Exhibit 420

Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease

"Based on currently available data, neither absence nor presence of signs or symptoms are accurate enough to rule in or rule out disease."

https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD013665/full?highlightAbstract=covid|diagnos|19|diagnosis|diagnosi|diagnoses

Cochrane Library

Cochrane Database of Systematic Reviews

Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease (Review)

Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MMG, Spijker R, Hooft L, Emperador D, Dittrich S, Domen J, Horn SRA, Van den Bruel A, Cochrane COVID-19 Diagnostic Test Accuracy Group

Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MMG, Spijker R, Hooft L, Emperador D, Dittrich S, Domen J, Horn SR A, Van den Bruel A.
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.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease

Thomas Struyf[1], Jonathan J Deeks[2,3], Jacqueline Dinnes[2,3], Yemisi Takwoingi[2,3], Clare Davenport[2,3], Mariska MG Leeflang[4,5], René Spijker[6,7], Lotty Hooft[7], Devy Emperador[8], Sabine Dittrich[8], Julie Domen[9], Sebastiaan R A Horn[9], Ann Van den Bruel[1], Cochrane COVID-19 Diagnostic Test Accuracy Group[3]

[1]Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium. [2]Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, Birmingham, UK. [3]NIHR Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK. [4]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands. [5]Biomarker and Test Evaluation Programme (BiTE), Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands. [6]Medical Library, Amsterdam UMC, University of Amsterdam, Amsterdam Public Health, Amsterdam, Netherlands. [7]Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands. [8]FIND, Geneva, Switzerland. [9]De Wijkpraktijk, Antwerp, Belgium

Contact address: Jonathan J Deeks, j.deeks@bham.ac.uk.

Editorial group: Cochrane Infectious Diseases Group.
Publication status and date: Edited (no change to conclusions), published in Issue 7, 2020.

Citation: Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MMG, Spijker R, Hooft L, Emperador D, Dittrich S, Domen J, Horn SR A, Van den Bruel A. 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.

ABSTRACT

Background

Some people with SARS-CoV-2 infection remain asymptomatic, whilst in others the infection can cause mild to moderate COVID-19 disease and COVID-19 pneumonia, leading some patients to require intensive care support and, in some cases, to death, especially in older adults. Symptoms such as fever or cough, and signs such as oxygen saturation or lung auscultation findings, are the first and most readily available diagnostic information. Such information could be used to either rule out COVID-19 disease, or select patients for further diagnostic testing.

Objectives

To assess the diagnostic accuracy of signs and symptoms to determine if a person presenting in primary care or to hospital outpatient settings, such as the emergency department or dedicated COVID-19 clinics, has COVID-19 disease or COVID-19 pneumonia.

Search methods

On 27 April 2020, we undertook electronic searches in the Cochrane COVID-19 Study Register and the University of Bern living search database, which is updated daily with published articles from PubMed and Embase and with preprints from medRxiv and bioRxiv. In addition, we checked repositories of COVID-19 publications. We did not apply any language restrictions.

Selection criteria

Studies were eligible if they included patients with suspected COVID-19 disease, or if they recruited known cases with COVID-19 disease and controls without COVID-19. Studies were eligible when they recruited patients presenting to primary care or hospital outpatient settings. Studies including patients who contracted SARS-CoV-2 infection while admitted to hospital were not eligible. The minimum eligible sample

size of studies was 10 participants. All signs and symptoms were eligible for this review, including individual signs and symptoms or combinations. We accepted a range of reference standards including reverse transcription polymerase chain reaction (RT-PCR), clinical expertise, imaging, serology tests and World Health Organization (WHO) or other definitions of COVID-19.

Data collection and analysis

Pairs of review authors independently selected all studies, at both title and abstract stage and full-text stage. They resolved any disagreements by discussion with a third review author. Two review authors independently extracted data and resolved disagreements by discussion with a third review author. Two review authors independently assessed risk of bias using the QUADAS-2 checklist. Analyses were descriptive, presenting sensitivity and specificity in paired forest plots, in ROC (receiver operating characteristic) space and in dumbbell plots. We did not attempt meta-analysis due to the small number of studies, heterogeneity across studies and the high risk of bias.

Main results

We identified 16 studies including 7706 participants in total. Prevalence of COVID-19 disease varied from 5% to 38% with a median of 17%. There were no studies from primary care settings, although we did find seven studies in outpatient clinics (2172 participants), and four studies in the emergency department (1401 participants). We found data on 27 signs and symptoms, which fall into four different categories: systemic, respiratory, gastrointestinal and cardiovascular. No studies assessed combinations of different signs and symptoms and results were highly variable across studies. Most had very low sensitivity and high specificity; only six symptoms had a sensitivity of at least 50% in at least one study: cough, sore throat, fever, myalgia or arthralgia, fatigue, and headache. Of these, fever, myalgia or arthralgia, fatigue, and headache could be considered red flags (defined as having a positive likelihood ratio of at least 5) for COVID-19 as their specificity was above 90%, meaning that they substantially increase the likelihood of COVID-19 disease when present.

Seven studies carried a high risk of bias for selection of participants because inclusion in the studies depended on the applicable testing and referral protocols, which included many of the signs and symptoms under study in this review. Five studies only included participants with pneumonia on imaging, suggesting that this is a highly selected population. In an additional four studies, we were unable to assess the risk for selection bias. These factors make it very difficult to determine the diagnostic properties of these signs and symptoms from the included studies.

We also had concerns about the applicability of these results, since most studies included participants who were already admitted to hospital or presenting to hospital settings. This makes these findings less applicable to people presenting to primary care, who may have less severe illness and a lower prevalence of COVID-19 disease. None of the studies included any data on children, and only one focused specifically on older adults. We hope that future updates of this review will be able to provide more information about the diagnostic properties of signs and symptoms in different settings and age groups.

Authors' conclusions

The individual signs and symptoms included in this review appear to have very poor diagnostic properties, although this should be interpreted in the context of selection bias and heterogeneity between studies. Based on currently available data, neither absence nor presence of signs or symptoms are accurate enough to rule in or rule out disease. Prospective studies in an unselected population presenting to primary care or hospital outpatient settings, examining combinations of signs and symptoms to evaluate the syndromic presentation of COVID-19 disease, are urgently needed. Results from such studies could inform subsequent management decisions such as self-isolation or selecting patients for further diagnostic testing. We also need data on potentially more specific symptoms such as loss of sense of smell. Studies in older adults are especially important.

PLAIN LANGUAGE SUMMARY

Can symptoms and medical examination accurately diagnose COVID-19 disease?

COVID-19 is an infectious disease caused by the SARS-CoV-2 virus. Most people with COVID-19 have a mild to moderate respiratory illness; others experience severe illness, such as COVID-19 pneumonia. Formal diagnosis requires laboratory analysis of nose and throat samples, or imaging tests like CT scans. However, the first and most accessible diagnostic information is from symptoms and signs from clinical examination. If initial diagnosis by symptoms and signs were accurate, the need for time-consuming, specialist diagnostic tests would be reduced.

Symptoms are experienced by patients. People with mild COVID-19 might experience cough, sore throat, high temperature, diarrhoea, headache, muscle or joint pain, fatigue, and loss of sense of smell and taste. Symptoms of COVID-19 pneumonia include breathlessness, loss of appetite, confusion, pain or pressure in the chest, and high temperature (above 38 °C).

Signs are evaluated by clinical examination, and include lung sounds, blood pressure and heart rate.

Often, people with mild symptoms visit their doctor (primary care physician) for an initial diagnosis. People with more severe symptoms might visit a hospital outpatient or emergency department. Depending on their symptoms and signs, patients may be sent home to isolate, may receive further tests or be hospitalised.

Why is accurate diagnosis important?

Accurate diagnosis ensures that people receive the correct treatment quickly; are not tested, treated or isolated unnecessarily; and do not risk spreading COVID-19. This is important for individuals and saves time and resources.

What did we want to find out?

We wanted to know how accurate diagnosis of COVID-19 and COVID-19 pneumonia is in a primary care or hospital setting, based on symptoms and signs from medical examination.

What did we do?

We searched for studies that assessed the accuracy of symptoms and signs to diagnose mild COVID-19 and COVID-19 pneumonia. Studies could include people with possible COVID-19, or people known to have – and not to have – COVID-19. Studies had to be in primary care or hospital outpatient settings only and include at least 10 participants with any symptom or sign that might be COVID-19.

The included studies

We found 16 relevant studies with 7706 participants. The studies assessed 27 separate signs and symptoms, but none assessed combinations of signs and symptoms. Seven were set in hospital outpatient clinics (2172 participants), four in emergency departments (1401 participants), but none in primary care settings. No studies included children, and only one focused on older adults. All the studies confirmed COVID-19 diagnosis by the most accurate tests available.

Main results

The studies did not clearly distinguish mild to moderate COVID-19 from COVID-19 pneumonia, so we present the results for both conditions together.

The results indicate that at least half of participants with COVID-19 disease had a cough, sore throat, high temperature, muscle or joint pain, fatigue, or headache. However, cough and sore throat were also common in people without COVID-19, so these symptoms alone are less helpful for diagnosing COVID-19. High temperature, muscle or joint pain, fatigue, and headache substantially increase the likelihood of COVID-19 disease when they are present.

How reliable are the results?

The accuracy of individual symptoms and signs varied widely across studies. Moreover, the studies selected participants in a way that meant the accuracy of tests based on symptoms and signs may be uncertain.

Conclusions

All studies were conducted in hospital outpatient settings, so the results are not representative of primary care settings. The results do not apply to children or older adults specifically, and do not clearly differentiate between milder COVID-19 disease and COVID-19 pneumonia.

The results suggest that a single symptom or sign included in this review cannot accurately diagnose COVID-19. Doctors base diagnosis on multiple symptoms and signs, but the studies did not reflect this aspect of clinical practice.

Further research is needed to investigate combinations of symptoms and signs; symptoms that are likely to be more specific, such as loss of sense of smell; and testing unselected populations, in primary care settings and in children and older adults.

How up to date is this review?

The review authors searched for studies published from January to April 2020.

SUMMARY OF FINDINGS

Summary of findings 1. Signs and symptoms to determine if a patient presenting in primary care or outpatient hospital setting has COVID-19 disease

Sign or symptom	Study design	Setting	Number of studies/number of participants	Sensitivity (ranges)	Specificity (ranges)	Strength of evidence Number of studies with high risk of bias per QUADAS-2 domain: participant selection/index test/reference standard/flow and timing
Cough	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	7/2554	0.43 to 0.71	0.14 to 0.54	3/7/1/2
		Hospital inpatients[a]	1/53	0.55	0.42	1/1/0/0
	Case-control	Primary care	-	-	-	
		Hospital outpatient clinics	1/262	0.36	0.49	0/1/0/0
		Hospital inpatients[a]	2/170	0.47 to 0.69	0.15 to 0.20	2/1/0/0
Sputum production	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	6/2467	0.16 to 0.33	0.50 to 0.86	3/6/1/2
		Hospital inpatients[a]	-	-	-	
Dyspnoea	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	7/2554	0.00 to 0.25	0.82 to 0.98	3/7/1/2
		Hospital inpatients[a]	-	-	-	

	Case-control	Primary care	-	-	-	-	
		Hospital outpatient clinics	1/262	0.12	0.77	0/1/0/0	
		Hospital inpatients[a]	-	-	-	-	
Hypoxia	Cross-sectional	Primary care	-	-	-	-	
		Hospital outpatient clinics	1/2929[b]	0.15	0.83	0/0/0/0	
		Hospital inpatients[a]	-	-	-	-	
Haemoptysis	Cross-sectional	Primary care	-	-	-	-	
		Hospital outpatient clinics	1/116	0.00	0.99	0/1/0/0	
		Hospital inpatients[a]	-	-	-	-	
Positive auscultation findings	Cross-sectional	Primary care	-	-	-	-	
		Hospital outpatient clinics	1/788	0.11	0.95	1/1/0/0	
		Hospital inpatients[a]	-	-	-	-	
	Case-control	Primary care	-	-	-	-	
		Hospital outpatient clinics	-	-	-	-	
		Hospital inpatients[a]	1/34	0.11	0.67	1/1/0/0	
Respiratory symptoms (not otherwise specified)	Cross-sectional	Primary care	-	-	-	-	
		Hospital outpatient clinics	1/788	0.04	0.95	1/1/0/0	
		Hospital inpatients[a]	-	-	-	-	
Sore throat	Cross-sectional	Primary care	-	-	-	-	
		Hospital outpatient clinics	6/2438	0.05 to 0.71	0.55 to 0.80	3/6/1/2	
		Hospital inpatients[a]	-	-	-	-	

	Case-control	Primary care	-	-	-	
		Hospital outpatient clinics	1/262	0.17	0.55	0/1/0/0
		Hospital inpatients[a]	2/170	0.13 to 0.21	0.73 to 0.91	2/1/0/0
Nasal symptoms	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	5/2405	0.00 to 0.22	0.69 to 0.92	2/5/1/2
		Hospital inpatients[a]	-	-	-	
	Case-control	Primary care	-	-	-	
		Hospital outpatient clinics	1/262	0.19	0.79	0/1/0/0
		Hospital inpatients[a]	1/136	0.03 for nasal obstruction 0.04 for rhinorrhoea	0.94 for nasal obstruction 0.95 for rhinorrhoea	1/0/0/0
Loss of sense of smell or taste	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	1/870	0.23	0.99	0/1/0/0
		Hospital inpatients[a]	-	-	-	
	Case-control	Primary care	-	-	-	
		Hospital outpatient clinics	1/262	0.22 for smell 0.20 for taste	0.96 for smell 0.95 for taste	0/1/0/0
		Hospital inpatients[a]	-	-	-	
Fever	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	8/5315	0.07 to 0.93	0.16 to 0.94	2/7/1/2
		Hospital inpatients[a]	1/53	0.80	0.48	1/1/0/0
	Case-control	Primary care	-	-	-	

		Hospital outpatient clinics	1/262	0.54	0.74	0/1/0/0
		Hospital inpatients[a]	1/34	0.79	0.07	1/1/0/0
Low body temperature	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	1/2929[b]	0.37	0.32	0/0/0/0
		Hospital inpatients[a]	-	-	-	
Shivers	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	1/132	0.14	0.86	0/1/1/1
		Hospital inpatients[a]	-	-	-	
Chills	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	2/1443	0.07 to 0.29	0.72 to 0.91	0/2/1/1
		Hospital inpatients[a]	-	-	-	
Myalgia or arthralgia	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	4/339	0.19 to 0.86	0.45 to 0.91	2/4/1/2
		Hospital inpatients[a]	-	-	-	
	Case-control	Primary care	-	-	-	
		Hospital outpatient clinics	1/262	0.34	0.81	0/1/0/0
		Hospital inpatients[a]	-	-	-	
Myalgia or fatigue	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	2/1427	0.16 to 0.31	0.82 to 0.93	0/2/0/0
		Hospital inpatients[a]	-	-	-	
Fatigue	Cross-sectional	Primary care	-	-	-	

		Hospital outpatient clinics	2/220	0.43 to 0.57	0.60 to 0.67	1/2/1/2
		Hospital inpatients[a]	1/53	0.10	0.94	1/1/0/0
	Case-control	Primary care	-	-	-	
		Hospital outpatient clinics	1/262	0.42	0.69	0/1/0/0
		Hospital inpatients[a]	2/170	0.11 to 0.31	0.88 to 1.00	2/1/0/0
Headache	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	4/1647	0.03 to 0.71	0.78 to 0.98	1/4/1/2
		Hospital inpatients[a]	1/53	0.15	0.97	1/1/0/0
Nausea/vomiting	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	2/436	0.00 to 0.04	0.97 to 0.97	0/2/1/1
		Hospital inpatients[a]	1/53	0.05	1.00	1/1/0/0
	Case-control	Primary care	-	-	-	
		Hospital outpatient clinics	2/778	0.05 to 0.23	0.81 to 0.96	0/2/0/0
		Hospital inpatients[a]	-	-	-	
Diarrhoea	Cross-sectional	Primary care	-	-	-	
		Hospital outpatient clinics	5/1680	0.00 to 0.14	0.86 to 0.99	2/5/1/2
		Hospital inpatients[a]	1/53	0.15	0.88	1/1/0/0
	Case-control	Primary care	-	-	-	
		Hospital outpatient clinics	2/778	0.08 to 0.20	0.85 to 0.92	0/2/0/0
		Hospital inpatients[a]	1/34	0.05	0.93	1/1/0/0
Abdominal pain	Cross-sectional	Primary care	-	-	-	

		Hospital outpatient clinics	1/132	0.00	0.96	0/1/1/1	
		Hospital inpatients[a]	1/53	0.05	1.00	1/1/0/0	
Gastrointestinal symptoms	Cross-sectional	Primary care	-	-	-		
		Hospital outpatient clinics	1/788	0.37	0.68	1/1/0/0	
		Hospital inpatients[a]	-	-	-		
	Case-control	Primary care	-	-	-		
		Hospital outpatient clinics	1/516	0.35	0.74	0/1/0/0	
		Hospital inpatients[a]	-	-	-		
Low systolic blood pressure	Cross-sectional	Primary care	-	-	-		
		Hospital outpatient clinics	1/3341[b]	0.11	0.90	0/0/0/0	
		Hospital inpatients[a]	-	-	-		
High systolic blood pressure	Cross-sectional	Primary care	-	-	-		
		Hospital outpatient clinics	1/3341	0.39	0.57	0/0/0/0	
		Hospital inpatients[a]	-	-	-		
Tachycardia	Cross-sectional	Primary care	-	-	-		
		Hospital outpatient clinics	1/3373	0.47	0.62	0/0/0/0	
		Hospital inpatients[a]	-	-	-		
Palpitations	Cross-sectional	Primary care	-	-	-		
		Hospital outpatient clinics	1/132	0.00	0.98	0/1/1/1	
		Hospital inpatients[a]	-	-	-		
Chest tightness	Cross-sectional	Primary care	-	-	-		

Hospital outpatient clinics	-	-	-	-	
Hospital inpatients[a]	-	-	-	-	
Case-control					
Primary care	-	-	-	-	
Hospital outpatient clinics	-	-	-	-	
Hospital inpatients[a]	1/34	0.05	1.00	1/1/0/0	

[a]'Hospital inpatients' refers to studies that recruited patients admitted to hospital with COVID-19 disease and in whom the signs and symptoms were assessed on admission.
[b]Setting not specified; assumed hospital outpatients considering the timing in the epidemic and sparse testing capacity outside hospitals at the time (Rentsch 2020).

BACKGROUND

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and resulting COVID-19 pandemic present important diagnostic evaluation challenges. These range from, on the one hand, understanding the value of signs and symptoms in predicting possible infection, assessing whether existing biochemical and imaging tests can identify infection and recognise patients needing critical care, and on the other hand, evaluating whether new diagnostic tests can allow accurate rapid and point-of-care testing. Also, the diagnostic aims are diverse, including identifying current infection, ruling out infection, identifying people in need of care escalation, or testing for past infection.

This review is part of a cluster of reviews on the diagnosis of SARS-CoV-2 infection and COVID-19 disease, and deals solely with the diagnostic accuracy of presenting clinical signs and symptoms for diagnosing COVID-19 disease.

Target condition being diagnosed

COVID-19 is the disease caused by infection with the SARS-CoV-2 virus. SARS-CoV-2 infection is diagnosed with reverse transcription polymerase chain reaction (RT-PCR), which is a test that detects the virus' genetic material, with imaging to identify lung abnormalities and with clinical signs and symptoms.

SARS-CoV-2 infection can be asymptomatic (no symptoms); mild or moderate; severe (causing breathlessness and increased respiratory rate indicative of pneumonia and oxygen need); or critical (requiring intensive support due to severe acute respiratory syndrome (SARS) or acute respiratory distress syndrome (ARDS), shock or other organ dysfunction). People with COVID-19 pneumonia (severe or critical disease), require different patient management, which makes it important to distinguish between mild or moderate COVID-19 disease and COVID-19 pneumonia.

In this review, we will examine the diagnostic value of signs and symptoms for symptomatic SARS-CoV-2 infection, which includes mild or moderate COVID-19 disease and COVID-19 pneumonia.

In planning review updates, we will consider the potential addition of another grouping, which is a subset of the above:

- whether tests exist that identify people requiring respiratory support (SARS or ARDS) or intensive care.

Index test(s)

Signs and symptoms

Signs and symptoms are used in the initial diagnosis of suspected COVID-19 disease, and to identify people with COVID-19 pneumonia. Symptoms are what is experienced by patients, for example cough or nausea. Signs are what can be evaluated by clinical assessment, for example lung auscultation findings, blood pressure or heart rate.

Key symptoms that have been associated with mild to moderate COVID-19 disease include: troublesome dry cough (for example, coughing more than usual over a one-hour period, or three or more coughing episodes in 24 hours), fever greater than 37.8 °C, diarrhoea, headache, breathlessness on light exertion, muscle pain, fatigue, and loss of sense of smell and taste. Red flags indicating possible pneumonia include breathlessness at rest, loss of appetite, confusion, pain or pressure in the chest, and temperature above 38 °C.

Clinical pathway

Important in the context of COVID-19 is that the pathway is multifaceted because it is designed to care for the diseased individual and to protect the community from further spread. Decisions about patient and isolation pathways for COVID-19 vary according to health services and settings, available resources, and stages of the epidemic. They will change over time, if and when 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 be able to inform rational decision making.

Prior test(s)

In this review on signs and symptoms, no prior tests are required because signs and symptoms are used in the initial diagnosis of suspected COVID-19 disease. Patients can, however, self-assess before presenting to healthcare services based on their symptoms. This is in contrast to contact tracing, in which patients or participants are tested based on a documented contact with a SARS-CoV-2-positive person and may themselves be asymptomatic.

Role of index test(s)

Signs and symptoms are used as triage tests, that is, to rule out COVID-19 disease, but also to identify patients with possible COVID-19 who may require further testing, care escalation or isolation.

Alternative test(s)

Chest X-ray, ultrasound, and computed tomography (CT) are widely used diagnostic imaging tests to diagnose COVID-19 pneumonia. Availability and usage varies between settings. We address these radiological tests in a separate review.

Rationale

It is essential to understand the accuracy of diagnostic tests including signs and symptoms to identify the best way they can be used in different settings to develop effective diagnostic and management pathways. We are producing a suite of Cochrane 'living systematic reviews', which will summarise evidence on the clinical accuracy of different tests and diagnostic features, grouped according to present research questions and settings, in the diagnosis of SARS-CoV-2 infection and COVID-19 disease. Summary estimates of accuracy from these reviews will help inform diagnostic, screening, isolation, and patient management decisions.

New tests are being developed and evidence is emerging at an unprecedented rate during the COVID-19 pandemic. We will aim to update these reviews as often as is feasible to ensure that they provide the most up-to-date evidence about test accuracy.

These reviews are being produced rapidly to assist in providing a central resource of evidence to assist in the COVID-19 pandemic, summarising available evidence on the accuracy of the tests and presenting characteristics.

OBJECTIVES

To assess the diagnostic accuracy of signs and symptoms to determine if a person presenting in primary care or to hospital outpatient settings, such as the emergency department or dedicated COVID-19 clinics, has COVID-19 disease or COVID-19 pneumonia.

Secondary objectives

Where data are available, we will investigate diagnostic accuracy (either by stratified analysis or meta-regression) according to:

- days since symptom onset, population (children; older adults), reference standard, study design, setting, severity of COVID-19 pneumonia (severe COVID-19 pneumonia/ARDS requiring intensive care support).

METHODS

Criteria for considering studies for this review

Types of studies

We kept the eligibility criteria purposely broad to include all patient groups and all variations of a test at this initial stage of reviewing the evidence (that is, if the patient population was unclear, we included the study).

We included studies of all designs that produce estimates of test accuracy or provide data from which estimates can be computed. We included both single-gate (studies that recruit from a patient pathway before disease status has been ascertained) and multi-gate (where people with and without the target condition are recruited separately) designs. This means that we included studies that were cross-sectional or diagnostic case-control type studies.

When interpreting the results, we made sure that the limitations of different study designs were carefully considered, using quality assessment and analysis.

Participants

Studies recruiting people presenting to primary care or outpatient hospital settings with suspicion of COVID-19 disease were eligible.

For the initial version of this review, we included studies that recruited symptomatic people either known to have SARS-CoV-2 infection or known not to have SARS-CoV-2 infection.

Studies had to have a sample size of a minimum of 10 participants.

Index tests

- All signs and symptoms, including:
  - signs such as oxygen saturation, measured by oximetry or blood pressure;
  - classic symptoms, such as fever or cough.
- We included combinations of signs and symptoms, but not when they were combined with laboratory, imaging, or other types of index tests as these will be covered in the other reviews.

Target conditions

To be eligible studies had to identify at least one of:

- mild or moderate COVID-19 disease;
- COVID-19 pneumonia.

Asymptomatic infection with SARS-CoV-2 infection is out of scope for this review, considering it is by definition not possible to detect this based on signs and symptoms.

Reference standards

We anticipated that studies would use a range of reference standards. Although RT-PCR is considered the best available test, due to rapidly evolving knowledge about the target conditions, multiple reference standards on their own as well as in combination have emerged.

We expected to encounter cases defined by:

- RT-PCR alone;
- RT-PCR, clinical expertise, and imaging (for example, CT thorax);
- repeated RT-PCR several days apart or from different samples;
- plaque reduction neutralisation test (PRNT) or enzyme-linked immunosorbent assay**(**ELISA) tests;
- information available at a subsequent time point;
- World Health Organization (WHO) and other case definitions (see Appendix 1).

This list is not exhaustive, and we recorded all reference standards encountered. With a group of methodological and clinical experts, we are producing a ranking of reference standards according to their ability to correctly classify participants using a consensus process. We will use the ranking for informing the assessment of methodological quality in the next update of this review.

Search methods for identification of studies

The final search date for this version of the review is 27 April 2020.

Electronic searches

We conducted a single literature search to cover our suite of Cochrane COVID-19 diagnostic test accuracy (DTA) reviews (Deeks 2020; McInnes 2020).

We conducted electronic searches using two primary sources. Both of these searches aimed to identify all published articles and preprints related to COVID-19, and were not restricted to those evaluating biomarkers or tests. Thus, there are no test terms, diagnosis terms, or methodological terms in the searches. Searches were limited to 2019 and 2020, and for this version of the review have been conducted to 27 April 2020.

Cochrane COVID-19 Study Register searches

We used the Cochrane COVID-19 Study Register (covid-19.cochrane.org/), for searches conducted to 28 March 2020. At that time, the register was populated by searches of PubMed, as well as trials registers at ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform (ICTRP).

Search strategies were designed for maximum sensitivity, to retrieve all human studies on COVID-19 and with no language limits. See Appendix 2.

COVID-19 Living Evidence Database from the University of Bern

From 28 March 2020, we used the COVID-19 Living Evidence database from the Institute of Social and Preventive Medicine (ISPM) at the University of Bern (www.ispm.unibe.ch), as the primary source of records for the Cochrane COVID-19 DTA reviews. This search includes PubMed, Embase, and preprints indexed in bioRxiv and medRxiv databases. The strategies as described on the ISPM website are described here (ispmbern.github.io/covid-19/). See Appendix 3.

The decision to focus primarily on the 'Bern' feed was due to the exceptionally large numbers of COVID-19 studies available only as preprints. The Cochrane COVID-19 Study Register has undergone a number of iterations since the end of March 2020 and we anticipate moving back to the Cochrane COVID-19 Study Register as the primary source of records for subsequent review updates.

Searching other resources

We obtained Embase records through Martha Knuth for the Centers for Disease Control and Prevention (CDC), Stephen B Thacker CDC Library, COVID-19 Research Articles Downloadable Database and de-duplicated them against the Cochrane COVID-19 Study Register up to 1 April 2020. See Appendix 4.

We also checked our search results against two additional repositories of COVID-19 publications including:

- the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) 'COVID-19: Living map of the evidence' (eppi.ioe.ac.uk/COVID19_MAP/covid_map_v4.html);
- the Norwegian Institute of Public Health 'NIPH systematic and living map on COVID-19 evidence' (www.nornesk.no/forskningskart/NIPH_diagnosisMap.html)

Both of these repositories allow their contents to be filtered according to studies potentially relating to diagnosis, and both have agreed to provide us with updates of new diagnosis studies added. For this iteration of the review, we examined all diagnosis studies from both sources up to 16 April 2020.

We did not apply any language restrictions.

Data collection and analysis

Selection of studies

Pairs of review authors independently screened studies. We resolved disagreements by discussion with a third, experienced review author for initial title and abstract screening, and through discussion between three review authors for eligibility assessments.

Data extraction and management

Pairs of review authors independently performed data extraction. We resolved disagreements by discussion between three review authors.

We intended to contact study authors where we needed to clarify details or obtain missing information.

Assessment of methodological quality

Pairs of review authors independently assessed risk of bias and applicability concerns using the QUADAS-2 (Quality Assessment tool for Diagnostic Accuracy Studies) checklist, which was common to the suite of reviews but tailored to each particular review (Whiting 2011; Table 1). For this review, we excluded the questions on the nature of the samples as these were not relevant, and we added a question on who assessed the signs. We resolved disagreements by discussion between three review authors.

Statistical analysis and data synthesis

We present results of estimated sensitivity and specificity using paired forest plots and summarised in tables as appropriate.

We present the results without meta-analysis, due to the small numbers of studies currently available, considerable heterogeneity across studies and the high risk of bias that we identified, as we felt doing so would otherwise produce a seemingly more accurate estimate than the underlying evidence is able to provide at this moment in time.

We present results of estimated sensitivity and specificity using paired forest plots in Review Manager 2014, and dumbbell plots to display the change in disease probability after a positive or negative result.

We disaggregated data by study design and organised by target condition, reporting results from cross-sectional studies separately from studies that used a diagnostic case-control or other design that we assessed as prone to high risk of bias.

When pooling does become possible in a future update of this review, we will estimate mean sensitivity and specificity using hierarchical models where tests either report binary results or at commonly reported thresholds. Where data are sparse, we will use methods described by Takwoingi 2017 for obtaining estimates from simplified models. We anticipate that over time sufficient data will accumulate to provide clear estimates of test accuracy for some tests. We will undertake meta-analysis in STATA version 16.0 (STATA), or SAS (SAS 2015), as detailed in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (Chapter 10; Macaskill 2013).

Investigations of heterogeneity

We have listed sources of heterogeneity that we investigated if adequate data were available in the Secondary objectives. In this version of the review, we used stratification to investigate heterogeneity as we considered it was inappropriate to combine studies. In future updates, if meta-analysis becomes possible, we will investigate heterogeneity through meta-regression.

We will stratify by reference standard and study design. In this version of the review we have stratified by study design only, as stratification by reference standard was not yet possible.

Sensitivity analyses

We aimed to undertake sensitivity analyses considering the impact of:

- unpublished studies;
- studies with inadequate reference standards.

However, neither were possible in this version of the review.

Assessment of reporting bias

We aimed to publish lists of studies that we know exist but for which we have not managed to locate reports, and request information to include in updates of these reviews. However, at the time of writing this version of the review, we are unaware of unpublished studies.

Summary of findings

We have listed our key findings in a 'Summary of findings' table to determine the strength of evidence for each test and findings, and to highlight important gaps in the evidence.

Updating

We will undertake the searches of published literature and preprints bi-weekly, and, dependent on the number of new and important studies found, we will consider updating each review with each search if resources allow.

RESULTS

Results of the search

The search yielded 10,965 records after removing duplicates. The first selection resulted in 658 records that were potentially eligible for this review on signs and symptoms. After screening on title and abstract, we excluded 457 records, leaving 201 to be assessed on full text. Of these, we included 16 studies in this review. The reasons for excluding 185 records are listed in the PRISMA flow chart (see Figure 1; Moher 2009).

Figure 1. Flow diagram

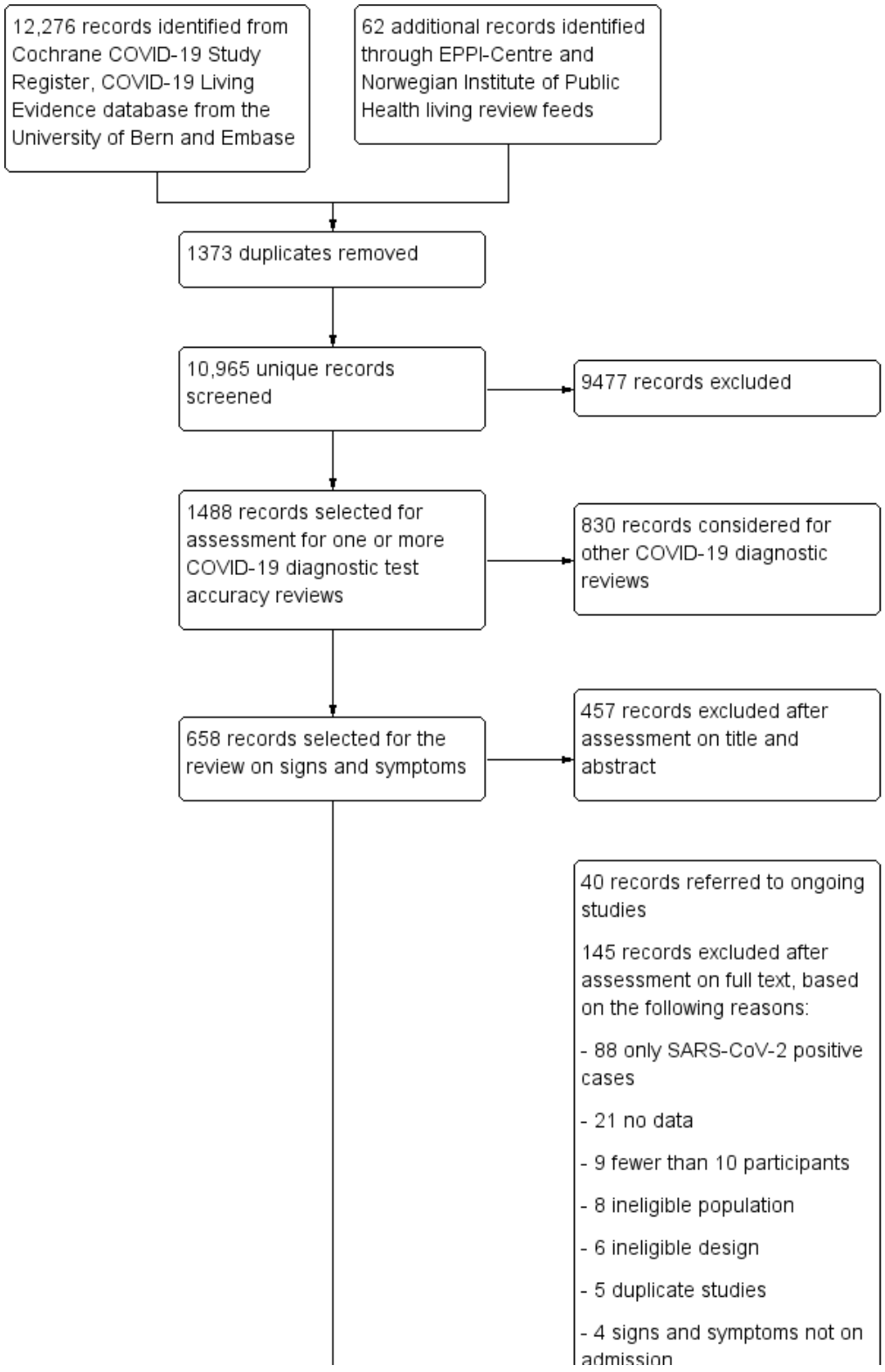
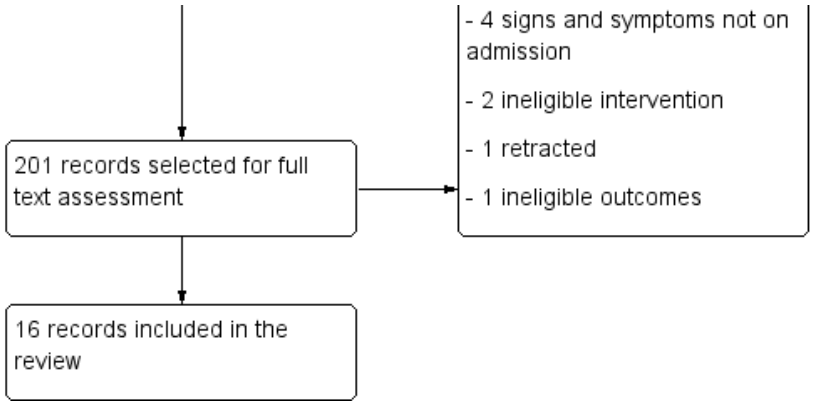


Figure 1. (Continued)



Two studies reported on the same cases while using a different control group (Chen X 2020; Yang 2020d). Chen X 2020 used a concurrent control group of pneumonia cases negative for SARS-CoV-2 on PCR testing but Yang 2020d used a historic control group of influenza pneumonia patients. For this reason we only included the Chen X 2020 results in the analyses.

One study reported a study that included a derivation and validation part for the development of a prediction rule (Song 2020b). The two parts are identical in set-up and only differ in respect to the time of data collection, that is, the derivation part recruited participants up to 5 February 2020 and the validation part recruited participants from 6 February 2020 onwards. As a result, we consider this to be one study and have entered all data on signs and symptoms as such.

Four studies were conducted in the USA, all other studies were from China. A summary of the main study characteristics can be found in Table 2.

Methodological quality of included studies

The results of the quality assessment are summarised in Figure 2 and Figure 3. We rated participant selection as introducing high risk of bias in seven studies. In five studies this was because a CT scan or other imaging was used to diagnose patients with pneumonia prior to inclusion in the study, leading to a highly selected patient population (Ai 2020a; Chen X 2020; Cheng 2020a; Liang 2020; Yang 2020d); RT-PCR results were subsequently used to distinguish between COVID-19 pneumonia and pneumonia from other causes. For all studies, testing was highly dependent on the local case definition and testing criteria that were in effect at the time of the study, meaning all patients that were included in studies had already gone through a referral/selection filter, which was not always described. The most extreme example of this is the study by Liang 2020, in which patients with radiological evidence of pneumonia and a clinical presentation compatible with COVID-19 were only tested for SARS-CoV-2 after a panel discussion.

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

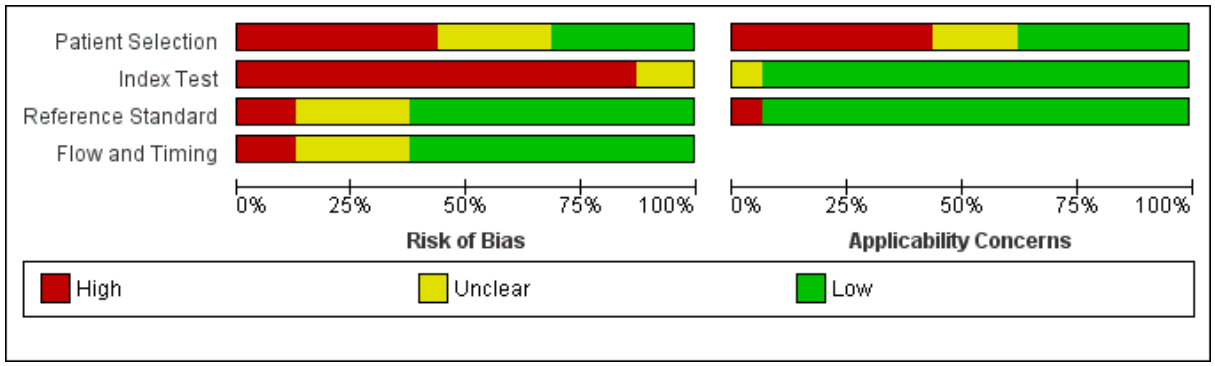


Figure 3. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study

Of the 16 studies included in this first version of the review, five studies did not use a cross-sectional design. Three studies were diagnostic case-control studies (Nobel 2020; Yang 2020d; Zhao 2020a), one study selected cases cross-sectionally in five hospitals but only selected cases in one hospital (Chen X 2020), and one study emailed patients who had undergone testing for SARS-CoV-2 about olfactory symptoms prior to the SARS-CoV-2 test, with a response rate of 58% in SARS-CoV-2 positive cases and 15% in negative cases (Yan 2020a).

We rated all studies except two as carrying a high risk of bias for the index tests because there was little to no detail on how, by whom, and when the signs and symptoms were measured. In addition, there is considerable uncertainty around the reference standard, with some studies providing little detail on the RT-PCR tests that they used or lack of clarity on blinding.

Participant flow was unclear in four studies (Yan 2020a; Yang 2020d; Zhao 2020a; Zhu 2020b), either because the timing of recording signs and symptoms and conduct of the reference standard was unclear, or because some tests received a second or third reference standard at unclear time points during hospital admission.

We rated applicability for participant selection as high risk when there was a risk of selection bias or studies did not describe selection. As for the applicability of the index tests and reference standard, we always scored this as low risk except for Chen X 2020, because blinding of the index tests was unclear, and Yang 2020d, because blinding and sample of the reference standard were unclear.

Findings

The main characteristics of all included studies are listed in Table 2. There were four studies in hospital inpatients (Ai 2020a; Chen X 2020; Yang 2020d; Zhao 2020a), seven studies in hospital outpatients (Cheng 2020a; Liang 2020; Nobel 2020; Peng 2020a; Song 2020b; Sun 2020a; Yan 2020a), and four studies in emergency departments (Feng 2020a; Tolia 2020; Wee 2020; Zhu 2020b). The setting was not specified in one study (Rentsch 2020); in the 'Summary of findings' table, we classified this study setting as being hospital outpatient under the assumption that at that time in the pandemic (February 2020 to March 2020) tests were not commonly available outside hospital clinics. There were no studies conducted in community primary care services.

Seven studies assessed the accuracy of signs and symptoms for the diagnosis of COVID-19 pneumonia (Ai 2020a; Chen X 2020; Cheng 2020a; Feng 2020a; Liang 2020; Yang 2020d; Zhao 2020a); the remaining studies had COVID-19 disease as the target condition, with no further description of the severity, meaning some patients could have suffered from mild or moderate COVID-19 disease and others from COVID-19 pneumonia. The distinction between these two target conditions was not always very clear, and a degree of overlap is to be assumed. All studies used RT-PCR testing as the reference standard, with some variation in the samples that were used.

In all, 7706 patients were included, the median number of participants was 134. Prevalence of infection varied from 5% to 38% with a median of 17%. There were no studies in children or elderly populations, except for Rentsch 2020, which included a cohort of a median age of 65.7 years old from the Veterans Affairs Healthcare System database.

We found data on 27 signs and symptoms, which fall into four different categories: systemic, respiratory, gastrointestinal and cardiovascular signs and symptoms. There were no analyses for combinations of tests, only for individual signs and symptoms. The results are summarised in Table 2. Results for the cross-sectional studies are presented in forest plots (Figure 4; Figure 5; Figure 6; Figure 7), and are plotted in ROC (receiver operating characteristic) space (Figure 8; Figure 9; Figure 10; Figure 11), results for the other studies are only listed in forest plots (Figure 12; Figure 13; Figure 14; Figure 15).

Figure 4. Forest plot of respiratory signs and symptoms (cross-sectional studies)

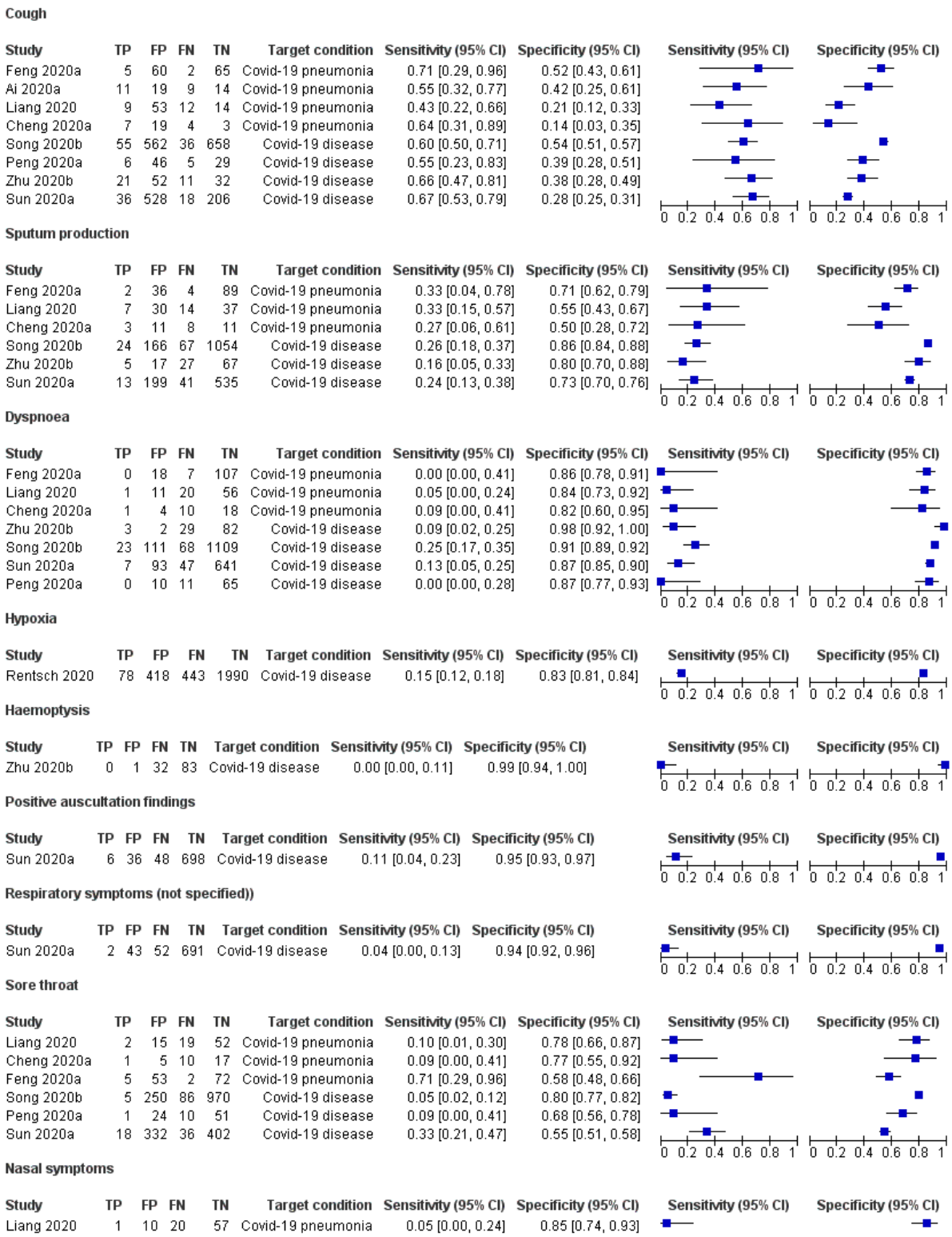
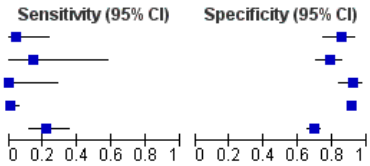


Figure 4. (Continued)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Liang 2020	1	10	20	57	Covid-19 pneumonia	0.05 [0.00, 0.24]	0.85 [0.74, 0.93]
Feng 2020a	1	27	6	98	Covid-19 pneumonia	0.14 [0.00, 0.58]	0.78 [0.70, 0.85]
Peng 2020a	0	6	11	69	Covid-19 disease	0.00 [0.00, 0.28]	0.92 [0.83, 0.97]
Song 2020b	1	107	90	1113	Covid-19 disease	0.01 [0.00, 0.06]	0.91 [0.90, 0.93]
Sun 2020a	12	226	42	508	Covid-19 disease	0.22 [0.12, 0.36]	0.69 [0.66, 0.73]



Loss of smell (anosmia) or loss of taste (ageusia)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Wee 2020	35	9	119	707	Covid-19 disease	0.23 [0.16, 0.30]	0.99 [0.98, 0.99]

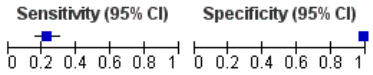
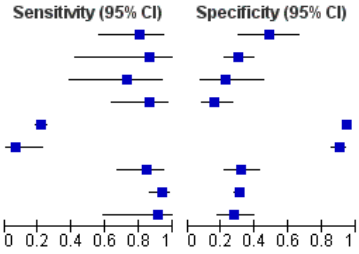


Figure 5. Forest plot of systemic signs and symptoms (cross-sectional studies)

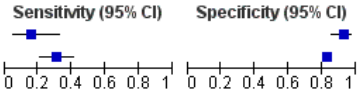
Fever

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	16	17	4	16	Covid-19 pneumonia	0.80 [0.56, 0.94]	0.48 [0.31, 0.66]
Feng 2020a	6	87	1	38	Covid-19 pneumonia	0.86 [0.42, 1.00]	0.30 [0.22, 0.39]
Cheng 2020a	8	17	3	5	Covid-19 pneumonia	0.73 [0.39, 0.94]	0.23 [0.08, 0.45]
Liang 2020	18	56	3	11	Covid-19 pneumonia	0.86 [0.64, 0.97]	0.16 [0.08, 0.27]
Rentsch 2020	120	169	431	2664	Covid-19 disease	0.22 [0.18, 0.25]	0.94 [0.93, 0.95]
Tolia 2020	2	25	27	227	Covid-19 disease	0.07 [0.01, 0.23]	0.90 [0.86, 0.93]
Zhu 2020b	27	57	5	27	Covid-19 disease	0.84 [0.67, 0.95]	0.32 [0.22, 0.43]
Song 2020b	85	844	6	376	Covid-19 disease	0.93 [0.86, 0.98]	0.31 [0.28, 0.33]
Peng 2020a	10	54	1	21	Covid-19 disease	0.91 [0.59, 1.00]	0.28 [0.18, 0.40]



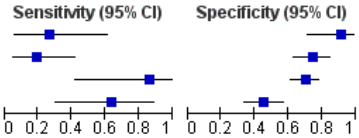
Myalgia or fatigue

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Zhu 2020b	5	6	27	78	Covid-19 disease	0.16 [0.05, 0.33]	0.93 [0.85, 0.97]
Song 2020b	28	214	63	1006	Covid-19 disease	0.31 [0.22, 0.41]	0.82 [0.80, 0.85]



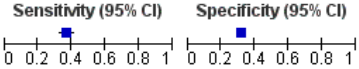
Myalgia or arthralgia

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Cheng 2020a	3	2	8	20	Covid-19 pneumonia	0.27 [0.06, 0.61]	0.91 [0.71, 0.99]
Liang 2020	4	17	17	50	Covid-19 pneumonia	0.19 [0.05, 0.42]	0.75 [0.63, 0.84]
Feng 2020a	6	37	1	88	Covid-19 pneumonia	0.86 [0.42, 1.00]	0.70 [0.62, 0.78]
Peng 2020a	7	41	4	34	Covid-19 disease	0.64 [0.31, 0.89]	0.45 [0.34, 0.57]



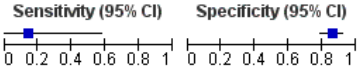
Low body temperature

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Rentsch 2020	204	1938	347	895	Covid-19 disease	0.37 [0.33, 0.41]	0.32 [0.30, 0.33]



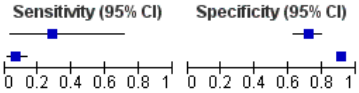
Shivers

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Feng 2020a	1	17	6	108	Covid-19 pneumonia	0.14 [0.00, 0.58]	0.86 [0.79, 0.92]



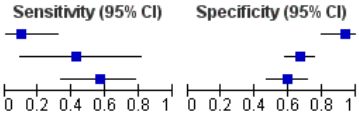
Chills

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Feng 2020a	2	35	5	90	Covid-19 pneumonia	0.29 [0.04, 0.71]	0.72 [0.63, 0.80]
Song 2020b	6	111	85	1109	Covid-19 disease	0.07 [0.02, 0.14]	0.91 [0.89, 0.92]



Fatigue

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	2	2	18	31	Covid-19 pneumonia	0.10 [0.01, 0.32]	0.94 [0.80, 0.99]
Feng 2020a	3	41	4	84	Covid-19 pneumonia	0.43 [0.10, 0.82]	0.67 [0.58, 0.75]
Liang 2020	12	27	9	40	Covid-19 pneumonia	0.57 [0.34, 0.78]	0.60 [0.47, 0.72]



Headache

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	3	1	17	32	Covid-19 pneumonia	0.15 [0.03, 0.38]	0.97 [0.84, 1.00]
Feng 2020a	5	23	2	102	Covid-19 pneumonia	0.71 [0.29, 0.96]	0.82 [0.74, 0.88]
Liang 2020	8	15	13	52	Covid-19 pneumonia	0.38 [0.18, 0.62]	0.78 [0.66, 0.87]
Zhu 2020b	1	2	31	82	Covid-19 disease	0.03 [0.00, 0.16]	0.98 [0.92, 1.00]
Song 2020b	9	158	82	1062	Covid-19 disease	0.10 [0.05, 0.18]	0.87 [0.85, 0.89]

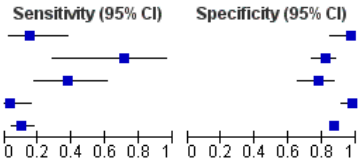


Figure 6. Forest plot of gastrointestinal signs and symptoms (cross-sectional studies)

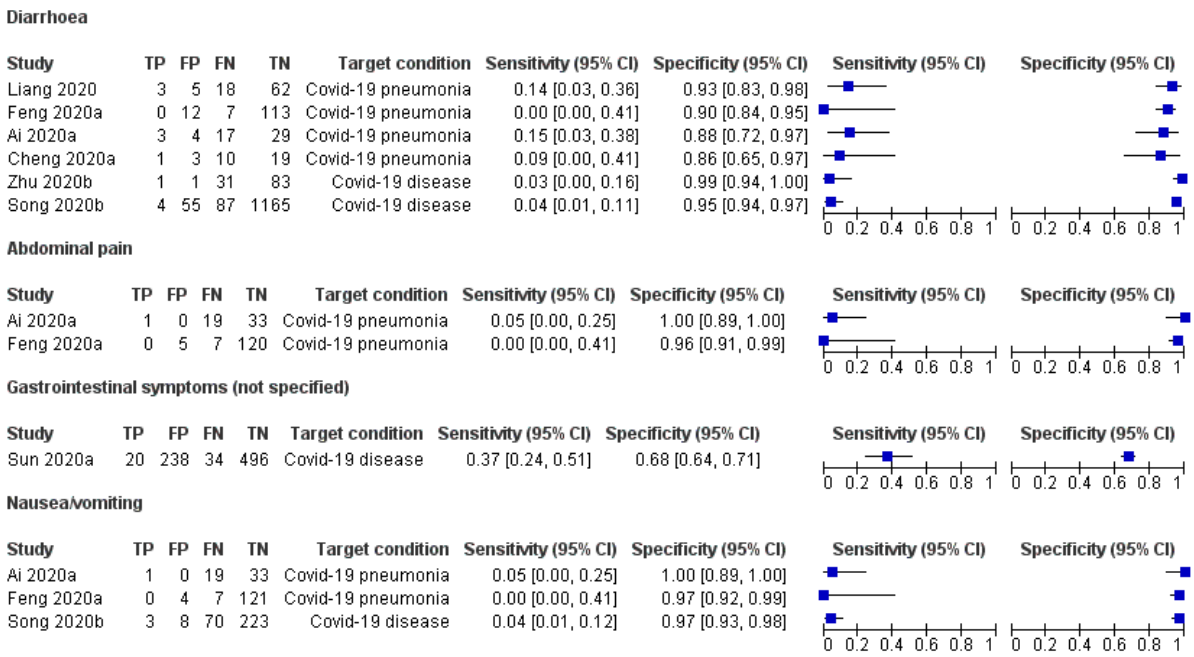


Figure 7. Forest plot of cardiovascular signs and symptoms (cross-sectional studies)

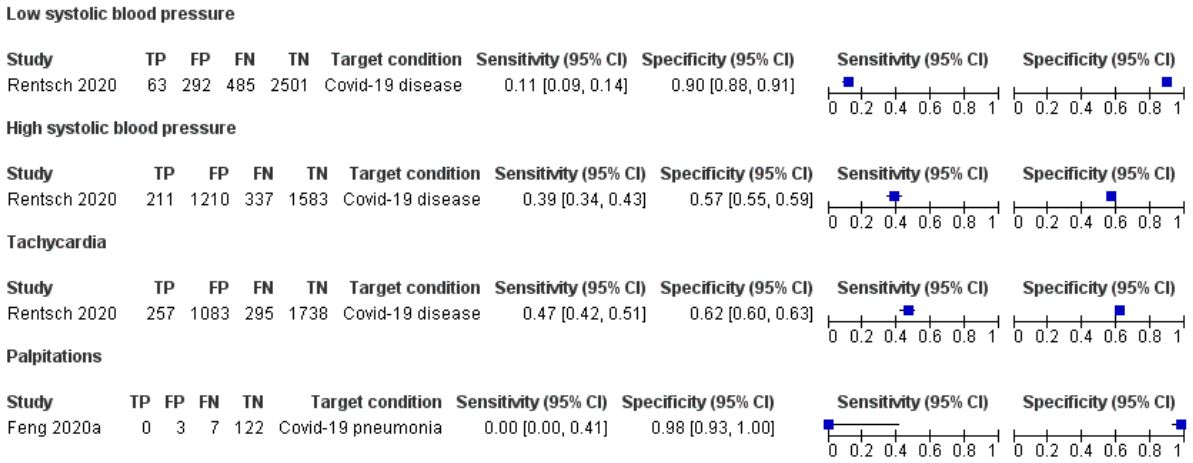


Figure 8. Summary ROC plot of respiratory signs and symptoms (cross-sectional studies)

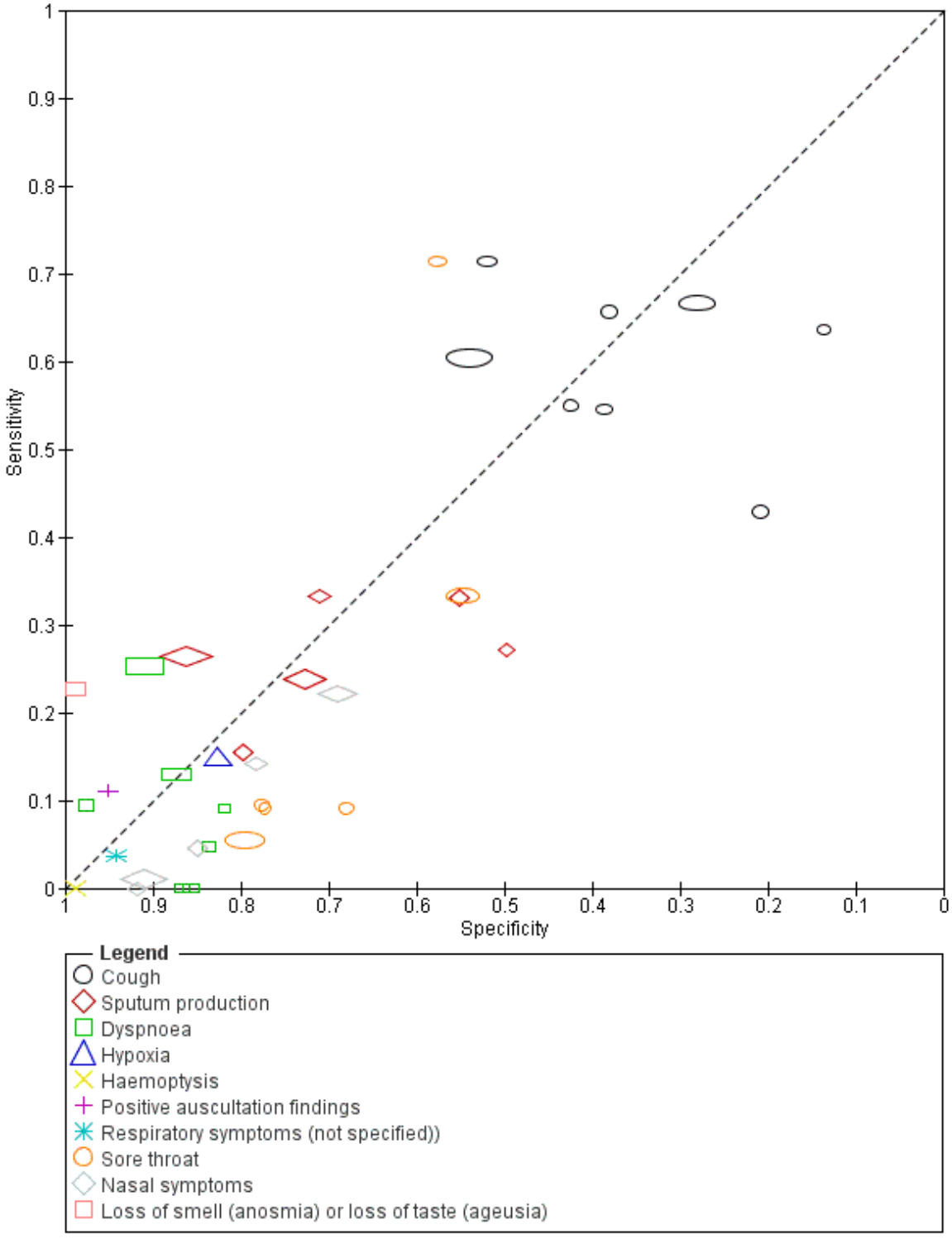


Figure 9. Summary ROC plot of systemic signs and symptoms (cross-sectional studies)

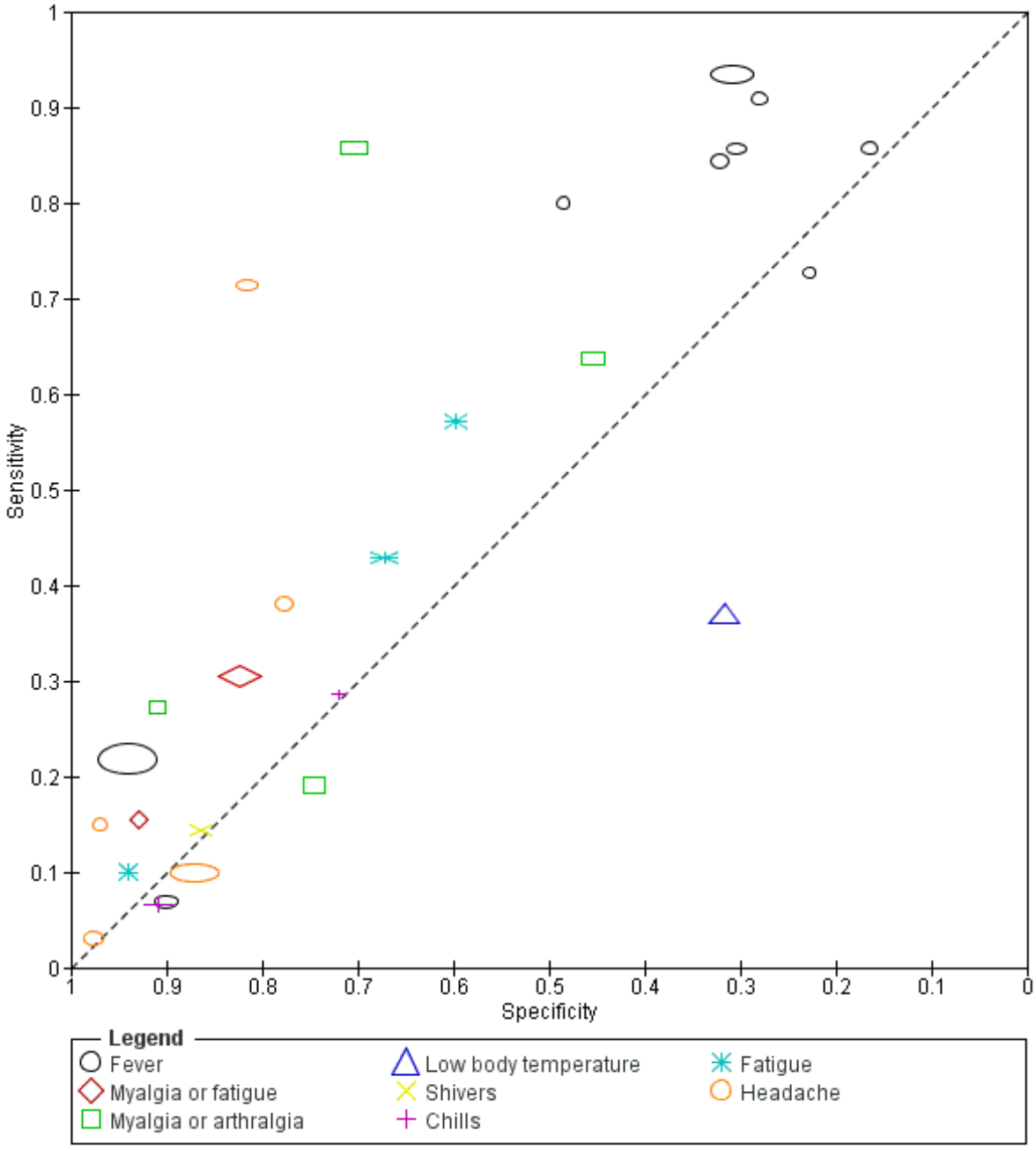


Figure 10. Summary ROC Plot of gastrointestinal signs and symptoms (cross-sectional studies)

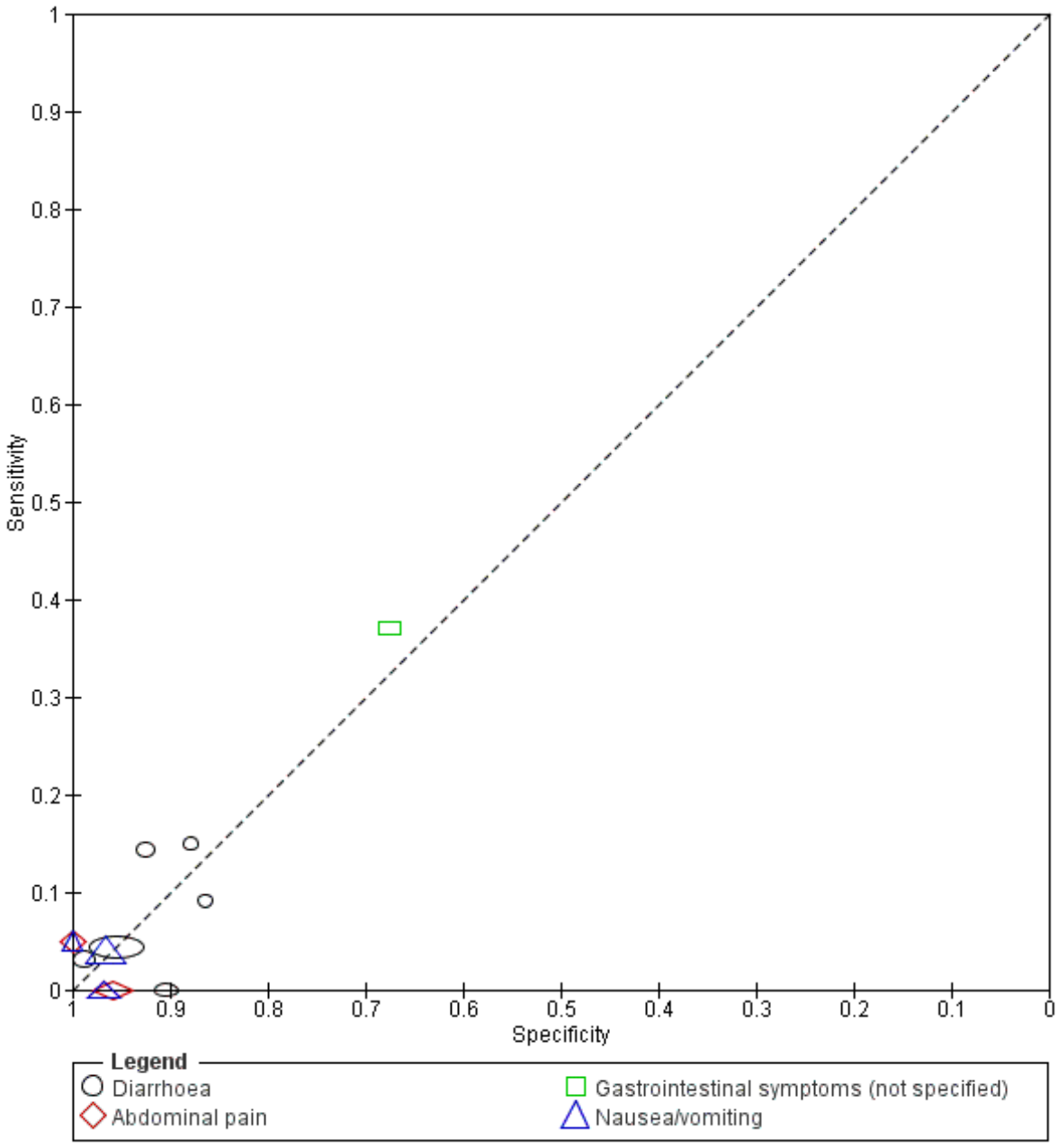


Figure 11. Summary ROC plot of cardiovascular signs and symptoms (cross-sectional studies)

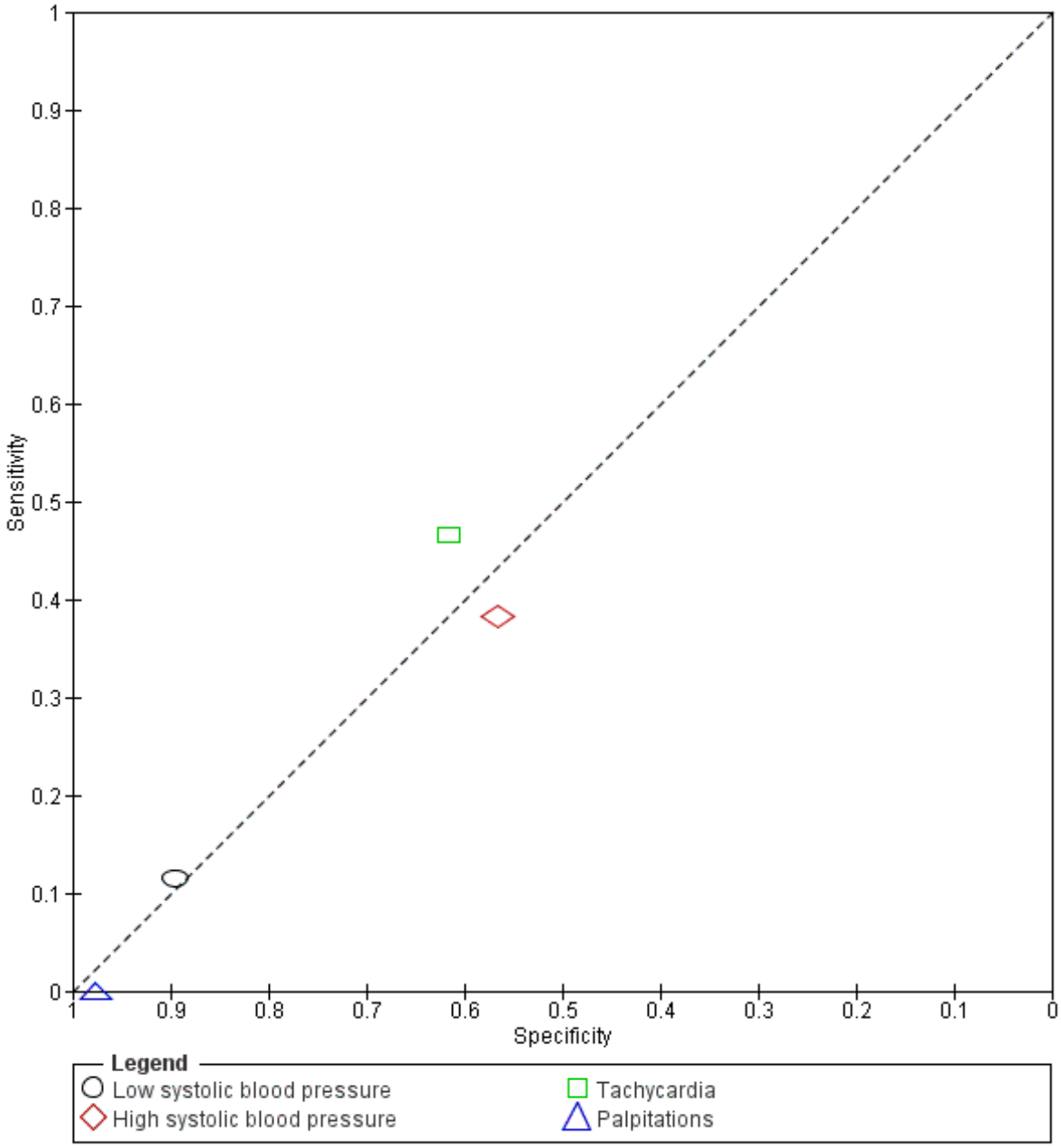
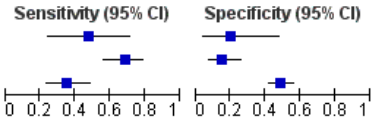


Figure 12. Forest plot of tests: 27 cough (non-cross-sectional study), 28 sore throat (non-cross-sectional study), 29 rhinorrhoea (non-cross-sectional study), 30 nasal obstruction (non-cross-sectional study), 34 dyspnoea (non-cross-sectional study), 31 loss of sense of smell (non-cross-sectional study), 32 loss of taste (non-cross-sectional study), 33 positive auscultation findings (non-cross-sectional study)

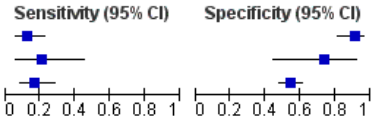
Cough (non-cross-sectional study)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Zhao 2020a	9	12	10	3	Covid-19 pneumonia	0.47 [0.24, 0.71]	0.20 [0.04, 0.48]
Chen X 2020	48	56	22	10	Covid-19 pneumonia	0.69 [0.56, 0.79]	0.15 [0.08, 0.26]
Yan 2020a	21	104	38	99	Covid-19 disease	0.36 [0.24, 0.49]	0.49 [0.42, 0.56]



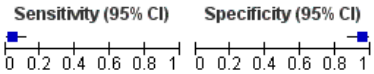
Sore throat (non-cross-sectional study)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Chen X 2020	9	6	61	60	Covid-19 pneumonia	0.13 [0.06, 0.23]	0.91 [0.81, 0.97]
Zhao 2020a	4	4	15	11	Covid-19 pneumonia	0.21 [0.06, 0.46]	0.73 [0.45, 0.92]
Yan 2020a	10	92	49	111	Covid-19 disease	0.17 [0.08, 0.29]	0.55 [0.48, 0.62]



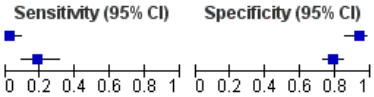
Rhinorrhoea (non-cross-sectional study)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Chen X 2020	3	3	67	63	Covid-19 pneumonia	0.04 [0.01, 0.12]	0.95 [0.87, 0.99]



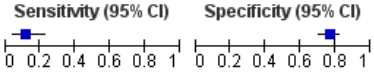
Nasal obstruction (non-cross-sectional study)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Chen X 2020	2	4	68	62	Covid-19 pneumonia	0.03 [0.00, 0.10]	0.94 [0.85, 0.98]
Yan 2020a	11	43	48	160	Covid-19 disease	0.19 [0.10, 0.31]	0.79 [0.73, 0.84]



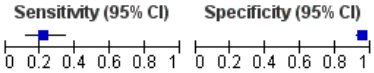
Dyspnoea (non-cross-sectional study)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Yan 2020a	7	47	52	156	Covid-19 disease	0.12 [0.05, 0.23]	0.77 [0.70, 0.82]



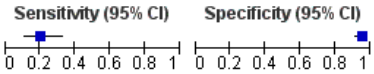
Loss of smell (anosmia) (non-cross-sectional study)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Yan 2020a	13	9	46	194	Covid-19 disease	0.22 [0.12, 0.35]	0.96 [0.92, 0.98]



Loss of taste (ageusia) (non-cross-sectional study)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Yan 2020a	12	10	47	193	Covid-19 disease	0.20 [0.11, 0.33]	0.95 [0.91, 0.98]



Positive auscultation findings (non-cross-sectional study)

Study	TP	FP	FN	TN	Target condition	Sensitivity (95% CI)	Specificity (95% CI)
Zhao 2020a	2	5	17	10	Covid-19 pneumonia	0.11 [0.01, 0.33]	0.67 [0.38, 0.88]

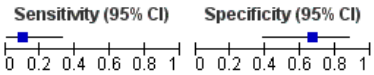


Figure 13. Forest plot of tests: 37 fatigue (non-cross-sectional study), 36 fever (non-cross-sectional study), 39 headache (non-cross-sectional study), 38 myalgia or arthralgia (non-cross-sectional study)

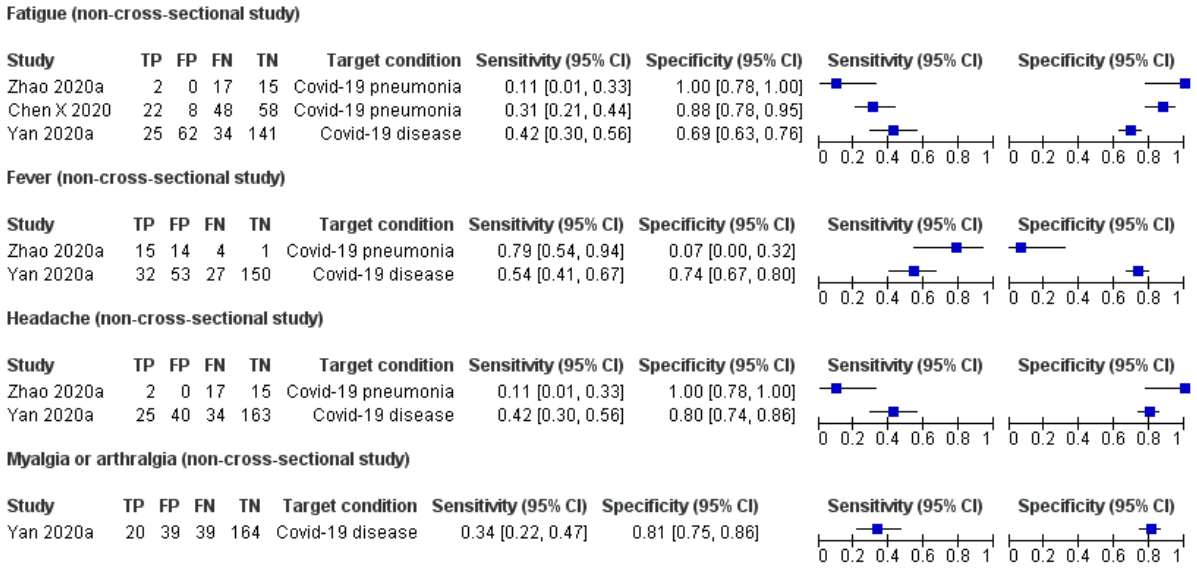


Figure 14. Forest plot of tests: 40 diarrhoea (non-cross-sectional study), 41 nausea/vomiting (non-cross-sectional study), 42 gastrointestinal symptoms, not specified (non-cross-sectional study)

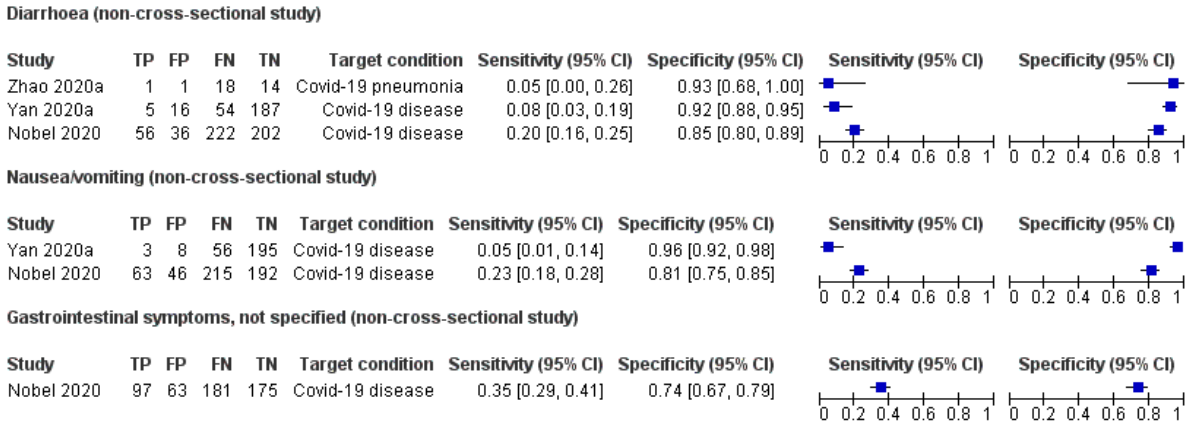
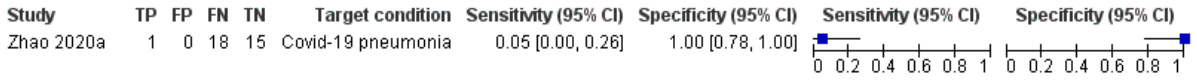


Figure 15. Forest plot of 35 chest tightness (non-cross-sectional study)



Overall, diagnostic accuracy of individual signs and symptoms is low, especially sensitivity. In addition, results were highly variable across studies, making it difficult to draw firm conclusions.

Signs and symptoms for which sensitivity was reported above 50% in at least one cross-sectional study are the following.

- Cough: sensitivity between 43% and 71%, specificity between 14% and 54%
- Sore throat: sensitivity between 5% and 71%, specificity between 55% and 80%
- Fever: sensitivity between 7% and 91%, specificity between 16% and 94%
- Myalgia or arthralgia: sensitivity between 19% and 86%, specificity between 45% and 91%
- Fatigue: sensitivity between 10% and 57%, specificity between 60% and 94%
- Headache: sensitivity between 3% and 71%, specificity between 78% and 98%

For fever, six of nine studies report a sensitivity of at least 80%, which is unsurprising considering fever was a key feature of COVID-19 that was used in selecting patients for further testing. As a result, most participants in these studies would have fever, both cases and non-cases. The same applies to cough, which was also listed as one of the main criteria for SARS-CoV-2 testing and may have contributed to inflated sensitivity estimates.

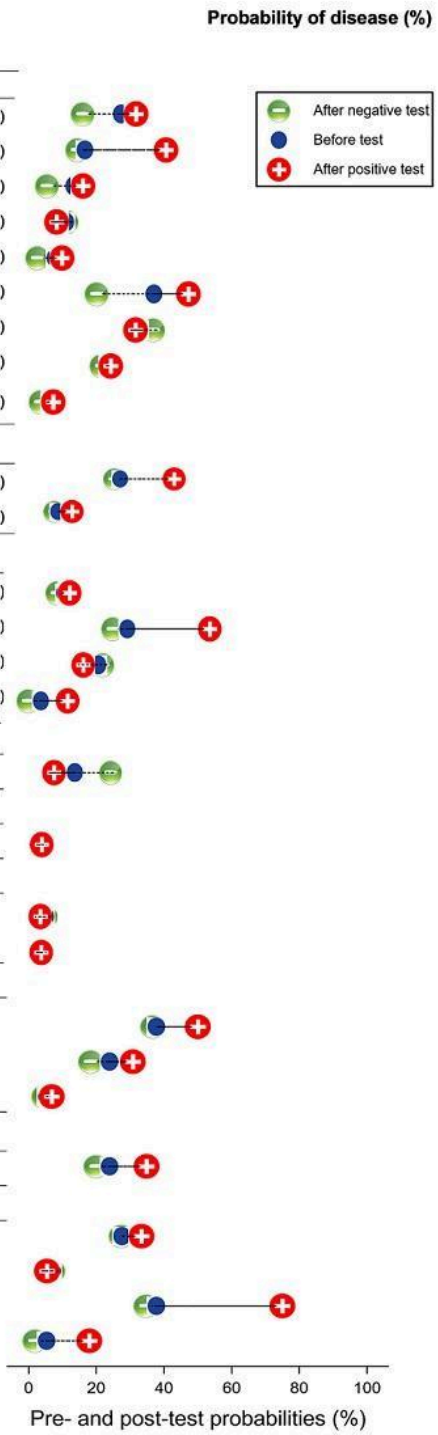
Specificity of at least 90% was achieved for 19 signs and symptoms. In only four signs and symptoms did this go along with sensitivity of at least 50% which would correspond to a positive likelihood ratio of at least 5, a commonly used arbitrary definition of a red flag. Using this definition, fever, myalgia or arthralgia, fatigue, or headache are to be considered red flags.

Strikingly, most of the respiratory symptoms such as cough, sore throat and sputum production are below the diagonal in ROC space (Figure 8). The diagonal line in ROC space is where sensitivity equals 1-specificity, meaning a test that is on the diagonal line has a positive likelihood ratio of 1 and is therefore not diagnostic because disease probability is left unchanged after conducting the test. Tests that lie below the diagonal line have a positive likelihood ratio that is smaller than 1, meaning the probability of COVID-19 disease decreases when this test is positive. For example, in Sun 2020a, pretest probability of COVID-19 is 6.9%; probability decreases to 6.4% when the patient has a cough and increases to 8.0% when the patient does not have a cough. We hypothesise on the reason for this counterintuitive finding in the discussion section. In contrast to respiratory features, systemic features are mostly above the diagonal line (Figure 9), suggesting that they do increase the probability of COVID-19 when present. Gastrointestinal symptoms and cardiovascular features are clustered in the bottom left corner or on the diagonal line suggesting that they have very little diagnostic value (Figure 10; Figure 11).

To further illustrate the systemic features' ability to either rule in or rule out COVID-19 disease or COVID-19 pneumonia, we constructed dumbbell plots showing pre- and post-test probabilities for each feature in each study (Figure 16). For each test, we have plotted the pre-test probability, which is the prevalence of COVID-19 disease (blue dot). Probability then changes depending on a positive test result (red dot marked +) or a negative test result (green dot marked -). The plot shows that fever, for example, increases the probability of COVID-19 in two studies (Ai 2020a; Rentsch 2020), makes little to no difference in five studies (Feng 2020; Liang 2020; Peng 2020; Song 2020; Zhu 2020), and decreases the probability of COVID-19 in two studies (Cheng 2020a; Tolia 2020).

Figure 16. Dumbbell plot: this plot shows how disease probability changes after a positive test result (red dot with plus sign) or after a negative test (green dot with minus sign). Pre-test probability or prevalence is the blue dot

Study	Target condition	Prevalence	Likelihood ratio (95%CI)	
			Positive	Negative
Fever				
Zhu 2020b	Covid-19 disease	27.6%	1.24 (1.01 to 1.53)	0.49 (0.21 to 1.15)
Rentsch 2020	Covid-19 disease	16.3%	3.65 (2.94 to 4.53)	0.83 (0.80 to 0.87)
Peng 2020a	Covid-19 disease	12.8%	1.26 (1.00 to 1.60)	0.32 (0.05 to 2.18)
Tolia 2020	Covid-19 disease	10.3%	0.70 (0.17 to 2.79)	1.03 (0.93 to 1.15)
Song 2020	Covid-19 disease	6.9%	1.35 (1.26 to 1.44)	0.21 (0.10 to 0.47)
Ai 2020a	Covid-19 pneumonia	37.7%	1.55 (1.04 to 2.31)	0.41 (0.16 to 1.06)
Cheng 2020a	Covid-19 pneumonia	33.3%	0.94 (0.61 to 1.44)	1.20 (0.35 to 4.13)
Liang Y 2020	Covid-19 pneumonia	23.9%	1.03 (0.84 to 1.26)	0.87 (0.27 to 2.83)
Feng 2020a	Covid-19 pneumonia	5.3%	1.23 (0.89 to 1.70)	0.47 (0.08 to 2.94)
Myalgia or fatigue				
Zhu 2020b	Covid-19 disease	27.6%	2.19 (0.72 to 6.67)	0.91 (0.77 to 1.07)
Song 2020	Covid-19 disease	6.9%	1.75 (1.26 to 2.44)	0.84 (0.73 to 0.97)
Myalgia or arthralgia				
Peng 2020a	Covid-19 disease	12.8%	1.16 (0.71 to 1.90)	0.80 (0.35 to 1.82)
Cheng 2020a	Covid-19 pneumonia	33.3%	3.00 (0.58 to 15.41)	0.80 (0.54 to 1.18)
Liang Y 2020	Covid-19 pneumonia	23.9%	0.75 (0.28 to 1.99)	1.08 (0.84 to 1.39)
Feng 2020a	Covid-19 pneumonia	5.3%	2.90 (1.93 to 4.34)	0.20 (0.03 to 1.25)
Low body temperature				
Rentsch 2020	Covid-19 disease	16.3%	0.54 (0.48 to 0.61)	1.99 (1.83 to 2.17)
Shivers				
Feng 2020a	Covid-19 pneumonia	5.3%	1.05 (0.16 to 6.80)	0.99 (0.73 to 1.35)
Chills				
Song 2020	Covid-19 disease	6.9%	0.72 (0.33 to 1.60)	1.03 (0.97 to 1.09)
Feng 2020a	Covid-19 pneumonia	5.3%	1.02 (0.31 to 3.40)	0.99 (0.61 to 1.61)
Fatigue				
Ai 2020a	Covid-19 pneumonia	37.7%	1.65 (0.25 to 10.81)	0.96 (0.81 to 1.14)
Liang Y 2020	Covid-19 pneumonia	23.9%	1.42 (0.89 to 2.27)	0.72 (0.42 to 1.22)
Feng 2020a	Covid-19 pneumonia	5.3%	1.31 (0.54 to 3.19)	0.85 (0.44 to 1.63)
Headache or dizziness				
Liang Y 2020	Covid-19 pneumonia	23.9%	1.70 (0.84 to 3.44)	0.80 (0.56 to 1.14)
Headache				
Zhu 2020b	Covid-19 disease	27.6%	1.31 (0.12 to 13.98)	0.99 (0.92 to 1.07)
Song 2020	Covid-19 disease	6.9%	0.76 (0.40 to 1.44)	1.04 (0.96 to 1.11)
Ai 2020a	Covid-19 pneumonia	37.7%	4.95 (0.55 to 44.41)	0.88 (0.72 to 1.06)
Feng 2020a	Covid-19 pneumonia	5.3%	3.88 (2.14 to 7.05)	0.35 (0.11 to 1.13)



DISCUSSION

Summary of main results

Individual signs and symptoms appear to have very poor diagnostic properties for COVID-19, although this has to be interpreted in the presence of a limited number of studies, heterogeneity between the studies precluding any firm conclusions and in a context of selection bias. Most features had very low sensitivity, while specificity was moderate to high.

We have identified four possible red flags, that is, symptoms that increased the probability of COVID-19 when present because of a positive likelihood ratio of at least 5 in at least one study: fever, myalgia or arthralgia, fatigue, and headache. When we apply the results of sensitivity and specificity of these systemic features to disease probabilities, we assess their value to rule in and rule out disease as shown in the dumbbell plots (Figure 16). These clearly show the limited effect on disease probability from these signs and symptoms. Importantly, we did not find any studies investigating the diagnostic accuracy of combinations of signs and symptoms. There were also no studies from community primary care settings.

Some of our findings are counterintuitive, for example that the majority of the studies that investigated cough found that cough decreases the probability of COVID-19 despite the fact that it is part of the case definition of COVID-19 in most countries. This is also the case for fever in two studies and myalgia in one study - even though these features were also red flags in at least one other study. We believe this may be caused by selection bias. Selection bias is present when selective and non-random inclusion and exclusion of participants apply and the resulting association between exposure and outcome (here the accuracy of the test) differs in the selected study population compared to the eligible study population, and it has been shown that this may decrease estimates of diagnostic accuracy (Rutjes 2006). For the diagnosis of COVID-19, rapidly and constantly changing, and widely variable test criteria have influenced who was referred for testing and who was not. Inclusion in the study of only a selective fraction of eligible patients can give a biased estimate of the real accuracy of the index test when measured against the reference standard and real disease status. Griffith 2020 reported on the problematic presence of collider stratification bias in the published studies on COVID-19. Appropriate sampling strategies need to be applied to avoid conclusions of spurious relationships, more specifically in our case, the biased accuracy estimates of signs and symptoms for the diagnosis of COVID-19 disease. Selection of patients based on the presence of specific pre-set symptoms, such as fever and cough, lead to biased associations between these symptoms and disease, and sensitivity and specificity estimates that differ from their true values. The example of collider bias for cough is illustrated in Figure 17. Grouping studies by diagnostic criteria for selection might clarify this issue, but studies do not clearly describe them, with study authors referring to the guidelines in general that were applicable at the time.

Figure 17. Directed acyclic graph on cough

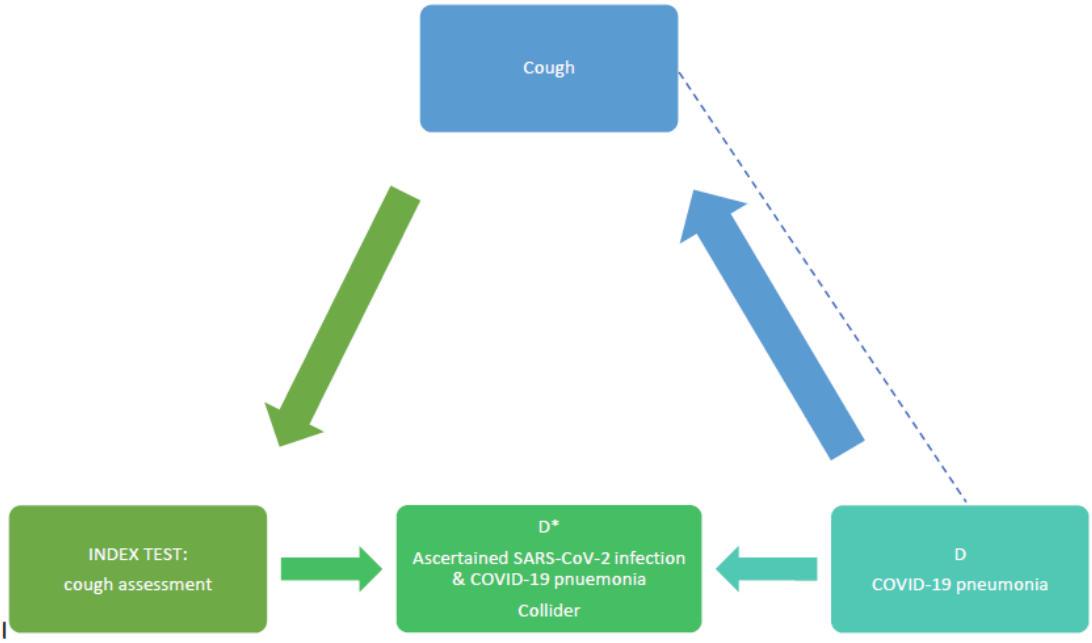


Figure Directed Acyclic Graph (DAG): the symptom, 'cough' is used to enter the study for cough assessment. Both cough and COVID-19 pneumonia (D) result in ascertained diagnosis of SARS-CoV-2 infection (D*). D* is a collider on the pathway between cough and COVID-19 pneumonia leading to a biased association between the symptom cough and COVID-19 pneumonia.

Another form of selection bias is spectrum bias, where the patients included in the studies do not reflect the patient spectrum to which the index test will be applied. The inclusion of hospitalised patients can lead to such a bias, when in these patients both the distribution of signs and symptoms differ and assessment with the reference standard is differential. In addition, the distribution and severity of alternative diagnoses may be different in hospitalised populations than in patients presenting to ambulatory care settings.

Strengths and weaknesses of the review

Strengths of our review are the systematic and broad search performed to include all possible studies, including those prior to peer-review, to gather the largest number of studies available at this point. Exclusion of cases-only studies, the largest number of the published cohorts of patients with COVID-19, limits the available data but also improves the quality of the evidence and the possibility to present both sensitivity and specificity (cases only cannot provide both accuracy measures). Because this is a living systematic review, future updates may offer the possibility to do a meta-analysis, which was not possible at this stage. In addition, further insights into this novel disease may lead to new evidence on signs and symptoms that are more diagnostic, which we will incorporate in future updates.

The lack of data on combinations of signs and symptoms is an important evidence gap. Consequently, there is no evidence on syndromic presentation and the value of composite signs and symptoms on the diagnostic accuracy measures. In addition, subgroup analyses by the pre-defined variables were not feasible due to lack of reporting.

We need to assess multiple variables for their possible confounding effect on the summary estimates. Possible confounders include the presence of other respiratory pathogens (seasonality), the phase of the epidemic, exposure to high versus low prevalence setting, high or low exposure risk, comorbidity of the participants, or time since infection. Seasonality may influence specificity, because alternative diagnoses such as influenza or other respiratory viruses are more prevalent in winter, leading to more non-COVID-19 patients displaying symptoms such as cough or fever, decreasing specificity. In this version of the review, all studies were conducted in Winter or early Spring, suggesting this may still have been at play. However, social distancing policies have shortened this year's influenza season in several countries (www.who.int/influenza/surveillance_monitoring/updates/en/), which may have led to higher specificity for signs and symptoms than what we may expect in the next influenza season. In future updates of the review, we will explore seasonality effects if data allow. As for time since onset, given that the moment of infection is more likely than not an unrecognisable and unmeasurable variable, time since onset of symptoms can be used as a proxy. Reporting of studies, with presentation of the 2x2 table stratified by time since onset of disease, is informative and might have the potential to increase accuracy of the signs and symptoms and their diagnostic differential potential.

Applicability of findings to the review question

The high risk of selection bias, with many studies including patients who had already been admitted to hospital or who presented to hospital settings with the intent to hospitalise, leads to findings that are less applicable to people presenting in primary care, who on average experience a shorter illness duration, less severe symptoms and have a lower probability of the target condition.

Our search did not find any articles providing data on children. Children have been underrepresented in the studies on diagnosing SARS-CoV-2 infection. Their absence seems related to the general mild presentation of the disease in the paediatric population, and the even more frequent asymptomatic course of COVID-19 in children. The full scope of disease presentation in children is therefore not known. It is important to identify signs and symptoms that can be used to clinically assess children with suspected COVID-19, especially because aspecific presentations and fever without a source are already common in this age group, and acute infection in children is a common cause for families to self-isolate or get tested. Misclassification of children, where children will be asked to remain in quarantine when they present with predefined, but not yet evidence-based symptoms needs to be avoided to decrease the possible damage done to children's health and education. Having separate data for neonates, young infants, toddlers, school-aged children and adolescents is therefore of value.

Another important patient group is older adults. They are most at risk of an adverse outcome of SARS-CoV-2 infection, including an increased risk of requiring intensive care support and and increased risk of mortality. In this version of the review, only one study focused on adults aged 55 to 75 years. All other studies included adults of all ages and did not present results separately for older age groups. The lack of a solid evidence base for the diagnosis of COVID-19 in older adults adds to the difficulty in diagnosing serious infections in this age group, as other serious infections such as bacterial pneumonia or urinary sepsis also tend to lead to aspecific presentations.

The association of a single sign or symptom with COVID-19 is highly uncertain, and we do not have data on combinations of signs and symptoms. Additionally, potentially more diagnostic symptoms such as loss of sense of smell have not yet been studied widely and remain to be assessed in well designed studies. Moreover, the nature of the signs and symptoms that are used to guide self-isolation decisions are such that people may end up being quarantined on a regular basis, leading to missed days at school or work, isolation and anxiety.

In future updates of this review, we intend to organise findings by age group, settings (in particular primary care settings versus hospital settings), and target condition, when evidence allows.

AUTHORS' CONCLUSIONS

Implications for practice

The results were highly variable across studies, making it difficult to draw firm conclusions. Selection bias further hinders interpretability. Until results of further studies become available, broad investigation of patients with suspected SARS-Cov-2 infection remains necessary. Neither absence nor presence of the individual signs and symptoms included in this review are accurate enough to rule in or rule out disease.

Implications for research

Our review reflects the need for improved study methodology in COVID-19 diagnostic accuracy research: appropriate patient

sampling strategies; prospective one-gate design; and investigating the presence or absence of clinical signs and symptoms in all suspected patients. In addition, we urgently need studies in community primary care settings, and studies investigating combinations of signs and symptoms. Evidence on signs and symptoms that are used for testing or referral decisions, such as loss of sense of smell, heart rate, breathing rate and oxygen saturation, should be included in future studies using clearly stated definitions and cut-offs. In order to inform self-isolation policies, studies in community settings, where prevalence is lower than in the included studies, will be needed to better determine the balance of risks arising from false positives and false negatives.

We also need improved reporting with studies clearly describing how they assessed signs and symptoms, when and by whom, and providing clearer definitions of what constitutes an abnormal sign or symptom. Studies also need to report reference standards more clearly.

In addition, more data on specific patient groups with comorbidities at higher risk of complications or severe disease are needed, especially older adults, as missing COVID-19 disease may have more serious consequences in these patients. We also need to have more studies in children.

We would like to recommend authors to adhere to the STARD guidelines when reporting new studies on this topic (Bossuyt 2015).

ACKNOWLEDGEMENTS

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

- 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 (Salameh JP, McGrath TA, van der Pol CB, Frank RA, Prager R, Hare SS, Dennie C, Jenniskens K, Korevaar DA, Cohen JF, van de Wijgert J, Damen JAAG, 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.

We would like to acknowledge Joanne Merckx for her contribution to the data extraction of nine papers in the initial stages of this review. Due to conflict of interest, a decision was taken to have a systematic reviewer, Nicholas Henschke, independently check all data extracted by her before publication.

The editorial process for this review was managed by Cochrane's EMD Editorial Service in collaboration with Cochrane Infectious Diseases. We thank Helen Wakeford, Anne-Marie Stephani and Deirdre Walshe for their comments and editorial management. We thank Robin Featherstone for comments on the search and Mike Brown and Paul Garner for sign-off comments. We thank Denise Mitchell for her efforts in copy-editing this review.

Thank you also to peer referees Trish Greenhalgh, Robert Walton, Chris Cates and Lynda Ware, consumer referee Jenny Negus, and methodological referees Gianni Virgili and Marta Roqué, for their insights.

The editorial base of Cochrane Infectious Diseases 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 UK National Institute for Health Research (NIHR) Senior Investigator Emeritus. Yemisi Takwoingi is supported by a NIHR Postdoctoral Fellowship. Jonathan Deeks, Jacqueline Dinnes, Yemisi Takwoingi, Clare Davenport and Malcolm Price are supported by the NIHR Birmingham Biomedical Research Centre. This paper presents independent research supported by the NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham. The views expressed are those of the author(s) 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 2020a *{published data only}*

Ai JW, Zhang HC, Xu T, Wu J, Zhu M, Yu YQ, et al. Optimizing diagnostic strategy for novel coronavirus pneumonia, a multi-center study in Eastern China. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.02.13.20022673]

Cheng 2020a *{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-center study in Shanghai, China. American Journal of Roentgenology 2020 Mar 14 [Epub ahead of print]. [DOI: 10.2214/AJR.20.22959]

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* 16/04/2020. [DOI: 10.1007/s00330-020-06829-2]

Feng 2020a *{published data only}*

Feng C, Huang Z, Wang L, Chen X, Zhai Y, Zhu F, et al. A novel triage tool of artificial intelligence assisted diagnosis aid system for suspected COVID-19 pneumonia in fever clinics. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.03.19.20039099]

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 [Preprint]* 2020. [DOI: 10.1101/2020.02.25.20027763]

Nobel 2020 *{published data only}*

Nobel YR, Phipps M, Zucker J, Lebwohl B, Wang TC, Sobieszczyk ME, et al. Gastrointestinal symptoms and COVID-19: case-control study from the United States. Gastroenterology 2020 Apr 12 [Epub ahead of print]. [DOI: 10.1053/j.gastro.2020.04.017]

Peng 2020a *{published data only}*

Peng L, Liu KY, Xue F, Miao YF, Tu PA, Zhou C. Improved early recognition of coronavirus disease-2019 (COVID-19): single-center data from a Shanghai screening hospital. *Archives of Iranian Medicine* 2020;**23**(4):272-6.

Rentsch 2020 *{published data only}*

Rentsch CT, Kidwai-Khan F, Tate JP, Park LS, King JT, Skanderson M, et al. Covid-19 testing, hospital admission, and intensive care among 2,026,227 United States veterans aged 54-75 years. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.04.09.20059964]

Song 2020b *{published data only}*

Song CY, Xu J, He JQ, Lu YQ. COVID-19 early warning score: a multi-parameter screening tool to identify highly suspected patients. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.03.05.20031906]

Sun 2020a *{published data only}*

Sun Y, Koh V, Marimuthu K, Ng OT, Young B, Vasoo S, et al. Epidemiological and clinical predictors of COVID-19. Clinical Infectious Diseases 2020 Mar 25 [Epub ahead of print]. [DOI: 10.1093/cid/ciaa322]

Tolia 2020 *{published data only}*

Tolia VM, Chan TC, Castillo EM. Preliminary results of initial testing for coronavirus (COVID-19) in the emergency department. *Western Journal of Emergency Medicine* 2020;**21**(3):503-6.

Wee 2020 *{published data only}*

Wee LE, Chan YZ, Teo NY, Cherng BZ, Thien SY, Wong M, et al. The role of self-reported olfactory and gustatory dysfunction as a screening criterion for suspected COVID-19. European Archives of Oto-Rhino-Laryngology 2020 Apr 24 [Epub ahead of print]. [DOI: 10.1007/s00405-020-05999-5]

Yan 2020a *{published data only}*

Yan CH, Faraji F, Prajapati DP, Boone CE, DeConde AS. Association of chemosensory dysfunction and Covid-19 in patients presenting with influenza-like symptoms. International Forum of Allergy & Rhinology 2020 Apr 12 [Epub ahead of print]. [DOI: 10.1002/alr.22579]

Yang 2020d *{published data only}*

Yang Z, Lin D, Chen X, Qiu J, Li S, Huang R, et al. Distinguishing COVID-19 from influenza pneumonia in the early stage through CT imaging and clinical features. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.04.17.20061242]

Zhao 2020a *{published data only}*

Zhao D, Yao F, Wang L, Zheng L, Gao Y, Ye J, et al. A comparative study on the clinical features of COVID-19 pneumonia to other pneumonias. Clinical Infectious Diseases 2020 Mar 12 [Epub ahead of print]. [DOI: 10.1093/cid/ciaa247]

Zhu 2020b *{published data only}*

Zhu W, Xie K, Lu H, Xu L, Zhou S, Fang S. Initial clinical features of suspected coronavirus disease 2019 in two emergency departments outside of Hubei, China. Journal of Medical Virology 2020 Mar 24 [Epub ahead of print]. [DOI: 10.1002/jmv.25763]

References to studies excluded from this review

Guan 2020 *{published data only}*

Guan, W, Ni, Z, Hu, Y, Liang, W, Ou, C, He, J, et al. Clinical characteristics of 2019 novel coronavirus infection in China. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.02.06.20020974]

Soares 2020 *{published data only}*

Soares F, Villavicencio A, Anzanello MJ, Fogliatto FS, Idiart M, Stevenson M. A novel high specificity COVID-19 screening

method based on simple blood exams and artificial intelligence. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.04.10.20061036]

Song 2020 *{published data only}*

Song F, Shi N, Shan F, Zhang Z, Shen J, Lu H, et al. Emerging coronavirus 2019-nCoV pneumonia. *Radiology* 2020;**295**(1):200274. [DOI: 10.1148/radiol.2020200274]

Wang 2020 *{published data only}*

Wang Y, Kang H, Liu X, Tong Z. Combination of RT-qPCR testing and clinical features for diagnosis of COVID-19 facilitates management of SARS-CoV-2 outbreak. *Journal of Medical Virology* 2020;**92**(6). [DOI: 10.1002/jmv.25721 10.1002/jmv.25721]

References to ongoing studies

ChiCTR2000029462 *{published data only}*

ChiCTR2000029462. Study for clinical characteristics and distribution of TCM syndrome of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=48922 (first received 27 April 2020).

ChiCTR2000029734 *{published data only}*

ChiCTR2000029734. Epidemiological investigation and clinical characteristics analysis of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=48868 (first received 27 April 2020).

ChiCTR2000029770 *{published data only}*

ChiCTR2000029770. Study for epidemiology, diagnosis and treatment of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/hvshowproject.aspx?id=23744 (first received 27 April 2020).

ChiCTR2000029839 *{published data only}*

ChiCTR2000029839. An observational study on the clinical characteristics, treatment and outcome of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=49439 (first received 27 April 2020).

ChiCTR2000029865 *{published data only}*

ChiCTR2000029865. Descriptive study on the clinical characteristics and outcomes of novel coronavirus pneumonia (COVID-19) in cardiovascular patients. www.chictr.org.cn/showproj.aspx?proj=49545 (first received 27 April 2020).

ChiCTR2000029866 *{published data only}*

ChiCTR2000029866. Early warning prediction of patients with severe novel coronavirus pneumonia (COVID-19) based on multiomics. www.chictr.org.cn/showproj.aspx?proj=49519 (first received 27 April 2020).

ChiCTR2000029959 *{published data only}*

ChiCTR2000029959. Clinical observation and research of Severe acute respiratory syndrome coronavirus 2(COVID-19) infection in perinatal newborns. www.chictr.org.cn/showproj.aspx?proj=49636 (first received 27 April 2020).

ChiCTR 2000030096 *{published data only}*

ChiCTR 2000030096. Study for establishment of correlation between virological dynamics and clinical features in novel coronavirus pneumonia (COVID-19). http://www.chictr.org.cn/showprojen.aspx?proj=49794 (first received 27 April 2020).

ChiCTR2000030256 *{published data only}*

ChiCTR2000030256. Epidemiological and clinical characteristics of COVID-19: a large-scale investigation in epicenter Wuhan, China. www.chictr.org.cn/showproj.aspx?proj=50078 (first received 27 April 2020).

ChiCTR2000030327 *{published data only}*

ChiCTR2000030327. Analysis of clinical characteristics of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=50214 (first received 27 April 2020).

ChiCTR2000030363 *{published data only}*

ChiCTR2000030363. Novel coronavirus infected disease (COVID-19) in children: epidemiology, clinical features and treatment outcome. www.chictr.org.cn/showproj.aspx?proj=49984 (first received 27 April 2020).

ChiCTR2000030387 *{published data only}*

ChiCTR2000030387. Clinical observation and research of multiple organs injury in severe patients with novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=50329 (first received 27 April 2020).

ChiCTR2000030464 *{published data only}*

ChiCTR2000030464. Study for the clinical characteristics of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=50382 (first received 27 April 2020).

ChiCTR2000030491 *{published data only}*

ChiCTR2000030491. A medical records based study for comparing differences of clinical features and outcomes of novel coronavirus pneumonia (COVID-19) patients between Sichuan Province and Wuhan City. www.chictr.org.cn/hvshowproject.aspx?id=23102 (first received 27 April 2020).

ChiCTR2000030519 *{published data only}*

ChiCTR2000030519. Study for the clinical characteristics and digestive system damage of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=50604 (first received 27 April 2020).

ChiCTR2000030544 *{published data only}*

ChiCTR2000030544. Study for the risk factors of critically ill patients with novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=50134 (first received 27 April 2020).

ChiCTR 2000030679 *{published data only}*

ChiCTR 2000030679. Cohort study of novel coronavirus infected diseases (COVID-19) in children. www.chictr.org.cn/hvshowproject.aspx?id=23417 (first received 27 April 2020).

ChiCTR2000030707 *{published data only}*

ChiCTR2000030707. Retrospective study on novel coronavirus pneumonia (COVID-19) in Tibetan Plateau. www.chictr.org.cn/showproj.aspx?proj=50160 (first received 27 April 2020).

ChiCTR2000030722 *{published data only}*

ChiCTR2000030722. Auscultatory characteristics of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=50338 (first received 27 April 2020).

ChiCTR2000030739 *{published data only}*

ChiCTR2000030739. Exploration of the clinical characteristics of patients with novel coronavirus pneumonia (COVID-19) and its differences from patients with severe influenza A and MERS. www.chictr.org.cn/showproj.aspx?proj=50896 (first received 27 April 2020).

ChiCTR2000030755 *{published data only}*

ChiCTR2000030755. A medical records based study for characteristics, prognosis of elderly patients with novel coronavirus pneumonia (COVID-19) in Wuhan area. www.chictr.org.cn/hvshowproject.aspx?id=23554 (first received 27 April 2020).

ChiCTR2000030778 *{published data only}*

ChiCTR2000030778. A medical records based study for epidemic and clinical features of novel coronavirus pneumonia (COVID-19) in Ningbo First Hospital. www.chictr.org.cn/hvshowproject.aspx?id=23642 (first received 27 April 2020).

ChiCTR2000030784 *{published data only}*

ChiCTR2000030784. A study for clinical characteristics of novel coronavirus pneumonia (COVID-19) patients follow-up in Guangxi. www.chictr.org.cn/showproj.aspx?proj=50307 (first received 27 April 2020).

ChiCTR2000030796 *{published data only}*

ChiCTR2000030796. Clinical characteristics and treatment of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=50991 (first received 27 April 2020).

ChiCTR 2000030798 *{published data only}*

ChiCTR 2000030798. A medical records based study for clinical characteristics of novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/hvshowproject.aspx?id=23687 (first received 27 April 2020).

ChiCTR2000030803 *{published data only}*

ChiCTR2000030803. Collection and analysis of clinical data in severe and critically ill patients with novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=51007 (first received 27 April 2020).

ChiCTR2000030807 *{published data only}*

ChiCTR2000030807. Clinical characteristics and prognosis of cancer patients with novel coronavirus pneumonia (COVID-19). www.chictr.org.cn/showproj.aspx?proj=51019 (first received 27 April 2020).

ChiCTR2000030818 *{unpublished data only}*

ChiCTR2000030818. A medical records based study for the value of Lymphocyte subsets in the diagnose and treatment. http://www.chictr.org.cn/hvshowproject.aspx?id=23742 (first received 27 April 2020).

ChiCTR2000030819 *{published data only (unpublished sought but not used)}*

ChiCTR2000030819. Retrospective analysis of digestive system symptoms in 600 cases of novel coronavirus pneumonia (COVID-19) in Guanggu district, Wuhan. www.chictr.org.cn/showproj.aspx?proj=51039 2020.

ChiCTR2000030834 *{published data only}*

ChiCTR2000030834. Epidemiological characteristics and antibody levels of novel coronavirus pneumonia (COVID-19) of pediatric medical staff working in quarantine area. www.chictr.org.cn/showproj.aspx?proj=51047 (frist received 27 April 2020).

ChiCTR2000030854 *{published data only}*

ChiCTR2000030854. A clinical multicenter study for the occurrence, development and prognosis of novel coronavirus pneumonia (COVID-19). http://www.chictr.org.cn/showprojen.aspx?proj=51083 (first received 27 April 2020).

ChiCTR2000030858 *{published data only}*

ChiCTR2000030858. Clinical characteristics and outcomes of 483 mild patients with novel coronavirus pneumonia (COVID-19) in Wuhan, China during the outbreak: a single-center, retrospective study from the mobile cabin hospital. www.chictr.org.cn/showproj.aspx?proj=51097 (first received 27 April 2020).

ChiCTR2000030863 *{published data only}*

ChiCTR2000030863. Clinical and CT imaging characteristics of novel coronavirus pneumonia (COVID-19): an multicenter cohort study. www.chictr.org.cn/showproj.aspx?proj=50767 (first received 27 April 2020).

NCT04270383 *{published data only}*

NCT04270383. Clinical characteristics and long-term prognosis of 2019-nCoV infection in children. clinicaltrials.gov/ct2/show/NCT04270383 (first received 17 February 2020).

NCT04279782 *{published data only}*

NCT04279782. Clinical features of suspected and confirmed patients of 2019 novel coronavirus infection. www.clinicaltrials.gov/ct2/show/NCT04279782 (first received 27 April 2020).

NCT04279899 *{published data only}*

NCT04279899. The investigation of the neonates with or with risk of COVID-19. clinicaltrials.gov/ct2/show/NCT04279899 (first received 27 April 2020).

NCT04285801 *{published data only}*

NCT04285801. Critically ill patients with COVID-19 in Hong Kong: a multicentre observational cohort study. clinicaltrials.gov/ct2/show/NCT04285801 (first received 27 April 2020).

NCT04292327 *{published data only}*

NCT04292327. Clinical progressive characteristics and treatment effects of 2019-novel coronavirus. clinicaltrials.gov/ct2/show/NCT04292327 (first received 27 April 2020).

NCT04292964 *{published data only}*

NCT04292964. Prognostic factors of patients with COVID-19. clinicaltrials.gov/ct2/show/NCT04292964 (first received 27 April 2020).

NCT04315870 *{published data only}*

NCT04315870. Clinical characteristics of coronavirus disease 2019 (COVID-19) in pregnancy: the Italian Registry on coronavirus in pregnancy. clinicaltrials.gov/ct2/show/NCT04315870 (first received 20 March 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. *BMJ* 2015;**351**:h5527. [DOI: 10.1136/bmj.h5527]

Griffith 2020

Griffith G, Morris TT, Tudball M, Herbert A, Mancano G, Pike L, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *medRxiv* 2020. [DOI: 10.1101/2020.05.04.20090506]

Macaskill 2013

Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2013. Available from srdta.cochrane.org.

McInnes 2020

McInnes M, Leeflang MM, Salameh J-P, McGrath T, Van der Pol CB, Frank RA, et al. Imaging tests for the diagnosis of COVID-19. *Cochrane Database of Systematic Reviews* 2020, Issue 6. [DOI: 10.1002/14651858.CD013639]

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, the PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Medicine* 2009;**6**(7):e1000097. [DOI: 10.1371/journal.pmed1000097]

Review Manager 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Rutjes 2006

Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, Van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. *Canadian Medical Association Journal* 2006;**174**(4):469476.

SAS 2015 [Computer program]

SAS Institute SAS. Version 9.4. Cary, NC, USA: SAS Institute, 2015.Available at www.sas.com.

STATA [Computer program]

STATA. Version 16. College Station, TX, USA: StataCorp, 2019.Available at www.stata.com.

Takwoingi 2017

Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Statistical Methods in Medical Research* 2017;**26**(4):1896-911.

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

References to other published versions of this review

Deeks 2020

Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MM, Spijker R, et al. Diagnosis of SARS-CoV-2 infection and COVID-19: accuracy of signs and symptoms; molecular, antigen, and antibody tests; and routine laboratory markers. *Cochrane Database of Systematic Reviews* 2020, Issue 4. [DOI: 10.1002/14651858.CD013596]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

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

Ai 2020a

Study characteristics	
Patient Sampling	**Purpose:** diagnosis of SARS-CoV-2 pneumonia
	Design: cross-sectional multicentre prospective study
	Recruitment: hospitalised pneumonia patients

Ai 2020a *(Continued)*

	Sample size: n = 53 (20 cases) **Inclusion criteria**: suspected SARS-COV-2 pneumonia patients, defined as having pneumonia after chest CT (with 1 of the 2 following criteria met: fever or respiratory symptoms, normal or decreased WBC counts/decreased **Exclusion criteria**: not defined
Patient characteristics and setting	**Facility cases:** confirmed case: a positive SARS-COV-2 nucleotides result either by metagenomic sequencing or RT-PCR assay for nasopharyngeal swab specimens **Facility controls**: pneumonia patients confirmed not to be infected by SARS-Cov2 (2 PCR tests, 2 days in between) **Country:** China **Dates**: 22 January 2020-19 February 2020 **Symptoms and severity**: suspected SARS-COV-2 pneumonia (NCP): having pneumonia after chest CT with 1 of the 2 following criteria met: fever or respiratory symptoms, normal or decreased WBC counts/decreased lymphocyte counts, and a travel history or contact with patients with fever or respiratory symptoms from Hubei Province or confirmed cases within 2 weeks **Demographics**: median age cases 37 years, controls 39 years, gender distribution cases (M/F: 50/50), controls (M/F: 48.5/51.5) **Exposure history**: not specified
Index tests	• Fever • Dry cough • Diarrhoea • Fatigue • Headache • Vomiting • Abdominal pain
Target condition and reference standard(s)	• TC: COVID-19 pneumonia • RS: a positive SARS-COV-2 nucleotides result either by metagenomic sequencing or RT-PCR assay for nasopharyngeal swab specimens, repeated after 2 days if negative on day 0
Flow and timing	Time interval not specified. Reference standard at day 0 and day 2, index tests from electronic medical records but stated at pneumonia onset
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 2020a *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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 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?		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?	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	

Chen X 2020

Study characteristics			

Chen X 2020 *(Continued)*

Patient Sampling	**Purpose:** diagnosis COVID-19 pneumonia - to identify differences in CT imaging and clinical manifestations between pneumonia patients with and without COVID-19, and to develop and validate a diagnostic model for COVID-19 based on radiological semantic and clinical features **Design**: cross-sectional multicentre retrospective study **Recruitment:** cases: consecutive patients with COVID-19 admitted in 5 independent hospitals controls: at the same period, another 66 consecutive pneumonia patients without COVID-19 from Meizhou People's Hospital **Sample size:** n = 136 (cases = 70) **Inclusion criteria**: patients admitted with COVID-19 pneumonia (cases) and patients admitted with non-COVID-19 pneumonia (controls) **Exclusion criteria**: not specified for cases except those from 1 hospital (Meizhou), for cases and controls in Meizhou: after chest CT neoplasm, tuberculosis, pulmonary oedema, pulmonary contusion, aspiration pneumonia, bronchitis, any local or systemic treatment before CT scan, normal CT image without epidemiological history
Patient characteristics and setting	**Facility cases:** pneumonia patients with positive SARS-CoV-2 test **Facility controls**: CT pneumonia patients with consecutive negative RT-PCR **Country:** China **Dates**: 1 January 2020-8 February 2020 **Symptoms and severity**: pneumonia patients for cases and control; unclear severity of cases **Demographics**: M/F: cases 41/29, controls 43/23 mean age: cases 42.9 range, 16-69 years, controls 46.7 range, 0.3-93 years **Exposure history**: data about exposure to epidemic centres collected, but no results in the study nor in appendices
Index tests	• Systolic BP • Diastolic BP • Respiration rate • Heart rate • Temperature • Dry cough • Fatigue • Sore throat • Stuffy • Runny nose
Target condition and reference standard(s)	• TC: COVID-19 pneumonia • RS: RT-PCR and next generation sequencing for SARS-Cov2
Flow and timing	Time interval not specified
Comparative	

Chen X 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		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
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?	Unclear		

Chen X 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

Cheng 2020a

Study characteristics	
Patient Sampling	**Purpose:** to identify the clinical features and CT manifestations of COVID-19 and compare them with those of pneumonia occurring in patients who do not have COVID-19 **Design**: cross-sectional single-centre retrospective study **Recruitment:** pneumonia patients who presented at a fever observation department in Shanghai **Sample size:** n = 33 (11 cases) **Inclusion criteria**: patients with clinical and radiological features of pneumonia, and a normal or reduced total leukocyte count or total lymphocyte count, plus an epidemiologic history that included travel or a history of residence in Hubei province or other areas where continuous transmission of local cases occurred within 14 days before onset of symptoms, a history of contact with patients who had fever or respiratory symptoms and were from Hubei province or other areas with continuous transmission of local cases within 14 days before onset of the disease, or clustering or epidemiologic association with the new coronavirus infection **Exclusion criteria**: not defined
Patient characteristics and setting	**Facility cases:** confirmed case: positive RT-PCR test result obtained by a throat swab. Test was repeated when the first test was negative **Facility controls**: pneumonia patients confirmed not to be infected by SARS-Cov2 (2 PCR tests) **Country:** China **Dates**: 19 January 2020-6 February 2020 **Symptoms and severity**: pneumonia was defined as patients with at least 1 clinical symptom (i.e. cough, sputum, fever, dyspnoea, or pleuritic chest pain), a finding of either coarse crackles on auscultation or elevated inflammatory biomarkers, and observation of a new pulmonary opacification on chest CT **Demographics**: median age +- SD cases 50.36 +- 15.5, controls 43.59 +- 16.02, gender distribution cases (M/F: 8/3), controls (M/F: 7/15) **Exposure history**: cases 8/11, controls 7/22 (in the last 14 days with patients with fever or respiratory symptoms or with known cases)
Index tests	• Fever • Cough • Sputum • Shortness of breath • Muscle ache

Cheng 2020a *(Continued)*

	• Diarrhoea • Sore throat • Peak body temperature
Target condition and reference standard(s)	• TC: COVID-19 pneumonia • RS: RT-PCR testing on throat swab specimens Tests were repeated if the first test was negative
Flow and timing	Time interval not specified, reference test at day 0 (or later when the first test was negative), index tests were questionnaired at day 0 for the presence of symptoms in the past period of time
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		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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 3: Reference Standard			

Cheng 2020a *(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?		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?	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	

Feng 2020a

Study characteristics	
Patient Sampling	**Purpose:** diagnosis COVID-19 pneumonia **Design**: cross-sectional, retrospective, single-centre study **Recruitment:** patients admitted to ED with history of exposure to COVID-19 **Sample size:** n = 132 (cases = 7) **inclusion criteria**: all patients admitted to the fever clinic of the ED of the First medical center, Chinese People's Liberation Army General Hospital (PLAGH) in Beijing with the epidemiological history of exposure to COVID-19 according to WHO interim guidance **Exclusion criteria**: < 14 years old, no other criteria specified
Patient characteristics and setting	**Facility cases:** among clinically suspected patients: those with a positive RT-PCR **Facility controls**: clinically non-suspected patients + suspected patients with negative RT-PCR **Country:** China **Dates**: 14 January 2020-9 February 2020

Feng 2020a *(Continued)*

	Symptoms and severity: all patients admitted, with exposure history to COVID-19, so all levels of severity; days from illness onset until admission (median, IQR): 2.0 (1.0-5.0); patient population with general mild disease and limited presence of comorbidities (range 0%-2.3% (COPD)) **Demographics**: age: controls median 40.0 years (IQR 32.5-54.5), cases median 39.0 years (IQR 37.0-41.5) M%/F%: cases 71.4/28.6, controls 63.2/36.8 **Exposure history**: epidemiological history of exposure to COVID-19 (as per WHO guidance)
Index tests	• Heart rate • Diastolic BP • Systolic BP • Fever (former: median only on all and cases - no control median given) • Highest temperature • Cough • Shortness of breath • Muscle ache • Headache • Sore throat • Rhinorrhoea • Diarrhoea • Nausea • Vomiting • Chills • Shiver • Expectoration • Abdominal pain • Fatigue • Palpitation
Target condition and reference standard(s)	• TC: COVID-19 pneumonia • RS: in-house RT-PCR (E-gene) - at 4 institutions
Flow and timing	Index test and RS both taken on admission
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		

Feng 2020a *(Continued)*

Did the study avoid inappropriate inclusions?	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 (All tests)			
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 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?		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?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Liang 2020

Study characteristics	
Patient Sampling	**Purpose:** to estimate the prevalence of COVID-19 in pneumonias during this period and to find the unique features of COVID-19 as compared to pneumonias caused by other agents **Design**: cross-sectional, single-centre, retrospective study

Liang 2020 *(Continued)*

	Recruitment: 342 cases of pneumonia were diagnosed in Fever Clinic in Peking University Third Hospital. From these patients, 88 were reviewed by panel discussion as possible or probable cases of COVID-19, and received 2019-nCoV detection by RT-PCR **Sample size:** n = 88 (21 cases) **Inclusion criteria**: patients visiting the Fever Clinic at Peking University Third Hospital. Based on epidemiological history, epidemiological evidence, fever and/or respiratory symptoms, chest radiological findings and WBC results, cases with possible or probable COVID-19 were sent for panel discussion and then for 2019-nCoV detection by RT-PCR **Exclusion criteria**: COVID-19 unlikely by panel discussion; lack of CT scan or no signs of pneumonia on CT scan; paediatric patients
Patient characteristics and setting	**Facility cases:** 2019-nCoV real-time PCR testing, which was positive in 19 cases (confirmed cases). In another 2 patients, though PCR testing was negative, a clinical diagnosis was made according to epidemiological evidence, consistent clinical and CT findings (clinical cases) **Facility controls**: for the cases with negative viral detection, the diagnosis of COVID-19 was excluded based on inconsistent epidemiological, clinical or radiological data **Country:** China **Dates**: 21 January 2020-15 February 2020 **Symptoms** • Fever with a mean body temperature of 37.8 C • Cough • Expectoration • Fatigue • Headache • Dizziness • Shortness of breath • Myalgia or arthralgia • Sore throat • Nasal symptoms and diarrhoea **Severity of COVID-19** • Mild-moderate: fever and/or respiratory symptoms with pneumonia in radiology examination, without signs of severe or very severe diseases • Severe: presence of 1 of the following: respiratory rate ≥ 30 beat/min; SpO_2 ≤ 93% at rest; PaO_2/FiO_2 ≤ 300 mmHg • Very severe: presence of 1 of the following: severe respiratory failure requiring mechanical ventilation; shock; complicated with other organ failure and requiring ICU admission **Demographics**: COVID-group only: median age was 42.0 years (25th-75th percentile, 34.5-66.0 years). Range 24-85. Male/female: 11 (52.4%)/10 (47.6%) **Exposure history**: 19/21 (90.5%) had a clear epidemiological history of COVID-19. 7 patients, from 5 family clusters, had close contact with their family members
Index tests	• Fever with a mean body temperature of 37.8 C • Cough • Expectoration • Fatigue • Headache • Dizziness

Liang 2020 *(Continued)*

	• Shortness of breath • Myalgia or arthralgia • Sore throat • Nasal symptoms and diarrhoea
Target condition and reference standard(s)	• TC: COVID-19 pneumonia • RS: 2019-nCoV real-time PCR testing or clinical diagnosis was made according to epidemiological evidence, consistent clinical and CT findings
Flow and timing	Time interval not specified
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		
Did the study avoid inappropriate inclusions?	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 (All tests)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it prespecified?	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 3: Reference Standard			

Liang 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?	No		
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?	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?		High risk	

Nobel 2020

Study characteristics	
Patient Sampling	**Purpose:** assess GI symptoms in COVID-19 and their association with short-term outcomes **Design**: diagnostic case-control, retrospective study **Recruitment:** adults who underwent nasopharyngeal swab testing for SARS-CoV-2 at outpatient settings: clinics or the ED, of New York-Presbyterian-Columbia or the medical centre's affiliates in New York **Sample size:** 516 (278 cases) **Inclusion criteria**: adults ≥ 18 years of age who underwent nasopharyngeal swab testing for SARS-CoV-2. Indications for testing during this period were respiratory symptoms (cough, fever, shortness of breath) with intent to hospitalise or the same symptoms in essential personnel. **Exclusion criteria**: if insufficient data were available in the electronic medical record or if testing was performed during a pre-existing inpatient admission
Patient characteristics and setting	**Facility cases:** SARS-CoV-2 PCR test result positive (1 test)

Nobel 2020 *(Continued)*

	Facility controls: SARS-CoV-2 PCR test result negative **Country:** USA **Dates**: 10 March 2020-21 March 2020 **Symptoms and severity**: respiratory symptoms (cough, fever, shortness of breath) with intent to hospitalise or in essential workers **Demographics**: median age: 51-70 years (cases and controls), gender distribution: cases (M/F(%): 52/48), controls (M/F(%): 45/55) **Exposure history**: not specified
Index tests	• GI symptoms: diarrhoea, vomiting/nausea
Target condition and reference standard(s)	• TC: SARS-Cov-2 infection • RS: SARS-CoV-2 PCR test, once (nasopharyngeal swab)
Flow and timing	Time interval: both taken at intake
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?	No		
Did the study avoid inappropriate exclusions?	Yes		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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

Nobel 2020 *(Continued)*

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?		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	

Peng 2020a

Study characteristics	
Patient Sampling	**Purpose:** analyse the clinical features and imaging manifestations of COVID-19 **Design**: cross-sectional, single-centre, retrospective study **Recruitment:** clinically suspected cases who were sent to hospital for screening **Sample size:** n = 86 (n = 11) **Inclusion criteria**: clinically suspected patients **Exclusion criteria**: not specified
Patient characteristics and setting	**Facility cases:** positive RT-PCR via nasopharyngeal swab **Facility controls**: negative RT-PCR via nasopharyngeal swab (1x) **Country:** China **Dates**: 23 January 2020-16 February 2020 **Symptoms and severity**: fever, cough, dyspnoea, sore throat, fatigue, systemic soreness, runny nose **Demographics**: M/F: total 39/47, cases: 5/6, controls 34/40 Case group: mean age 40.73 ± 11.32 years, 5 men. Control group: mean age 39.67 ± 13.90 years, 34 men

Peng 2020a *(Continued)*

	Exposure history: 7/11 COVID-19 patients (63.6%) had a history of travel to Hubei (5 Wuhan, 1 Huanggang, 1 Xiaogan), 2 patients had close contact with the COVID-19 patients, and 2 taxi drivers
Index tests	• Fever • Cough • Dyspnoea • Sore throat • Fatigue • Systemic soreness • Runny nose
Target condition and reference standard(s)	• TC: SARS-Cov-2 infection • RS: RT-PCR (nasopharyngeal swab)
Flow and timing	Time interval not specified
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		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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 3: Reference Standard			

Peng 2020a *(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?		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?	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	

Rentsch 2020

Study characteristics	
Patient Sampling	**Purpose:** diagnosis SARS-CoV-2 test positives **Design**: cross-sectional, retrospective study **Recruitment:** electronic health record data from the national Veterans Affairs Healthcare System - national Corporate Data Warehouse (USA) **Sample size:** 3789 (585 cases) **Inclusion criteria**: all patients in the Veterans Affairs cohort, born between 1945 and 1965 and active in care, tested for COVID-19 between 8 February and 30 March 2020 **Exclusion criteria**: patients for whom results were pending (n = 93) or inconclusive (n = 33) were excluded
Patient characteristics and setting	**Facility cases:** tested positive for SARS-CoV-2 **Facility controls**: tested negative for SARS-CoV-2 **Country:** USA **Dates**: 8 February 2020-30 March 2020 **Symptoms and severity**: all patients who were tested were included **Demographics**: median age overall: 65.7 years (IQR 60.5-70.7) (cases: 66.1 years, controls: 65.6 years);

Rentsch 2020 *(Continued)*

	gender overall (M%/F%): 90.2/9.8, cases 95.4/4.6, controls 89.2/10.8 **Exposure history**: not specified (all over USA)
Index tests	• Hypoxia (oxygen saturation ≤ 93%) • Body temperature (3 categories)
Target condition and reference standard(s)	• TC: SARS-CoV-2 infection • RS: no data on reference PCR test used, multiple different reference tests used with unknown test characteristics (samples: nasopharyngeal swabs)
Flow and timing	Time interval maximum 2 days
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		
Did the study avoid inappropriate inclusions?	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?			Unclear
DOMAIN 2: Index Test (All tests)			
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 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Rentsch 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?		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?	Unclear		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Song 2020b

Study characteristics	
Patient Sampling	**Purpose:** to develop a tool for early diagnosis of SARS-CoV2-infected patients **Design**: cross-sectional, retrospective, single-centre (2 time frame study: training - validation data set) **Recruitment:** 1311 patients who presented to the First Affiliated Hospital, School of Medicine, Zhejiang University with at least 1 SARS-CoV-2 RT-PCR test **Sample size:** n = 304 (73 cases) (= subset of the study including training dataset only) n = 95 (18 cases) (= validation dataset) **Inclusion criteria** • All RT-PCR-positive cases; 1311 • All RT-PCR-negative patients who came to the First Affiliated Hospital, School of Medicine, Zhejiang University and performed with at least 1 SARS-CoV-2 nucleic acid detection for analysis RT-PCR • First 60% of negative outpatients sorted by 'Z-A' based on Chinese first name from Qingchun District (training dataset), and then final 40% who presented (validation dataset) **Exclusion criteria** • Asymptomatic patients without history of exposure but had strong willingness for detection • Patients with "important" missing data
Patient characteristics and setting	**Facility cases:** positive SARS-CoV-2 **Facility controls**: negative SARS-CoV-2 **Country:** China

Song 2020b *(Continued)*

	Dates: 20 January 2020-05 February 2020
	
	Symptoms and severity: in positives: non-severe (n = 31), including mild or moderate patients to severe (n = 42) including severe or critical patients
	
	• Mild: patients had no pneumonia on imaging (CT)
	• Moderate: patients with symptoms and imaging examination showing pneumonia
	• Severe: patients meet any of the following:
	  * respiratory rate ≥ 30/min
	  * resting pulse SpO_2 ≤ 93%
	  * PaO_2/FiO2 ≤ 300 mmHg (1 mmHg = 0.133 kPa)
	  * multiple pulmonary lobes showing more than 50% progression of lesion in 24-48 hours on imaging
	• Critical: patients meet any of the following:
	  * respiratory failure requiring mechanical ventilation
	  * shock
	  * combination of other organ failure that requires admission to ICU
	
	Demographics: M/F: cases 46/27, controls 104/127
	median age: cases 53.0 years (43.5-62.0) controls 34 years (29-49)
	
	Exposure history: Wuhan-related exposure and or close contact to confirmed COV-ID-19 case: cases 40.7%, controls 57.5%

Index tests	• Fever • Cough • Expectoration • Headache • Myalgia or fatigue • Chill • Rhinobyon/rhinorrhoea • Pharyngalgia • Dyspnoea • Diarrhoea • Nausea/vomiting • Temperature (maximum) • Body temperature • SpO_2 • Respiratory rate • Heart rate • Mean arterial pressure
Target condition and reference standard(s)	• TC: SARS-CoV-2 infection • RS: RT-PCR for SARS-CoV-2 (test not specified: "using emergency use authorization approved SARS-CoV-2 assays)" (following WHO protocol, 2 target RT-PCR (ORF1 and N)
Flow and timing	Within 3 h for RS, first in-hospital stay for index tests
Comparative	
Notes	
Methodological quality	

Song 2020b *(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?	Unclear		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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 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			

Song 2020b *(Continued)*

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

Sun 2020a

Study characteristics	
Patient Sampling	**Purpose:** algorithm development for estimating risk COVID-19 **Design**: cross-sectional, retrospective study **Recruitment:** patients presenting at the designated national outbreak screening centre and tertiary care hospital in Singapore for SARS-CoV-2 testing. Patients were either self-referred, referred from primary care facilities, or were at-risk cases identified by national contact tracing efforts (recruited n = 991) **Sample size:** n = 788 (n = 54) **Inclusion criteria**: patients presenting to the centre: • self-referred • referred from primary care facilities • at-risk cases identified by national contact tracing efforts **Exclusion criteria**: PCR results not available at time of data collection - no electronic medical records - unavailable vital sign records
Patient characteristics and setting	**Facility cases:** positive SARS-CoV2 RT-PCR test **Facility controls**: all SARS-CoV-2 RT-PCR results were negative (minimum 2 test negatives in high-risk patients, minimum 1 test low-risk patients) **Country:** Singapore **Dates**: 26 January 2020-16 February 2020 **Symptoms and severity**: 252 (33.2%) symptoms > 5 days at presentation, 75 (9.5%) any comorbidity • body temperature • heart rate • respiratory rate • systolic BP • diastolic BP • cough • sputum production • shortness of breath • rhinnorhoea or nasal congestion • sore throat

Sun 2020a *(Continued)*

	• auscultation finding of pneumonia • other respiratory symptoms • gastrointestinal symptoms **Demographics**: median age 34 years (range 7 years-98 years, IQR 27-45) (cases median 42 years, range 16-79; controls 34 years (range 7-98); M/F: 48.3%/51.7% F (cases M: 88 (88.9%)) **Exposure history**: contact with a known COVID-19 case (20.1% (32/54 cases (59.3%)); 126/734 controls (17.2%), contact with travellers from China (22.1%, 15/54 cases (27.8%); 42/734 controls (5.7%)), recent travel history, and visit to hospital in China within 14 days prior to symptom onset (0.8%)
Index tests	• Body temperature • Heart rate • Respiratory rate • Systolic BP • Diastolic BP • Cough • Sputum production • Shortness of breath • Rhinnorhea or nasal congestion • Sore throat • Auscultation finding of pneumonia • Other respiratory symptoms • GI symptoms
Target condition and reference standard(s)	• TC: SARS-CoV-2 infection • RS: SARS-CoV-2 2 commercial assays 2-target (1 assay: Orf1ab and N - other unclear) RT-PCR
Flow and timing	Time interval not specified
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?	No		
Did the study avoid inappropriate exclusions?	Yes		
Did the study avoid inappropriate inclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Sun 2020a *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
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 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?	No		
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?	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	

Tolia 2020

Study characteristics	
Patient Sampling	**Purpose:** diagnosis of acute SARS-CoV-2 infection **Design**: cross-sectional, retrospective study

Tolia 2020 *(Continued)*

	Recruitment: all patients presenting to 1 of 2 EDs, located at an urban teaching hospital, and academic quaternary medical centre, within the same healthcare system who had targeted testing based on clinician's decision during the initial 10 days of test availability **Sample size:** n = 283 (29 cases) **Inclusion criteria**: • patients presenting with symptoms related to COVID-19 infection (fever and cough or shortness of breath) • travel within 14 days to countries with high rates of infection (at that time China, Iran, Italy, Japan, and South Korea) or • risk factors for infection complications (including age or comorbid conditions) or • the patient was a healthcare worker who could potentially expose others at risk and clinician made decision for testing **Exclusion criteria**: not specified
Patient characteristics and setting	**Facility cases:** positive SARS-CoV-2 test **Facility controls**: negative SARS-CoV-2 test, visiting the same EDs and being tested **Country:** USA (San Diego, CA) **Dates**: 10 March 2020-19 March 2020 **Symptoms and severity**: • all patients presenting to ED who were eligible for targeted testing (= patients presenting with symptoms related to COVID-19 infection (fever and cough or shortness of breath) • travel within 14 days to countries with high rates of infection (at that time China, Iran, Italy, Japan, and South Korea) or • risk factors for infection complications (including age or comorbid conditions) or • the patient was a healthcare worker who could potentially expose others at risk • comorbidities 101/235 (43.0%) (cases: 8/27 (29.6%), controls 93/208 (44.7%)) **Demographics**: age (< 18 years: 0.7%, 18-64 years: 83.4%, > 65 years: 15.9%); gender: cases M/F%: 55.2/44.8; controls M/F%: 52.8/47.2; all M/F%: 53.0/47.0 **Exposure history**: recent travel (5.5%), 90.6% symptom-based criteria for testing, no known exposure history based
Index tests	• Fever
Target condition and reference standard(s)	• TC: SARS-CoV-2 infection • RS: Commercial RT-PCR test - ePLex SARS-CoV-2 test (nasopharyngeal swab)
Flow and timing	Probably no time interval between index test and RS, but not specified
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Tolia 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?	Unclear		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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 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		

Tolia 2020 *(Continued)*

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

Wee 2020

Study characteristics	
Patient Sampling	**Purpose:** to analyse OTDs as a diagnostic criterion for COVID-19 **Design**: cross-sectional, prospective single-centre study **Recruitment:** all suspected cases presenting to the ED **Sample size:** n = 870 (cases = 154) **Inclusion criteria**: • presence of respiratory symptoms and suspicious epidemiological links or travel history or • new onset OTD **Exclusion criteria**: not specified
Patient characteristics and setting	**Facility cases:** positive RT-PCR for 2019-nCov **Facility controls**: negative RT-PCR for 2019-nCov **Country:** Singapore **Dates**: 26 March 2020-10 April 2020 **Symptoms and severity**: loss of sense of smell/taste **Demographics**: not specified **Exposure history**: close contact of a confirmed COVID-19 case: cases 42/112, controls 37/679
Index tests	• Loss of sense of smell/taste
Target condition and reference standard(s)	• TC: SARS-Cov-2 infection • RS: RT-PCR (oropharyngeal swabs)
Flow and timing	Time interval: same day
Comparative	
Notes	

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**

Wee 2020 *(Continued)*

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		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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 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?		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	

Yan 2020a

Study characteristics	
Patient Sampling	**Purpose:** to evaluate association of patient-reported symptoms with a focus on sense of smell and taste and SARS-CoV-2 infection **Design**: internet survey of patients after presentation to a single-centre **Recruitment:** email invitation with 1 phone call follow-up to everyone who was tested for COVID-19 between 3 March 2020 and 29 March 2020 **Sample size:** n = 262 (cases: 59) **Inclusion criteria**: • adult patients who presented to the institution and got tested for COVID-19 • analysis on responders to email survey (responses: cases 59/102, controls 203/1378) **Exclusion criteria**:
Patient characteristics and setting	**Facility cases:** SARS-CoV-2-positive **Facility controls**: SARS-CoV-2-negative **Country:** USA, San Diego **Dates**: 3 March 2020-29 March 2020 **Symptoms and severity**: • larger representation of ambulatory patients (higher response rate to survey) • severity - hospital admission: cases 4/59, controls 14/203 **Demographics**: adults only, M/F: cases 29/29, controls 69/132 **Exposure history**: not specified
Index tests	• Fatigue • Loss of taste • Fever • Loss of sense of smell • Cough • Headache • Myalgia • Dyspnoea • Diarrhoea • Nasal obstruction • Sore throat • Rhinorrhoea • Nausea
Target condition and reference standard(s)	• TC: SARS-CoV-2 infection • RS: PCR for SARS-CoV-2 (sample not specified)

Yan 2020a *(Continued)*

Flow and timing	PCR taken at presentation, not specified when the questionnaire was sent. Patients had to list their symptoms at presentation.
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		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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?	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 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?	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			

Yan 2020a *(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

Yang 2020d

Study characteristics	
Patient Sampling	**Purpose:** to identify differences in CT imaging and clinical features between COVID-19 and influenza pneumonia in the early stage, and to identify the most valuable features in the differential diagnosis **Design**: diagnostic case-control study, retrospective multicentre with historic control group **Recruitment:** cases: confirmed SARS-CoV-2 patients; controls: influenza pneumonia patients (1 January 2015-30 September 2019 from 2 hospitals) **Sample size:** n = 121 (cases = 73) **Inclusion criteria**: patients confirmed with SARS-CoV-2; controls: patients who had 9 respiratory pathogen IgM antibody tested from January 2015-September 2019 **Exclusion criteria**: cases: not specified controls: • parainfluenza • respiratory syncytial virus • adenovirus • *Legionella spp* • *Mycoplasma pneumoniae* • *Chlamydia pneumoniae* • *Coxiella burnetii* • aspiration pneumonia • radiation pneumonia • pulmonary contusion • pulmonary oedema • neoplasm No CT date, no clinical date
Patient characteristics and setting	**Facility cases:** positive RT-PCR for 2019-nCov **Facility controls**: influenza pneumonia **Country:** China

Yang 2020d *(Continued)*

	Dates: 1 January 2020-15 February 2020 **Symptoms and severity**: all patients in early stages of COVID-19 or influenza pneumonia **Demographics**: M/F: cases 41/32, controls 30/18 mean age: cases 41.9, controls 40.4 **Exposure history**: not specified
Index tests	• Body temperature • Cough • Fatigue • Sore throat • Stuffy and runny nose
Target condition and reference standard(s)	• TC: COVID-19 pneumonia • RS: RT-PCR (sample not specified)
Flow and timing	Time interval unclear
Comparative	
Notes	Overlaps with Chen X 2020

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?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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

Yang 2020d *(Continued)*

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?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
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	

Zhao 2020a

Study characteristics	
Patient Sampling	**Purpose:** to compare and assess the clinical features of COVID-19 pneumonia with features in non-COVID-19 pneumonia patients **Design**: diagnostic case control, retrospective study **Recruitment:** patients with similar duration between symptom onset to admission were selected as controls **Sample size:** n = 34 (n = 15) **Inclusion criteria**: admitted pneumonia cases with a history of travel to Hubei or exposure to a PCR SARS-CoV-2-confirmed-positive patient **Exclusion criteria**: not specified
Patient characteristics and setting	**Facility cases:** single sputum or throat swab test RT-PCR-positive pneumonia **Facility controls**: for non-COVID-19 confirmation: 3 consecutive negative throat swabs or sputum sampling every other day during first 7 days of admission **Country:** China, Anhui **Dates**: 23 January 2020-5 February 2020 **Symptoms and severity**: • fever

Zhao 2020a *(Continued)*

	• cough • sore throat • headache • fatigue • diarrhoea • chest tightness • abnormal lung auscultation **Demographics**: mean age (cases/controls): 48 (IQR 27~56)/35 (IQR 27~46) in COVID-19 and non-COVID-19 patients, respectively; F/M (cases/controls): 8 (42.11%) **Exposure history**: all patients had a history of exposure to confirmed cases of 2019-nCoV or travel to Hubei before illness. Investigators interviewed each patient and their relatives, where necessary, to determine exposure or close contact histories during the 2 weeks before the illness onset
Index tests	• Fever • Cough • Sore throat • Headache • Fatigue • Diarrhoea • Chest tightness • Abnormal lung auscultation
Target condition and reference standard(s)	• TC: COVID-19 pneumonia • RS: real-time RT-PCR (unknown assay) (sample: throat swabs or/and sputa)
Flow and timing	Time interval not specified
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?	Unclear		
Did the study avoid inappropriate inclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Zhao 2020a *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All tests)			
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 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?		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	

Zhu 2020b

Study characteristics	
Patient Sampling	**Purpose:** description of initial clinical features in patients with suspected and confirmed SARS-CoV-2 infection **Design**: cross-sectional, retrospective study **Recruitment:** all patients with suspected COVID-19 who presented to the ED of the First Affiliated Hospital of USTC and the Infectious Hospital of the First Affiliated Hospital of USTC for the first time **Sample size:** n = 116 (32 cases)

Zhu 2020b *(Continued)*

	Inclusion criteria: • patients defined as suspected SARS-CoV-2 infection based on guidelines for the diagnosis and treatment of pneumonia caused by novel coronavirus infection (trial version III) • presentation to, clinical observation and quarantine in our ED • nucleic acid amplification test performed in the ED **Exclusion criteria**: transfer from another hospital or previous visit to our hospital and previous diagnosis of COVID-19
Patient characteristics and setting	**Facility cases:** positive nucleic acid amplification test on admission or 24 h later **Facility controls**: SARs-CoV-2 PCR test negative **Country:** China, Anhui **Dates**: 24 January 2020-20 February 2020 **Symptoms and severity**: all suspected COVID-19 patients included; days since onset of symptoms median 5 (IQR 2-7) **Demographics**: median age: all: 40 years (IQR 27-53), cases: 46 years (IQR 35-52), controls: 35 years (IQR 27-53); gender distribution M%/F%: all 46/54, cases 47/53, controls 46/54 **Exposure history**: no specific exposure history common to all patients with suspected disease: 8 (25%) diagnosed patients had visited Wuhan in the previous 2 weeks and 12 (38%) had been exposed to patients with infection in the previous 2 weeks
Index tests	• Fever • Cough • Myalgia or fatigue • Experctoration • Chest stuffiness (congestion) • Haemoptysis • Headache • Diarrhoea
Target condition and reference standard(s)	• TC: SARS-CoV-2 infection • RS: nucleic acid amplification test not further specified (twice in case negatives) (samples: swabs, origin not specified)
Flow and timing	Index tests and RS both taken on admission or after 24 h
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		

Zhu 2020b *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Did the study avoid inappropriate inclusions?	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 (All tests)			
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 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?		Unclear risk	

BP: blood pressure; **COPD:** constructive obstructive pulmonary disease; **COVID-19:** coronavirus disease 2019; **CT:** computed tomography; **ED:** emergency department; **F**: female; **FiO_2:** fraction of inspired oxygen; **GI:** gastrointestinal; **ICU:** intensive care unit; **IgM:** immunoglobulin M;**IQR:** interquartile range; **M:** male; **NCP:** novel coronavirus pneumonia; **OTD:** olfactory and taste disorder; **PaO_2:** partial pressure of oxygen; **RS:** reference standard; **RT-PCR:** reverse transcription polymerase chain reaction; **SARS-CoV-2:** severe acute respiratory syndrome

coronavirus 2; **SD:** standard deviation;**SpO_2:** oxygen saturation; **TC:** target condition; **WBC:** blood white blood cell; **WHO:** World Health Organization; **2019-nCoV:** 2019 novel coronavirus

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

Study	Reason for exclusion
Guan 2020	SARS-CoV-2-positive cases only
Soares 2020	No data
Song 2020	SARS-CoV-2-positive cases only
Wang 2020	No data

D A T A

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

Table Tests. Data tables by test

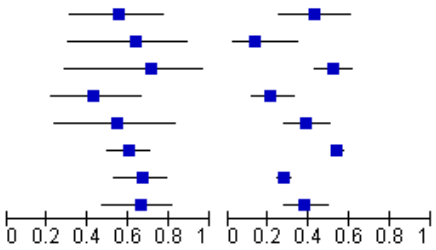
Test	No. of studies	No. of participants
1 Cough	8	2607
2 Sputum production	6	2467
3 Dyspnoea	7	2554
4 Hypoxia	1	2929
5 Haemoptysis	1	116
6 Positive auscultation findings	1	788
7 Respiratory symptoms (not specified))	1	788
8 Sore throat	6	2438
9 Nasal symptoms	5	2405
10 Loss of smell (anosmia) or loss of taste (ageusia)	1	870
11 Fever	9	5484
12 Low body temperature	1	3384
13 Shivers	1	132
14 Chills	2	1443
15 Myalgia or arthralgia	4	339

Test	No. of studies	No. of participants
16 Myalgia or fatigue	2	1427
17 Fatigue	3	273
18 Headache	5	1700
19 Nausea/vomiting	3	489
20 Diarrhoea	6	1733
21 Abdominal pain	2	185
22 Gastrointestinal symptoms (not specified)	1	788
23 Low systolic blood pressure	1	3341
24 High systolic blood pressure	1	3341
25 Tachycardia	1	3373
26 Palpitations	1	132
27 Cough (non-cross-sectional study)	3	432
28 Sore throat (non-cross-sectional study)	3	432
29 Rhinorrhoea (non-cross-sectional study)	1	136
30 Nasal obstruction (non-cross-sectional study)	2	398
31 Loss of smell (anosmia) (non-cross-sectional study)	1	262
32 Loss of taste (ageusia) (non-cross-sectional study)	1	262
33 Positive auscultation findings (non-cross-sectional study)	1	34
34 Dyspnoea (non-cross-sectional study)	1	262
35 Chest tightness (non-cross-sectional study)	1	34
36 Fever (non-cross-sectional study)	2	296
37 Fatigue (non-cross-sectional study)	3	432
38 Myalgia or arthralgia (non-cross-sectional study)	1	262
39 Headache (non-cross-sectional study)	2	296
40 Diarrhoea (non-cross-sectional study)	3	812
41 Nausea/vomiting (non-cross-sectional study)	2	778
42 Gastrointestinal symptoms, not specified (non-cross-sectional study)	1	516

Test 1. Cough

Cough

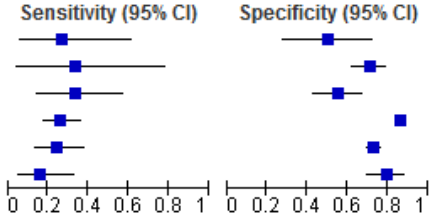
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	11	19	9	14	0.55 [0.32, 0.77]	0.42 [0.25, 0.61]
Cheng 2020a	7	19	4	3	0.64 [0.31, 0.89]	0.14 [0.03, 0.35]
Feng 2020a	5	60	2	65	0.71 [0.29, 0.96]	0.52 [0.43, 0.61]
Liang 2020	9	53	12	14	0.43 [0.22, 0.66]	0.21 [0.12, 0.33]
Peng 2020a	6	46	5	29	0.55 [0.23, 0.83]	0.39 [0.28, 0.51]
Song 2020b	55	562	36	658	0.60 [0.50, 0.71]	0.54 [0.51, 0.57]
Sun 2020a	36	528	18	206	0.67 [0.53, 0.79]	0.28 [0.25, 0.31]
Zhu 2020b	21	52	11	32	0.66 [0.47, 0.81]	0.38 [0.28, 0.49]



Test 2. Sputum production

Sputum production

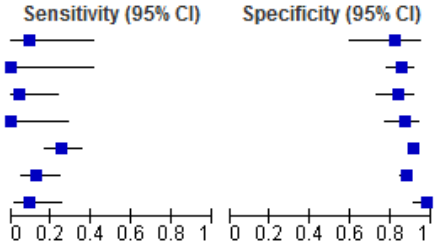
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Cheng 2020a	3	11	8	11	0.27 [0.06, 0.61]	0.50 [0.28, 0.72]
Feng 2020a	2	36	4	89	0.33 [0.04, 0.78]	0.71 [0.62, 0.79]
Liang 2020	7	30	14	37	0.33 [0.15, 0.57]	0.55 [0.43, 0.67]
Song 2020b	24	166	67	1054	0.26 [0.18, 0.37]	0.86 [0.84, 0.88]
Sun 2020a	13	199	41	535	0.24 [0.13, 0.38]	0.73 [0.70, 0.76]
Zhu 2020b	5	17	27	67	0.16 [0.05, 0.33]	0.80 [0.70, 0.88]



Test 3. Dyspnoea

Dyspnoea

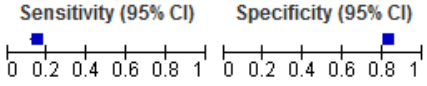
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Cheng 2020a	1	4	10	18	0.09 [0.00, 0.41]	0.82 [0.60, 0.95]
Feng 2020a	0	18	7	107	0.00 [0.00, 0.41]	0.86 [0.78, 0.91]
Liang 2020	1	11	20	56	0.05 [0.00, 0.24]	0.84 [0.73, 0.92]
Peng 2020a	0	10	11	65	0.00 [0.00, 0.28]	0.87 [0.77, 0.93]
Song 2020b	23	111	68	1109	0.25 [0.17, 0.35]	0.91 [0.89, 0.92]
Sun 2020a	7	93	47	641	0.13 [0.05, 0.25]	0.87 [0.85, 0.90]
Zhu 2020b	3	2	29	82	0.09 [0.02, 0.25]	0.98 [0.92, 1.00]



Test 4. Hypoxia

Hypoxia

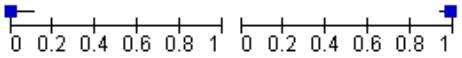
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rentsch 2020	78	418	443	1990	0.15 [0.12, 0.18]	0.83 [0.81, 0.84]



Test 5. Haemoptysis

Haemoptysis

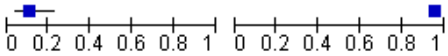
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Zhu 2020b	0	1	32	83	0.00 [0.00, 0.11]	0.99 [0.94, 1.00]		



Test 6. Positive auscultation findings

Positive auscultation findings

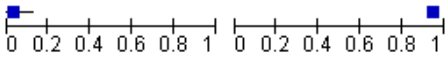
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Sun 2020a	6	36	48	698	0.11 [0.04, 0.23]	0.95 [0.93, 0.97]		



Test 7. Respiratory symptoms (not specified))

Respiratory symptoms (not specified))

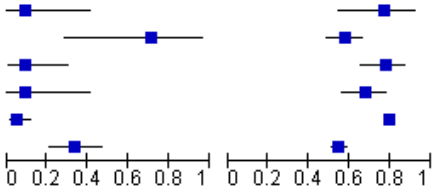
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Sun 2020a	2	43	52	691	0.04 [0.00, 0.13]	0.94 [0.92, 0.96]		



Test 8. Sore throat

Sore throat

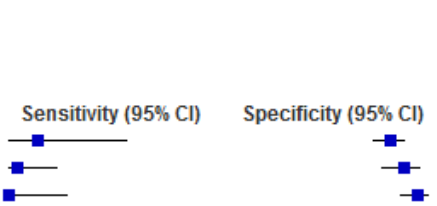
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Cheng 2020a	1	5	10	17	0.09 [0.00, 0.41]	0.77 [0.55, 0.92]		
Feng 2020a	5	53	2	72	0.71 [0.29, 0.96]	0.58 [0.48, 0.66]		
Liang 2020	2	15	19	52	0.10 [0.01, 0.30]	0.78 [0.66, 0.87]		
Peng 2020a	1	24	10	51	0.09 [0.00, 0.41]	0.68 [0.56, 0.78]		
Song 2020b	5	250	86	970	0.05 [0.02, 0.12]	0.80 [0.77, 0.82]		
Sun 2020a	18	332	36	402	0.33 [0.21, 0.47]	0.55 [0.51, 0.58]		



Test 9. Nasal symptoms

Nasal symptoms

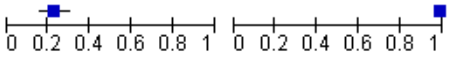
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Feng 2020a	1	27	6	98	0.14 [0.00, 0.58]	0.78 [0.70, 0.85]		
Liang 2020	1	10	20	57	0.05 [0.00, 0.24]	0.85 [0.74, 0.93]		
Peng 2020a	0	6	11	69	0.00 [0.00, 0.28]	0.92 [0.83, 0.97]		
Song 2020b	1	107	90	1113	0.01 [0.00, 0.06]	0.91 [0.90, 0.93]		
Sun 2020a	12	226	42	508	0.22 [0.12, 0.36]	0.69 [0.66, 0.73]		



Test 10. Loss of smell (anosmia) or loss of taste (ageusia)

Loss of smell (anosmia) or loss of taste (ageusia)

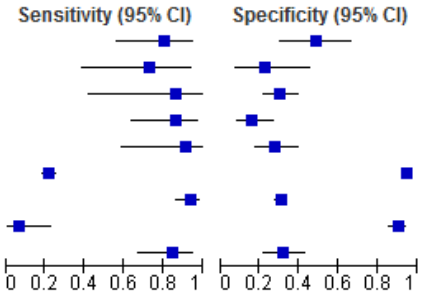
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Wee 2020	35	9	119	707	0.23 [0.16, 0.30]	0.99 [0.98, 0.99]		



Test 11. Fever

Fever

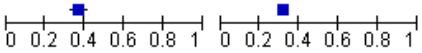
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	16	17	4	16	0.80 [0.56, 0.94]	0.48 [0.31, 0.66]		
Cheng 2020a	8	17	3	5	0.73 [0.39, 0.94]	0.23 [0.08, 0.45]		
Feng 2020a	6	87	1	38	0.86 [0.42, 1.00]	0.30 [0.22, 0.39]		
Liang 2020	18	56	3	11	0.86 [0.64, 0.97]	0.16 [0.08, 0.27]		
Peng 2020a	10	54	1	21	0.91 [0.59, 1.00]	0.28 [0.18, 0.40]		
Rentsch 2020	120	169	431	2664	0.22 [0.18, 0.25]	0.94 [0.93, 0.95]		
Song 2020b	85	844	6	376	0.93 [0.86, 0.98]	0.31 [0.28, 0.33]		
Tolia 2020	2	25	27	227	0.07 [0.01, 0.23]	0.90 [0.86, 0.93]		
Zhu 2020b	27	57	5	27	0.84 [0.67, 0.95]	0.32 [0.22, 0.43]		



Test 12. Low body temperature

Low body temperature

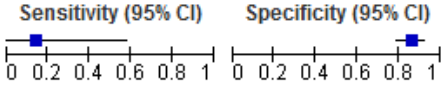
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Rentsch 2020	204	1938	347	895	0.37 [0.33, 0.41]	0.32 [0.30, 0.33]		



Test 13. Shivers

Shivers

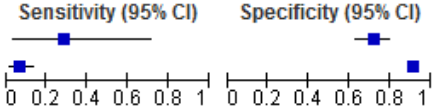
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Feng 2020a	1	17	6	108	0.14 [0.00, 0.58]	0.86 [0.79, 0.92]		



Test 14. Chills

Chills

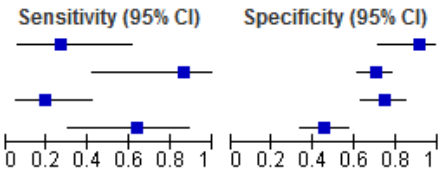
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Feng 2020a	2	35	5	90	0.29 [0.04, 0.71]	0.72 [0.63, 0.80]		
Song 2020b	6	111	85	1109	0.07 [0.02, 0.14]	0.91 [0.89, 0.92]		



Test 15. Myalgia or arthralgia

Myalgia or arthralgia

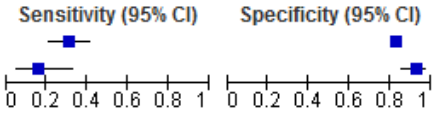
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Cheng 2020a	3	2	8	20	0.27 [0.06, 0.61]	0.91 [0.71, 0.99]
Feng 2020a	6	37	1	88	0.86 [0.42, 1.00]	0.70 [0.62, 0.78]
Liang 2020	4	17	17	50	0.19 [0.05, 0.42]	0.75 [0.63, 0.84]
Peng 2020a	7	41	4	34	0.64 [0.31, 0.89]	0.45 [0.34, 0.57]



Test 16. Myalgia or fatigue

Myalgia or fatigue

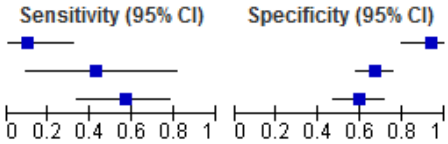
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Song 2020b	28	214	63	1006	0.31 [0.22, 0.41]	0.82 [0.80, 0.85]
Zhu 2020b	5	6	27	78	0.16 [0.05, 0.33]	0.93 [0.85, 0.97]



Test 17. Fatigue

Fatigue

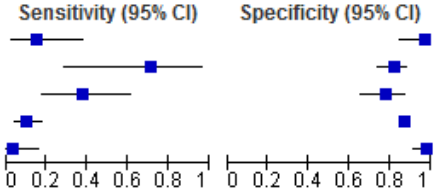
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	2	2	18	31	0.10 [0.01, 0.32]	0.94 [0.80, 0.99]
Feng 2020a	3	41	4	84	0.43 [0.10, 0.82]	0.67 [0.58, 0.75]
Liang 2020	12	27	9	40	0.57 [0.34, 0.78]	0.60 [0.47, 0.72]



Test 18. Headache

Headache

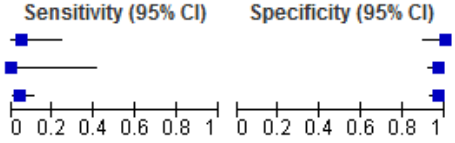
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	3	1	17	32	0.15 [0.03, 0.38]	0.97 [0.84, 1.00]
Feng 2020a	5	23	2	102	0.71 [0.29, 0.96]	0.82 [0.74, 0.88]
Liang 2020	8	15	13	52	0.38 [0.18, 0.62]	0.78 [0.66, 0.87]
Song 2020b	9	158	82	1062	0.10 [0.05, 0.18]	0.87 [0.85, 0.89]
Zhu 2020b	1	2	31	82	0.03 [0.00, 0.16]	0.98 [0.92, 1.00]



Test 19. Nausea/vomiting

Nausea/vomiting

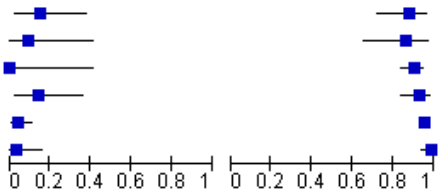
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	1	0	19	33	0.05 [0.00, 0.25]	1.00 [0.89, 1.00]
Feng 2020a	0	4	7	121	0.00 [0.00, 0.41]	0.97 [0.92, 0.99]
Song 2020b	3	8	70	223	0.04 [0.01, 0.12]	0.97 [0.93, 0.98]



Test 20. Diarrhoea

Diarrhoea

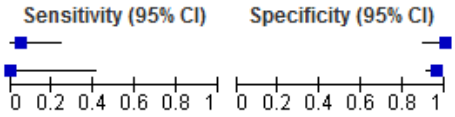
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	3	4	17	29	0.15 [0.03, 0.38]	0.88 [0.72, 0.97]
Cheng 2020a	1	3	10	19	0.09 [0.00, 0.41]	0.86 [0.65, 0.97]
Feng 2020a	0	12	7	113	0.00 [0.00, 0.41]	0.90 [0.84, 0.95]
Liang 2020	3	5	18	62	0.14 [0.03, 0.36]	0.93 [0.83, 0.98]
Song 2020b	4	55	87	1165	0.04 [0.01, 0.11]	0.95 [0.94, 0.97]
Zhu 2020b	1	1	31	83	0.03 [0.00, 0.16]	0.99 [0.94, 1.00]



Test 21. Abdominal pain

Abdominal pain

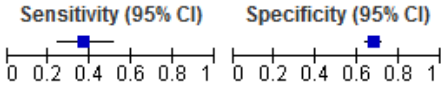
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ai 2020a	1	0	19	33	0.05 [0.00, 0.25]	1.00 [0.89, 1.00]
Feng 2020a	0	5	7	120	0.00 [0.00, 0.41]	0.96 [0.91, 0.99]



Test 22. Gastrointestinal symptoms (not specified)

Gastrointestinal symptoms (not specified)

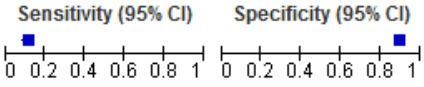
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Sun 2020a	20	238	34	496	0.37 [0.24, 0.51]	0.68 [0.64, 0.71]



Test 23. Low systolic blood pressure

Low systolic blood pressure

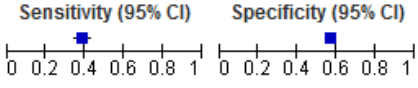
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rentsch 2020	63	292	485	2501	0.11 [0.09, 0.14]	0.90 [0.88, 0.91]



Test 24. High systolic blood pressure

High systolic blood pressure

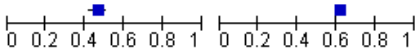
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rentsch 2020	211	1210	337	1583	0.39 [0.34, 0.43]	0.57 [0.55, 0.59]



Test 25. Tachycardia

Tachycardia

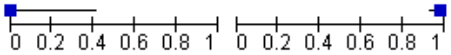
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Rentsch 2020	257	1083	295	1738	0.47 [0.42, 0.51]	0.62 [0.60, 0.63]		



Test 26. Palpitations

Palpitations

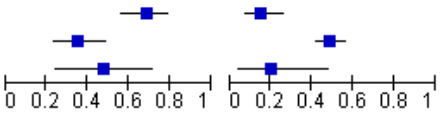
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Feng 2020a	0	3	7	122	0.00 [0.00, 0.41]	0.98 [0.93, 1.00]		



Test 27. Cough (non-cross-sectional study)

Cough (non-cross-sectional study)

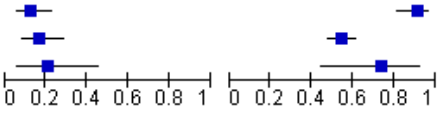
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Chen X 2020	48	56	22	10	0.69 [0.56, 0.79]	0.15 [0.08, 0.26]		
Yan 2020a	21	104	38	99	0.36 [0.24, 0.49]	0.49 [0.42, 0.56]		
Zhao 2020a	9	12	10	3	0.47 [0.24, 0.71]	0.20 [0.04, 0.48]		



Test 28. Sore throat (non-cross-sectional study)

Sore throat (non-cross-sectional study)

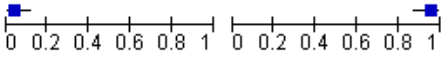
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Chen X 2020	9	6	61	60	0.13 [0.06, 0.23]	0.91 [0.81, 0.97]		
Yan 2020a	10	92	49	111	0.17 [0.08, 0.29]	0.55 [0.48, 0.62]		
Zhao 2020a	4	4	15	11	0.21 [0.06, 0.46]	0.73 [0.45, 0.92]		



Test 29. Rhinorrhoea (non-cross-sectional study)

Rhinorrhoea (non-cross-sectional study)

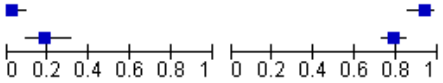
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Chen X 2020	3	3	67	63	0.04 [0.01, 0.12]	0.95 [0.87, 0.99]		



Test 30. Nasal obstruction (non-cross-sectional study)

Nasal obstruction (non-cross-sectional study)

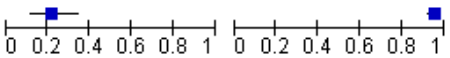
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Chen X 2020	2	4	68	62	0.03 [0.00, 0.10]	0.94 [0.85, 0.98]		
Yan 2020a	11	43	48	160	0.19 [0.10, 0.31]	0.79 [0.73, 0.84]		



Test 31. Loss of smell (anosmia) (non-cross-sectional study)

Loss of smell (anosmia) (non-cross-sectional study)

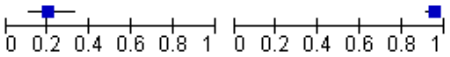
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Yan 2020a	13	9	46	194	0.22 [0.12, 0.35]	0.96 [0.92, 0.98]		



Test 32. Loss of taste (ageusia) (non-cross-sectional study)

Loss of taste (ageusia) (non-cross-sectional study)

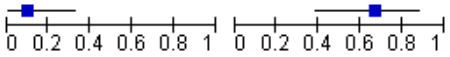
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Yan 2020a	12	10	47	193	0.20 [0.11, 0.33]	0.95 [0.91, 0.98]		



Test 33. Positive auscultation findings (non-cross-sectional study)

Positive auscultation findings (non-cross-sectional study)

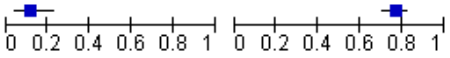
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Zhao 2020a	2	5	17	10	0.11 [0.01, 0.33]	0.67 [0.38, 0.88]		



Test 34. Dyspnoea (non-cross-sectional study)

Dyspnoea (non-cross-sectional study)

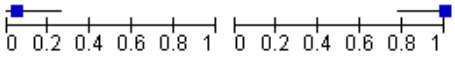
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Yan 2020a	7	47	52	156	0.12 [0.05, 0.23]	0.77 [0.70, 0.82]		



Test 35. Chest tightness (non-cross-sectional study)

Chest tightness (non-cross-sectional study)

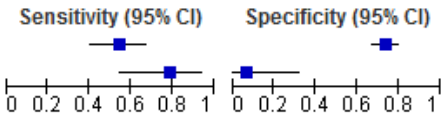
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Zhao 2020a	1	0	18	15	0.05 [0.00, 0.26]	1.00 [0.78, 1.00]		



Test 36. Fever (non-cross-sectional study)

Fever (non-cross-sectional study)

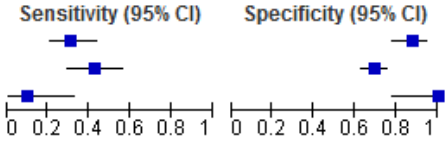
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Yan 2020a	32	53	27	150	0.54 [0.41, 0.67]	0.74 [0.67, 0.80]
Zhao 2020a	15	14	4	1	0.79 [0.54, 0.94]	0.07 [0.00, 0.32]



Test 37. Fatigue (non-cross-sectional study)

Fatigue (non-cross-sectional study)

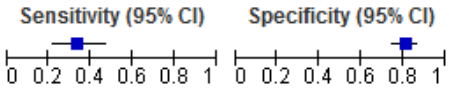
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Chen X 2020	22	8	48	58	0.31 [0.21, 0.44]	0.88 [0.78, 0.95]
Yan 2020a	25	62	34	141	0.42 [0.30, 0.56]	0.69 [0.63, 0.76]
Zhao 2020a	2	0	17	15	0.11 [0.01, 0.33]	1.00 [0.78, 1.00]



Test 38. Myalgia or arthralgia (non-cross-sectional study)

Myalgia or arthralgia (non-cross-sectional study)

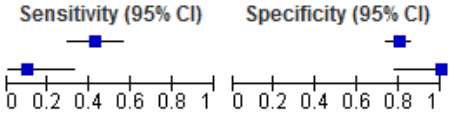
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Yan 2020a	20	39	39	164	0.34 [0.22, 0.47]	0.81 [0.75, 0.86]



Test 39. Headache (non-cross-sectional study)

Headache (non-cross-sectional study)

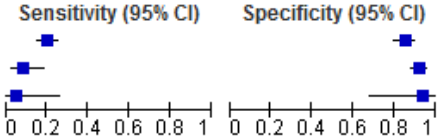
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Yan 2020a	25	40	34	163	0.42 [0.30, 0.56]	0.80 [0.74, 0.86]
Zhao 2020a	2	0	17	15	0.11 [0.01, 0.33]	1.00 [0.78, 1.00]



Test 40. Diarrhoea (non-cross-sectional study)

Diarrhoea (non-cross-sectional study)

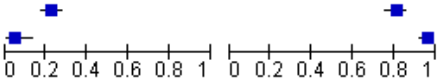
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Nobel 2020	56	36	222	202	0.20 [0.16, 0.25]	0.85 [0.80, 0.89]
Yan 2020a	5	16	54	187	0.08 [0.03, 0.19]	0.92 [0.88, 0.95]
Zhao 2020a	1	1	18	14	0.05 [0.00, 0.26]	0.93 [0.68, 1.00]



Test 41. Nausea/vomiting (non-cross-sectional study)

Nausea/vomiting (non-cross-sectional study)

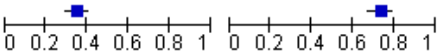
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Nobel 2020	63	46	215	192	0.23 [0.18, 0.28]	0.81 [0.75, 0.85]
Yan 2020a	3	8	56	195	0.05 [0.01, 0.14]	0.96 [0.92, 0.98]



Test 42. Gastrointestinal symptoms, not specified (non-cross-sectional study)

Gastrointestinal symptoms, not specified (non-cross-sectional study)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Nobel 2020	97	63	181	175	0.35 [0.29, 0.41]	0.74 [0.67, 0.79]



ADDITIONAL TABLES

Table 1. QUADAS-2 checklist

Index test(s)	Signs and symptoms
Patients (setting, intended use of index test, presentation, prior testing)	Primary care, hospital outpatient settings including emergency departments Inpatients presenting with suspected COVID-19 No prior testing Signs and symptoms often used for triage or referral
Reference standard and target condition	The focus will be on the diagnosis of COVID-19 disease and COVID-19 pneumonia. For this review, the focus will not be on prognosis.
Participant selection	
Was a consecutive or random sample of patients enrolled?	This will be similar for all index tests, target conditions, and populations. YES: if a study explicitly stated 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 was clear that a different selection procedure was employed; for example, selection based on clinician's preference, or based on institutions. UNCLEAR: if the selection procedure was not clear or not reported.
Was a case-control design avoided?	This will be similar for all index tests, target conditions, and populations. YES: if a study explicitly stated that all participants came from the same group of (suspected) patients. NO: if it was clear that a different selection procedure was employed for the participants depending on their COVID-19 (pneumonia) status or SARS-CoV-2 infection status. UNCLEAR: if the selection procedure was not clear or not reported.

Table 1. QUADAS-2 checklist *(Continued)*

Did the study avoid inappropriate exclusions?	Studies may have excluded participants, or selected participants in such a way that they avoided including those who were difficult to diagnose or likely to be borderline. Although the inclusion and exclusion criteria will be different for the different index tests, inappropriate exclusions and inclusions will be similar for all index tests: for example, only elderly patients excluded, or children (as sampling may be more difficult). This needs to be addressed on a case-by-case basis. YES: if a high proportion of eligible patients was included without clear selection. NO: if a high proportion of eligible patients was excluded without providing a reason; if, in a retrospective study, participants without index test or reference standard results were excluded; if exclusion was based on severity assessment post-factum or comorbidities (cardiovascular disease, diabetes, immunosuppression). UNCLEAR: if the exclusion criteria were not reported.
Did the study avoid inappropriate inclusions?	YES: if samples included were likely to be representative of the spectrum of disease. NO: if the study oversampled patients with particular characteristics likely to affect estimates of accuracy. UNCLEAR: if the exclusion criteria were 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?	HIGH: if accuracy of signs and symptoms were assessed in a case-control design, or in an already highly selected group of participants, or the study was able to only estimate sensitivity or specificity. LOW: any situation where signs and symptoms were the first assessment/test to be done on the included participants. UNCLEAR: if a description about the participants was lacking.
Index tests	
Were the index test results interpreted without knowledge of the results of the reference standard?	This will be similar for all index tests, target conditions, and populations. 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. UNCLEAR: if blinding was unclearly reported.
If a threshold was used, was it prespecified?	This will be similar for all index tests, target conditions, and populations. YES: if the test was dichotomous by nature, or if the threshold was stated in the methods section, or if authors stated that the threshold as recommended by the manufacturer was used. NO: if a receiver operating characteristic curve was drawn or multiple threshold reported in the results section; and the final result was based on one of these thresholds; if fever was not defined beforehand. UNCLEAR: if threshold selection was not clearly reported.

Table 1. QUADAS-2 checklist *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?	HIGH: if one or more signalling questions were answered with NO, as even in a laboratory situation knowledge of the reference standard may lead to bias. 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?	This will probably be answered 'LOW' in all cases except when assessments were made in a different setting, or using personnel not available in practice.
Reference standard	
Is the reference standard likely to correctly classify the target condition?	We will define acceptable reference standards using a consensus process once the list of reference standards that have been used has been obtained from the eligible studies. For severe pneumonia, we will consider how well processes adhered to the WHO case definition in Appendix 1.
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. UNCLEAR: if blinding was unclearly reported.
Did the definition of the reference standard incorporate results from the index test(s)?	YES: if results from the index test were a component of the reference standard definition. NO: if the reference standard did not incorporate the index standard test. UNCLEAR: if it was unclear whether the results of the index test formed part of the reference standard.
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: if the target condition was COVID-19 pneumonia, but only RT-PCR was used; if alternative diagnosis was highly likely and not excluded (will happen in paediatric cases, where exclusion of other respiratory pathogens is also necessary); if tests used to follow up viral load in known test-positives. LOW: if above situations were not present. UNCLEAR: if intention for testing was not reported in the study.
Flow and timing	
Was there an appropriate interval between index test(s) and reference standard?	YES: this will be similar for all index tests, populations for the current infection target conditions: as the situation of a patient, including clinical presentation and disease progress, evolves rapidly and new/ongoing exposure can result in case status change, an appropriate time interval will be within 24 hours. NO: if there was more than 24 hours between the index test and the reference standard or if participants were otherwise reported to be assessed with the index versus reference standard test at moments of different severity.

Table 1. QUADAS-2 checklist *(Continued)*

	UNCLEAR: if the time interval was not reported.
Did all patients receive a reference standard?	YES: if all participants 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 was not reported.
Did all patients receive the same reference standard?	YES: if all participants received the same reference standard (clearly no differential verification). NO: if (part of) the index test-positives or index test-negatives received a different reference standard. UNCLEAR: if it was not reported.
Were all patients included in the analysis?	YES: if all included participants were included in the analyses. 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 was not clear from the reported numbers.
Could the patient flow 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.

ICU: intensive care unit; **RT-PCR:** reverse transcription polymerase chain reaction; **SARS-CoV-2:** severe acute respiratory syndrome coronavirus 2; **WHO:** World Health Organization

Table 2. Summary of study characteristics

Study ID	Target condition	Sample size	Prevalence	Setting	Population	Design	Reference standard
Ai 2020a	COVID-19 pneumonia	53	38%	Hospital inpatients[a]	Patients hospitalised with pneumonia diagnosed by imaging	Cross-sectional	PCR on nasopharyngeal swabs
Chen X 2020	COVID-19 pneumonia	136	Not applicable	Hospital inpatients[a]	Patients admitted with pneumonia	Cases selected cross-sectionally in 5 hospitals, non-cases from 1 hospital only	PCR, samples not specified
Cheng 2020a	COVID-19 pneumonia	33	33%	Hospital outpatients	Patients presenting to a fever observation department with pneumonia	Cross-sectional	PCR on throat swabs
Feng 2020a	COVID-19 pneumonia	132	5%	Emergency department	Patients presenting to fever clinic of emergency department	Cross-sectional	PCR on throat swabs
Liang 2020	COVID-19 pneumonia	88	24%	Hospital outpatients	Patients with pneumonia and presenting to fever clinic	Cross-sectional	PCR, sample not specified; conducted after panel discussion
Nobel 2020	COVID-19 disease	516	Not applicable	Hospital outpatients	Patients who underwent SARS-CoV-2 testing with intent to hospitalise or in essential personnel	Case-control	PCR on nasopharyngeal swabs
Peng 2020a	COVID-19 disease	86	13%	Hospital outpatients	Patients clinically suspected and referred for testing	Cross-sectional	PCR on nasopharyngeal swabs
Rentsch 2020	COVID-19 disease	3789	15%	Unclear	Patients tested for SARS-CoV-2 in the Veterans Affairs Cohort born between 1945 and 1965	Cross-sectional	PCR on nasopharyngeal swabs

Table 2. Summary of study characteristics *(Continued)*

Song 2020b	COVID-19 disease	399	7%	Hospital outpatients	Patients tested for SARS-CoV-2	Cross-sectional	PCR on sputum samples
Sun 2020a	COVID-19 disease	788	7%	Hospital outpatients	Patients presenting to testing centre, either self-referred, referred from primary care or at-risk cases identified by national contact tracing	Cross-sectional	PCR on sputum, endotracheal aspirate, nasopharyngeal swabs or throat swabs
Tolia 2020	COVID-19 disease	283	10%	Emergency department	Patients presenting with symptoms, travel history, risk factors or healthcare workers	Cross-sectional	PCR on nasopharyngeal swabs
Wee 2020	COVID-19 disease	870	18%	Emergency department	Patients presenting with respiratory symptoms or travel history	Cross-sectional	PCR on oropharyngeal swabs
Yan 2020a	COVID-19 disease	262	23%	Hospital outpatient	Patients presenting hospital for SARS-CoV-2 testing, not otherwise specified	Internet survey after presentation	PCR, samples not specified
Yang 2020d	COVID-19 pneumonia	121	Not applicable	Hospital inpatients[a]	Patient with pneumonia from SARS-CoV-2 and patients with pneumonia from influenza in 2015-2019	Case-control	PCR, samples not specified
Zhao 2020a	COVID-19 pneumonia	34	Not applicable	Hospital inpatients[a]	Patients with pneumonia and admitted to hospital	Case-control	PCR on throat or sputum swabs
Zhu 2020b	COVID-19 disease	116	28%	Emergency department	Patients suspected of SARS-CoV-2 and presenting to the emergency department	Cross-sectional	PCR, samples not specified

PCR: polymerase chain reaction; **SARS-CoV-2:** severe acute respiratory syndrome coronavirus 2

[a] Hospital inpatients' refers to studies that recruited patients admitted to hospital with COVID-19 disease and in whom the signs and symptoms were assessed on admission.

APPENDICES

Appendix 1. World Health Organization case definitions

Severe pneumonia

Adolescent or adult: fever or suspected respiratory infection, plus one of the following: respiratory rate > 30 breaths/minute; severe respiratory distress; or oxygen saturation (SpO_2) ≤ 93% on room air. Child with cough or difficulty in breathing, plus at least one of the following: central cyanosis or SpO_2 < 90%; severe respiratory distress (for example, grunting, very severe chest indrawing); signs of pneumonia with a general danger sign: inability to breastfeed or drink, lethargy or unconsciousness, or convulsions.

Other signs of pneumonia may be present: chest indrawing, fast breathing (in breaths/minute): aged < 2 months: ≥ 60; aged 2 to 11 months: ≥ 50; aged 1 to 5 years: ≥ 40. While the diagnosis is made on clinical grounds; chest imaging may identify or exclude some pulmonary complications.

Acute respiratory distress syndrome (ARDS)

Onset within one week of a known clinical insult or new or worsening respiratory symptoms.

Chest imaging (that is, X-ray, computed tomography scan, or lung ultrasound): bilateral opacities, not fully explained by volume overload, lobar or lung collapse, or nodules.

Origin of pulmonary infiltrates: respiratory failure not fully explained by cardiac failure or fluid overload. Need objective assessment (for example, echocardiography) to exclude hydrostatic cause of infiltrates/oedema if no risk factor present.

Oxygenation impairment in adults:

- mild acute respiratory distress syndrome (ARDS): 200 mmHg < ratio of arterial oxygen partial pressure/fractional inspired oxygen (PaO_2/FiO_2) ≤ 300 mmHg (with positive end-expiratory pressure (PEEP) or continuous positive airway pressure (CPAP) ≥ 5 cmH_2O, or non-ventilated);
- moderate ARDS: 100 mmHg < PaO_2/FiO_2 ≤ 200 mmHg (with PEEP ≥ 5 cmH_2O, or non-ventilated);
- severe ARDS: PaO_2/FiO_2 ≤ 100 mmHg (with PEEP ≥ 5 cmH_2O, or non-ventilated);
- when PaO_2 is not available, SpO_2/FiO_2 ≤ 315 mmHg suggests ARDS (including in non-ventilated patients).

Oxygenation impairment in children: note OI = Oxygenation Index and OSI = Oxygenation Index using SpO_2. Use PaO_2-based metric when available. If PaO_2 not available, wean FiO_2 to maintain SpO_2 ≤ 97% to calculate OSI or SpO_2/FiO_2 ratio:

- bilevel (non-invasive ventilation or CPAP) ≥ 5 cmH_2O via full-face mask: PaO_2/FiO_2 ≤ 300 mmHg or SpO_2/FiO_2 ≤ 264;
- mild ARDS (invasively ventilated): 4 ≤ OI < 8 or 5 ≤ OSI < 7.5;
- moderate ARDS (invasively ventilated): 8 ≤ OI < 16 or 7.5 ≤ OSI < 12.3;
- severe ARDS (invasively ventilated): OI ≥ 16 or OSI ≥ 12.3.

Appendix 2. Cochrane COVID-19 Study Register searches

Source	Strategy
ClinicalTrials.gov	COVID-19[a]
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 corona 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])

[a]Automatic term mapping links results for 2019-nCoV, 2019 novel coronavirus, SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

Appendix 3. Living search from the University of Bern

We took the following information from the university of Bern website (see: ispmbern.github.io/covid-19/living-review/collectingdata.html).

The register is updated daily and CSV file downloads are made available.

1 April 2020

From 1 April 2020, we will retriev the curated BioRxiv/MedRxiv dataset (connect.medrxiv.org/relate/content/181).

26 to 31 March 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 (www.unibe.ch/university/services/university_library/faculty_libraries/medicine/public_health_amp_primary_care_library_phc/index_eng.html), and following guidance of the Medical Library Association (www.mlanet.org/p/cm/ld/fid=1713).

1 January 2020 to 25 March 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

Appendix 4. 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
7 July 2020	Amended	Resolution of two figures improved

HISTORY

Review first published: Issue 7, 2020

CONTRIBUTIONS OF AUTHORS

JD, JDi, YT, CD, ML, RS, LH, AV, DE, and SD contributed clinical, methodological and/or technical expertise to drafting the protocol. JD co-ordinated contributions from all co-authors and drafted the protocol. ML drafted the QUADAS-2 criteria. AVDB oversaw the overall progress of this review, drafted the manuscript and participated in the selection and data extraction. TS participated in the data extraction, analyses and drafting of the manuscript. JD and BH also participated in the data extraction, interpretation of the findings and commented on the manuscript.

DECLARATIONS OF INTEREST

Thomas Struyf: none known

Jonathan J Deeks: none known

Jacqueline Dinnes: none known

Yemisi Takwoingi: none known

Clare Davenport: none known

Mariska MG Leeflang: none known

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.

Lotty Hooft: none known

Devy Emperador: is employed by FIND. FIND has several clinical research projects to evaluate multiple new diagnostic tests against published Target Product Profiles that have been defined through consensus processes. These studies are for diagnostic products developed by private sector companies who provide access to know-how, equipment/reagents, and contribute through unrestricted donations as per FIND policy and external SAC review.

Sabine Dittrich: is employed by FIND with funding from DFID and Australian Aid. FIND is a global non-for profit product development partnership and WHO Diagnostic Collaboration Centre. It is FIND's role to accelerate access to high quality diagnostic tools for low resource settings and this is achieved by supporting both R&D and access activities for a wide range of diseases, including COVID-19. .FIND has several clinical research projects to evaluate multiple new diagnostic tests against published Target Product Profiles that have been defined through consensus processes. These studies are for diagnostic products developed by private sector companies who provide access to know-how, equipment/reagents, and contribute through unrestricted donations as per FIND policy and external SAC review.

Julie Domen: none known

Sebastiaan Horn: none known

Ann Van den Bruel: none known

SOURCES OF SUPPORT

Internal sources

- Liverpool School of Tropical Medicine, UK
- University of Birmingham, UK

External sources

- Department for International Development, UK

  Project number: 300342-104
- National Institute for Health Research (NIHR), UK
- NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham, UK

INDEX TERMS

Medical Subject Headings (MeSH)

*Ambulatory Care; Arthralgia [diagnosis] [etiology]; *Betacoronavirus; Coronavirus Infections [complications] [*diagnosis] [epidemiology]; COVID-19; Fatigue [diagnosis] [etiology]; Fever [diagnosis] [etiology]; Headache [diagnosis]; Myalgia [diagnosis] [etiology]; Outpatient Clinics, Hospital [statistics & numerical data]; Pandemics; Physical Examination; Pneumonia, Viral [complications] [*diagnosis] [epidemiology]; *Primary Health Care; SARS-CoV-2; Selection Bias; *Symptom Assessment [classification] [statistics & numerical data]

MeSH check words

Humans