease, diabetes, immunosuppression). If the study oversampled patients with particular characteristics likely to affect estimates of accuracy. UNCLEAR: if the exclusion criteria are not reported.
Could the selection of patients have introduced bias?	HIGH: if one or more signalling questions were answered with NO, as any deviation from the selection process may lead to bias. LOW: if all signalling questions were answered with YES. UNCLEAR: all other instances
Is there concern that the included patients do not match the review question?	This needs to be addressed on a case-to-case basis, based on the objective the included study answers to. HIGH: if accuracy was assessed in a case-control design, or the study was able to only estimate sensitivity or specificity. LOW: any situation where imaging is generally available. UNCLEAR: if a description about the participants is lacking.
Index tests	
Were the index test results interpreted without knowledge of the results of the reference standard?	YES: if blinding was explicitly stated or index test was recorded before the results from the reference standard were available NO: if it was explicitly stated that the index test results were interpreted with knowledge of the results of the reference standard

(Continued)

	UNCLEAR: if blinding was unclearly reported.
If a threshold was used, was it prespecified?	YES: for any of these index tests it is highly unlikely that any numerical threshold is used. Still we expect studies to report their criteria for test-positivity (e.g. the constellation of imaging findings used). If these criteria are reported in the methods section, we will score 'YES' for this question. NO: if the optimal criterion for test-positivity was based on the reported data (for example, different scores on a quantitative scoring system) we will score 'NO'. UNCLEAR: if the criteria for test positivity were not or unclearly reported.
Could the conduct or interpretation of the index test have introduced bias?	HIGH: if one or more signalling questions were answered with NO. LOW: if all signalling questions were answered with YES. UNCLEAR: all other instances
Is there concern that the index test, its conduct, or interpretation differ from the review question?	There is not a huge amount of variability from a technical perspective. Therefore, this question will probably be answered 'LOW' in all cases except when assessments are made using personnel not available in practice, or personnel not trained for the job, or using modalities that are uncommon in practice. We will consult expert clinicians on a case-to-case basis to judge this question.
Reference standard	
Is the reference standard likely to correctly classify the target condition?	YES: for COVID-19: RT-PCR, done by trained personnel, and repeated after a first negative RT-PCR, following guidelines for confirmed cases and done with an assay targeting minimum 2 targets in the genes N, E, S or RdRP (one target even acceptable in zone with known transmission). To clarify, a low risk of bias reference standard for true negative would require 2 (or more) negative RT-PCR results. NO: any other test UNCLEAR: if no reference standard was reported, or if it was just reported that RT-PCR was done.
Were the reference standard results interpreted without knowledge of the results of the index test?	YES: if it was explicitly stated that the reference standard results were interpreted without knowledge of the results of the index test, or if the result of the index test was obtained after the reference standard. NO: if it was explicitly stated that the reference standard results were interpreted with knowledge of the results of the index test or if the index test was used to make the final diagnosis (incorporation bias). UNCLEAR: if blinding was unclearly reported.
Could the conduct or interpretation of the reference standard have introduced bias?	HIGH: if one or more signalling questions were answered with NO. LOW: if all signalling questions were answered with YES. UNCLEAR: all other instances
Is there concern that the target condition as defined by the reference standard does not match the review question?	HIGH: there is a high concern regarding applicability of the reference standard if the reference standard actually measures a different target condition than the one we are interested in for the review. For example, if the diagnosis is only based on clinical picture, without excluding other possible causes of this clinical picture (e.g. other respiratory pathogens), then there is considerable concern that the reference standard is actually measuring something else than COVID-19. In addition, a positive RT-PCR only measures SARS-CoV-2 infection and not COVID-19 and therefore the reference standard for COVID-19 is a combination of positive RT PCR and symptoms and/or imaging findings. LOW: if above situations not present UNCLEAR: if intention for testing is not reported in the study

(Continued)

Flow and timing	
Was there an appropriate interval between index test(s) and reference standard?	YES: as the situation of a patient, including clinical presentation and disease progress, evolves rapidly and new/ongoing exposure can result in case status change. On the other hand, negative PCR results need to be repeated for several days. Therefore, an appropriate time interval will be within 7 days. NO: if there is more than 7 days between the index test and the reference standard or if patients are otherwise reported to be assessed with the index versus reference standard test at moments of different severity. UNCLEAR: if the time interval is not reported
Did all participants receive a reference standard?	YES: if all patients received a reference standard (clearly no partial verification) NO: if only (part of) the index test positives or index test negatives received the complete reference standard UNCLEAR: if it is not reported.
Did all participants receive the same reference standard?	YES: if all patients received the same reference standard (clearly no differential verification). Verification of negative PCR result with a second PCR measurement is considered to be one reference standard. NO: if (part of) the index test positives or index test negatives received a different reference standard UNCLEAR: If it is not reported.
Were all participants included in the analysis?	YES: if all included participants were included in the analyses as well NO: if after the inclusion/exclusion process, participants were removed from the analyses for different reasons: no reference standard done, no index test done, intermediate results of both index test or reference standard, indeterminate results of both index test or reference standard, samples unusable. UNCLEAR: If this is not clear from the reported numbers.
Could the patient flow have introduced bias?	HIGH: if one or more signalling questions were answered with NO, or if one question answered with NO was judged to have little impact on the methodological quality of the study (this should be justified in the scoring). LOW: if all signalling questions were answered with YES. UNCLEAR: all other instances

CT: computed tomography; **CXR:** chest X-ray; **ICU:** intensive care unit; **RT-PCR:** reverse transcriptase polymerase chain reaction; **SARS-CoV-2:** severe acute respiratory syndrome coronavirus 2; **US:** ultrasound

Appendix 3. Search classification model

A more efficient approach was required to keep up with the rapidly increasing volume of COVID-19 literature. A classification model for COVID-19 diagnostic studies was built with the model building function within Eppi Reviewer, which uses the standard SGCClassifier in Scikit-learn on word trigrams. As outputs, new documents receive a percentage (from the predict_proba function) where scores close to 100 indicate a high probability of belonging to the class 'relevant document' and scores close to 0 indicate a low probability of belonging to the class 'relevant document'. We used three iterations of manual screening (title and abstract screening, followed by full-text review) to build and test classifiers. The final included studies were used as relevant documents, while the remainder of the COVID-19 studies were used as irrelevant documents. The classifier was trained on the first round of selected articles, and tested and retrained on the second round of selected articles. Testing on the second round of selected articles revealed poor positive predictive value but 100% sensitivity at

a cut-off of 10. The poor positive predictive value is mainly due to the broad scope of our topic (all diagnostic studies in COVID-19), poor reporting in abstracts, and a small set of included documents. The model was retrained using the articles selected of the second and third rounds of screening, which added a considerable number of additional documents. This led to a large increase in positive predictive value, at the cost of a lower sensitivity, which led us to reduce the cut-off to 5. The largest proportion of documents had a score between 0-5. This set did not contain any of the relevant documents. This version of the classifier with a cut-off 5 was used in subsequent rounds and accounted for approximately 80% of the screening burden.

Appendix 4. Search strategies

1. Living search from the University of Bern

27 April 2020

From 27 April 2020, we retrieved the curated bioRxiv/medRxiv dataset link

26 March 2020 to 27 April 2020

MEDLINE: (\"Wuhan coronavirus\" [Supplementary Concept] OR \"COVID-19\" OR \"2019 ncov\"[tiab] OR ((\"novel coronavirus\"[tiab] OR \"new coronavirus\"[tiab]) AND (wuhan[tiab] OR 2019[tiab])) OR 2019-nCoV[All Fields] OR (wuhan[tiab] AND coronavirus[tiab])))))

Embase: (nCoV or 2019-nCoV or ((new or novel or wuhan) adj3 coronavirus) or covid19 or covid-19 or SARS-CoV-2).mp

bioRxiv/medRxiv: ncov or corona or wuhan or COVID or SARS-CoV-2

With the kind support of the Public Health & Primary Care Library PHC, and following guidance of the Medical Library Association

01 January 2020 to 27 April 2020

MEDLINE: ("Wuhan coronavirus" [Supplementary Concept] OR "COVID-19" OR "2019 ncov"[tiab] OR (("novel coronavirus"[tiab] OR "new coronavirus"[tiab]) AND (wuhan[tiab] OR 2019[tiab])) OR 2019-nCoV[All Fields] OR (wuhan[tiab] AND coronavirus[tiab])))))

Embase: ncov OR (wuhan AND corona) OR COVID

bioRxiv/medRxiv: ncov or corona or wuhan or COVID

2. Cochrane COVID-19 Study Register searches

Source	Strategy
CT.gov	COVID-19*
WHO ICTRP	Health topic: 2019-nCov / COVID-19
PubMed	(("2019 nCoV"[tiab] OR 2019nCoV[tiab] OR "2019 novel coronavirus"[tiab] OR "COVID 19"[tiab] OR COVID19[tiab] OR "new coronavirus"[tiab] OR "novel coronavirus"[tiab] OR "novel coro-na virus"[tiab] OR "SARS CoV-2"[tiab] OR (Wuhan[tiab] AND (coronavirus[tiab] OR "corona virus"[tiab])) OR "COVID-19"[Supplementary Concept] OR "severe acute respiratory syndrome coronavirus 2"[Supplementary Concept]) NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])) NOT (editorial[pt] OR comment[pt] OR letter[pt] OR newspaper article[pt])

*Automatic term mapping links results for 2019-nCoV, 2019 novel coronavirus, SARS-CoV-2, severe acute respiratory syndrome coron-avirus

3. CDC Library, COVID-19 Research Articles Downloadable Database

Embase records from the Stephen B. Thacker CDC Library, Covid-19 Research articles Downloadable database.

Records were obtained by the CDC Library by searching Embase through Ovid using the following search strategy.

Source	Strategy
Embase	(coronavir* OR corona virus* OR betacoronavir* OR covid19 OR covid 19 OR nCoV OR novel CoV OR CoV 2 OR CoV2 OR sarscov2 OR 2019nCoV OR wuhan virus*).mp. OR ((wuhan OR hubei OR huanan) AND (severe acute respiratory OR pneumonia*) AND outbreak*).mp. OR Coronavirus infection/ OR coronavirinae/ OR exp betacoronavirus/ Limits: 2020- OR (novel coronavir* OR novel corona virus* OR covid19 OR covid 19 OR nCoV OR novel CoV OR CoV 2 OR CoV2 OR sarscov2 OR 2019nCoV OR wuhan virus*).mp. OR ((wuhan OR hubei OR huanan) AND (severe acute respiratory OR pneumonia*) AND outbreak*).mp. OR ((wuhan OR hubei OR huanan) AND (coronavir* OR betacoronavir*)).mp. Limits: 2019-

WHAT'S NEW

Date	Event	Description
23 October 2020	New search has been performed	This is a 'living' systematic review'; searches are run and screened monthly. The last search date was 22 June 2020. Results of all new studies identified have been incorporated. The conclusions of this Cochrane Review are therefore considered up to date.
23 October 2020	New citation required and conclusions have changed	The results for chest computed tomography (CT) have changed.

HISTORY

Protocol first published: Issue 6, 2020
Review first published: Issue 9, 2020

CONTRIBUTIONS OF AUTHORS

All authors reviewed, edited, contributed to, and approved this review.

The search was performed by RS, MMGL and LH.

DECLARATIONS OF INTEREST

Jean-Paul Salameh has no known conflicts of interest.

Mariska MG Leeflang has no known conflicts of interest.

Lotty Hooft has no known conflicts of interest.

Nayaar Islam has no known conflicts of interest.

Trevor McGrath has no known conflicts of interest.

Christian B van der Pol has no known conflicts of interest.

Robert A Frank has no known conflicts of interest.

Sakib Kazi has no known conflicts of interest.

Ross Prager has no known conflicts of interest.

Samanjit Singh Hare has no known conflicts of interest.

Carole Dennie has no known conflicts of interest.

René Spijker: the Dutch Cochrane Centre (DCC) has received grants for performing commissioned systematic reviews. In no situation, the commissioner had any influence on the results of the work.

Jonathan J Deeks has no known conflicts of interest.

Jacqueline Dinnes has no known conflicts of interest.

Kevin Jenniskens has no known conflicts of interest.

Daniel Korevaar has no known conflicts of interest.

Jérémie F Cohen has no known conflicts of interest.

Ann Van den Bruel has no known conflicts of interest.

Yemisi Takwoingi has no known conflicts of interest.

Janneke van de Wijgert has no known conflicts of interest.

Junfeng Wang received a consultancy fee from Biomind, an Artificial Intelligence (AI) company providing machine intelligence solutions in medical imaging. The consultancy service was about design of clinical studies, not related to this review. The company had no influence on the results of the work.

Matthew McInnes has no known conflicts of interest.

SOURCES OF SUPPORT

Internal sources

- Liverpool School of Tropical Medicine, UK

External sources

- Foreign, Commonwealth and Development Office (FCDO), UK

  Project number: 300342-104
- National Institute for Health Research (NIHR), UK
- Government of Ontario Ministry of Health COVID-19 Rapid Response Research Grant program, Canada

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Secondary objectives

Several planned secondary objectives were not addressed due to insufficient available data (McInnes 2020). These objectives include: evaluating the rate of positive imaging in patients with initial RT-PCR negative results who have a positive result on a follow-up RT-PCR test; determining if there is an association between number of days after symptom onset, symptom severity and the findings on thoracic imaging for patients with COVID-19; determining the rate of discrepancy or agreement between CT, chest X-ray and ultrasound findings; and determining the rate of alternative diagnoses identified by thoracic imaging.

Sensitivity analyses

We had planned to undertake additional sensitivity analyses to determine whether low risk of bias for all QUADAS-2 domains had an effect on findings.

Since all of the included studies had a high or unclear risk of bias due to study design, it was not possible to undertake these analyses.

Investigations of heterogeneity

Our protocol included additional sources to be evaluated, such as: disease prevalence, participant symptoms (severity), timing of symptom onset, participant co-morbidities and other potential candidate variables.

Due to the lack of available data, these covariates were not investigated.

Limitations of previous review and changes in this update

The initial version of this review included studies focusing on patients with either confirmed or suspected COVID-19, as well as studies including patients that were either proven to have the target condition (i.e. only sensitivity was estimated). A high proportion (almost 85%) of studies, comprised of only confirmed cases, were included and this limited our ability to evaluate both the sensitivity and specificity of the test. In this update, we only included studies focusing on patients with suspected COVID-19, from which both sensitivity and specificity estimates can be computed, as the body of evidence has grown to the point that sufficient studies meeting these preferred criteria are now available.

Investigations of variability (i.e. subgroup analyses) were limited in the initial review due to limited available data. The assessment of secondary objectives such as the impact of threshold effect (Irwig 1995), or any association between number of days after symptom onset, symptom severity and the findings on thoracic imaging for patients with COVID-19 was also not possible. In this update, we evaluated the impact of publication status (preprint versus published), but were unable to conduct further investigations of variability due to limited available data. We also began exploring the impact of threshold effects in this update, particularly that of the CO-RADS classification system, but were unable to formally evaluate the varying thresholds due to the limited number of included studies that used the CO-RADS system.

Of the studies included in the initial review, several failed to clearly report key information about their study design, as well as their methods for recruiting participants and delivering the reference standard. Therefore, data derived from these studies are likely at high risk of bias and this quality of reporting and weaknesses in the primary studies reflected the overall degree of robustness of our study. In this update, the majority of included studies also failed to report key information and had a high or unclear risk of bias with respect to participant selection, index test, reference standard, and participant flow.

The interpretation of the accuracy estimates in the previous review involved several uncertainties. While RT-PCR is considered the best available test, the results of the RT-PCR are not always sensitive; sensitivity depends on the timing of specimen collection, with high sensitivity around the onset of symptoms and during the symptomatic period but lower sensitivity before and after that window (Kucirka 2020), and collection of an appropriate specimen for testing can also be challenging. RT-PCR alone may not be the ideal reference standard (Li 2020g; Loeffelholz 2020), and it is possible that chest CT may be more sensitive than the reference standard in some patients, as some patients identified as having a false positive diagnosis on CT may have been missed by the RT-PCR test. The quality of reporting and the design of the included studies also affected the generalisability and ability to assess the validity of our findings. Because the majority of included studies recruited mainly confirmed COVID-19 cases, the accuracy of imaging tests in diagnosing COVID-19 is likely to be influenced by the prevalence of comparable viral pneumonias in a given setting. In addition, the majority of the studies included in the initial review (90%) were conducted in China, which may have impacted the generalisability of our findings. In this update, similar uncertainties with respect to the use of RT-PCR as the reference standard exist. However, this update addressed the issues involved with including studies with only confirmed cases by limiting inclusion to studies with participants that are suspected of having COVID-19. This update also includes a lower proportion of studies conducted in China compared to the previous review (53% versus 90%).

A quarter of the studies (21/84) included in the previous review were only available as preprint at the time of the search and had not yet been through the peer-review process. Data extracted from these studies will be updated and included in future versions of our review as these studies become published in peer-reviewed journals. This update includes a similar proportion of preprint studies (9/34; 26%); of the six preprint studies that were included in the previous review and also included in this update, one has been published (publication statuses are updated as of 1 October 2020).

INDEX TERMS

Medical Subject Headings (MeSH)

Bias; Case-Control Studies; COVID-19 [*diagnostic imaging]; Cross-Sectional Studies [statistics & numerical data]; Diagnostic Errors [statistics & numerical data]; Lung [diagnostic imaging]; *Radiography, Thoracic [statistics & numerical data]; Reverse Transcriptase Polymerase Chain Reaction [statistics & numerical data]; *SARS-CoV-2; Sensitivity and Specificity; *Tomography, X-Ray Computed [statistics & numerical data]; *Ultrasonography [statistics & numerical data]

MeSH check words

Adult; Child; Humans