Diagnosis of SARS-CoV-2 infection and COVID-19

Deeks, Jon; Dinnes, Jacqueline; Takwoingi, Yemisi; Davenport, Clare; Leeflang, Mariska M G; Spijker , René ; Hooft, Lotty; Van den Bruel, Ann; Emperador , Devy; Dittrich , Sabine

DOI:
10.1002/14651858.CD013596

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

Citation for published version (Harvard):
Deeks, J, Dinnes, J, Takwoingi, Y, Davenport, C, Leeflang, MMG, Spijker , R, Hooft, L, Van den Bruel, A, Emperador , D & Dittrich , S 2020, '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*, vol. 2020, no. 4, CD013596. https://doi.org/10.1002/14651858.CD013596

Cochrane Library

Cochrane Database of Systematic Reviews

Diagnosis of SARS-CoV-2 infection and COVID-19: accuracy of signs and symptoms; molecular, antigen, and antibody tests; and routine laboratory markers (Protocol)

Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MMG, Spijker R, Hooft L, Van den Bruel A, Emperador D, Dittrich S

Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MMG, Spijker R, Hooft L, Van den Bruel A, Emperador D, Dittrich S. Diagnosis of SARS-CoV-2 infection and COVID-19: accuracy of signs and symptoms; molecular, antigen, and antibody tests; and routine laboratory markers (Protocol).
Cochrane Database of Systematic Reviews 2020, Issue 4. Art. No.: CD013596.
DOI: 10.1002/14651858.CD013596.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Protocol]

Diagnosis of SARS-CoV-2 infection and COVID-19: accuracy of signs and symptoms; molecular, antigen, and antibody tests; and routine laboratory markers

Jonathan J Deeks[1,2], Jacqueline Dinnes[1,2], Yemisi Takwoingi[1,2], Clare Davenport[1,2], Mariska MG Leeflang[3,4], René Spijker[4,5], Lotty Hooft[5], Ann Van den Bruel[6], Devy Emperador[7], Sabine Dittrich[7]

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

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

Editorial group: Cochrane Infectious Diseases Group.
Publication status and date: New, published in Issue 4, 2020.

Citation: Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MMG, Spijker R, Hooft L, Van den Bruel A, Emperador D, Dittrich S. Diagnosis of SARS-CoV-2 infection and COVID-19: accuracy of signs and symptoms; molecular, antigen, and antibody tests; and routine laboratory markers (Protocol). *Cochrane Database of Systematic Reviews* 2020, Issue 4. Art. No.: CD013596. DOI: 10.1002/14651858.CD013596.

ABSTRACT

Objectives

This is a protocol for a Cochrane Review (diagnostic). The objectives are as follows:

- To assess the diagnostic accuracy of laboratory real-time polymerase chain reaction (RT-PCR) and other laboratory molecular tests to determine if a person presenting in the community or in secondary care has SARS-CoV-2 infection.
- To assess the diagnostic accuracy of each rapid PCR and antigen test to determine if a person presenting in the community or in secondary care has SARS-CoV-2 infection.
- To assess the diagnostic accuracy of each antibody test to determine if a person presenting in the community or in secondary care has SARS-CoV-2 infection, or has previously had SARS-CoV-2 infection.
- To assess the diagnostic accuracy of signs and symptoms to determine if a person presenting in the community, general practice, or at the emergency department has SARS-CoV-2 infection, COVID-19 pneumonia, or severe COVID-19 pneumonia/ARDS requiring hospital admission.
- To assess the diagnostic accuracy of routine laboratory testing to determine if a person has COVID-19 pneumonia or SARS-CoV-2 infection.

ZZZ <section xml:id="CD013596-abs2-0001"> ZZZ
ZZZ <title type="main"> ZZZSecondary objectives
Where data are available, for reviews #1 to #5, we will investigate the accuracy (either by stratified analysis or meta-regression) according to:

- laboratory method, days of symptoms, severity of symptoms, reference standard, sample type, study design, setting;
- test brand and version, days of symptoms, severity of symptoms, reference standard, sample type, study design, setting;
- current infection or past infection, test brand and version, days of symptoms or days since symptoms resolved, reference standard, study design, setting;

- days of symptoms, reference standard, study design, setting;
- specific measurement or biomarker, days of symptoms, severity of symptoms, reference standard, sample type, study design, setting.

BACKGROUND

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus and resulting COVID-19 pandemic present important diagnostic evaluation challenges. These range from understanding the value of signs and symptoms in predicting possible infection, assessing whether existing biochemical and imaging tests can identify infection and people needing critical care, and evaluating whether new diagnostic tests can allow accurate rapid and point-of-care testing, either to identify current infection, rule out infection, identify people in need of care escalation, or to test for past infection and immunity.

Target condition being diagnosed

COVID-19 is the disease caused by infection with the SARS-CoV-2. The key target conditions for this suite of reviews are current SARS-CoV-2 infection, current COVID-19, and past SARS-CoV-2 infection.

For current infection, the severity of the disease is of importance. SARS-CoV-2 infection can be asymptomatic (no symptoms); mild or moderate (symptoms such as fever, cough, aches, lethargy but without difficulty breathing at rest); severe (symptoms with breathlessness and increased respiratory rate indicative of pneumonia); or critical (requiring respiratory support due to severe acute respiratory syndrome (SARS) or ARDS). People with COVID-19 pneumonia (severe or critical disease) require different patient management, and it is important to be able to identify them.

Thus, there are two target conditions for current infection:

- SARS-CoV-2 infection (asymptomatic or symptomatic of any severity);
- COVID-19 pneumonia (severe or critical).

In planning review updates, we will consider the potential addition of two further groupings (which are subsets of the above):

- whether tests exist that identify people requiring respiratory support (SARS or ARDS);
- identification of asymptomatic SARS-CoV-2 infection.

Tests for past SARS-CoV-2 infection will be used to assess whether an individual is likely to be immune.

Index test(s)

Molecular and antigen tests

Testing for presence of the SARS-CoV-2 virus has been undertaken using quantitative reverse transcription polymerase chain reaction (qRT-PCR). Polymerase chain reaction (PCR) tests for SARS-CoV-2 identify viral ribonucleic acid (RNA). Reagents for the assay were very quick to produce once the viral RNA sequence was published. Testing is undertaken in central laboratories and is very labour intensive, with several points along the path of performing a single test where errors may occur, although some automation of parts of the process is possible. Test results are typically available in 24 to 48 hours, although faster processes are being implemented. Other nucleic amplification methods such as loop-mediated isothermal amplification (LAMP), or CRISPR-based nucleic acid detection methods are also being developed with the potential to produce results within minutes as opposed to hours. PCR tests use upper and lower respiratory samples. Sputum is currently considered better than oropharynx swabs or nasopharynx swabs but more difficult (and hazardous) to obtain.

Point-of-care PCR devices can also be used for identification of infection. These are portable analyzer devices (which can be based in the clinic) and matching test cartridges. Several companies have suitable existing portable technology systems and are producing the required new cartridges for diagnosis of SARS-CoV-2 infection. Test results are based on the same samples as for qRT-PCR, with results available within minutes or hours. Use of these tests requires investment in the technology platform, which will be in place in some settings as they are also used to diagnose other infectious diseases such as tuberculosis.

Point-of-care testing can also be undertaken using disposable lateral flow assays, akin to a pregnancy test, which indicate by colour change whether a SARS-CoV-2 antigen has been detected in a swab sample or sample of bodily fluids. These tests are portable and do not require laboratory facilities or technology platforms.

Antibody tests

Testing for antibody response to infection is typically based on measuring immunoglobulin (Ig)M and IgG in serology (IgA and total antibodies may also be measured). Most tests assess both IgG and IgM. IgM typically rises quickly with infection and declines soon after an infection is cleared. IgG is thought to rise more slowly but may persist and reflect longer term immunity. Antibody tests are likely to be available in laboratory form using enzyme-linked immunosorbent assay (ELISA) methods, but also as point-of-care tests using a disposable device using one or two spots of blood from a thumb prick on a testing strip, and take around 10 minutes for a positive answer. There is no infection risk to sampling over and above that of a finger prick.

Signs and symptoms

Signs and symptoms are used in the initial diagnosis of suspected COVID-19 infection, and in identifying people with COVID-19 pneumonia. Key symptoms that have been associated with mild-to-moderate COVID-19 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, breathless on light exertion, muscle pain, fatigue, and loss of sense of smell and taste. Red flags indicating possible pneumonia include: breathlessness at rest, increased respiratory rate (above 20 breaths/minute), increased heart rate (above 100 beats per minute), chest tightness, loss of appetite, confusion, pain or pressure in the chest, blue lips or face, and temperature above 38°C. Hypoxia based on measuring pulse oximetry is often used with various arbitrary thresholds (for example, 93%).

Routinely available biomarkers

Routinely available biomarkers for infection and inflammation may be considered in the investigation of people with possible COVID-19 infection. For example, many healthcare facilities have access to standard laboratory tests for infection such as C-reactive protein (CRP), procalcitonin, measures of anticoagulation, and white blood cell count with differential. Evaluation of these commonly available tests, particularly in low-resource settings, may be helpful for the triage of people with potential COVID-19.

Clinical pathway

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 decisions.

Prior test(s)

Prior testing will depend on whether people are being investigated for SARS-CoV-2 infection or COVID-19 pneumonia, as well as which index tests are being evaluated in a particular review. For example, in the review on signs and symptoms, there are no prior tests because signs and symptoms are used in the initial diagnosis of suspected COVID-19 infection, and in identifying people with COVID-19 pneumonia.

Role of index test(s)

This protocol addresses several index tests that may be used for different purposes, in different settings, and at different points of the diagnostic pathway. The potential role of the index tests as replacement, add-on or triage tests will be considered in each review.

Alternative test(s)

Chest X-ray, ultrasound, and computed tomography (CT) are widely used diagnostic imaging tests to identify COVID-19 pneumonia. Availability and usage varies between settings. These will be addressed in a separate protocol.

Rationale

It is essential to understand the accuracy of tests and diagnostic features 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 summarize evidence on the clinical accuracy of different tests and diagnostic features, grouped according to the research questions and settings that we are aware of. 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 current evidence about test accuracy.

This is a generic protocol for Cochrane Diagnostic Test Accuracy (DTA) Reviews of SARS-CoV-2 (the infection) and COVID-19 (the symptomatic disease). This version of the protocol covers five review titles.

1. Laboratory-based molecular tests for diagnosis of SARS-CoV-2 infection.
2. Rapid point-of-care tests for diagnosis of SARS-CoV-2 infection.
3. Antibody tests for identification of current and past infection with SARS-CoV-2.
4. Signs and symptoms to determine if a patient presenting in general practice or at the emergency department has COVID-19, COVID-19 pneumonia, or severe COVID-19 pneumonia/acute respiratory distress syndrome (ARDS) requiring intensive care unit (ICU) admission.
5. Routine laboratory testing to determine if a patient has COVID-19 pneumonia or SARS-CoV-2 infection.

We anticipate a sixth review on CT scanning and other diagnostic imaging to be added in due course.

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

OBJECTIVES

- To assess the diagnostic accuracy of laboratory real-time polymerase chain reaction (RT-PCR) and other laboratory molecular tests to determine if a person presenting in the community or in secondary care has SARS-CoV-2 infection.
- To assess the diagnostic accuracy of each rapid PCR and antigen test to determine if a person presenting in the community or in secondary care has SARS-CoV-2 infection.
- To assess the diagnostic accuracy of each antibody test to determine if a person presenting in the community or in secondary care has SARS-CoV-2 infection, or has previously had SARS-CoV-2 infection.
- To assess the diagnostic accuracy of signs and symptoms to determine if a person presenting in the community, general practice, or at the emergency department has SARS-CoV-2 infection, COVID-19 pneumonia, or severe COVID-19 pneumonia/ARDS requiring hospital admission.
- To assess the diagnostic accuracy of routine laboratory testing to determine if a person has COVID-19 pneumonia or SARS-CoV-2 infection.

Secondary objectives

Where data are available, for reviews #1 to #5, we will investigate the accuracy (either by stratified analysis or meta-regression) according to:

- laboratory method, days of symptoms, severity of symptoms, reference standard, sample type, study design, setting;
- test brand and version, days of symptoms, severity of symptoms, reference standard, sample type, study design, setting;
- current infection or past infection, test brand and version, days of symptoms or days since symptoms resolved, reference standard, study design, setting;
- days of symptoms, reference standard, study design, setting;
- specific measurement or biomarker, days of symptoms, severity of symptoms, reference standard, sample type, study design, setting.

METHODS

Criteria for considering studies for this review

Types of studies

For all questions, we will keep the eligibility criteria broad to include all patient groups and all variations of a test (that is, if patient population is unclear, we will include the study).

We will include studies of all designs that produce estimates of test accuracy or provide data from which estimates can be computed.

Unlike standard DTA reviews, in initial versions of these rapid reviews we will include studies that include only participants confirmed to have the target condition (to estimate sensitivity) or not to have the target condition (to estimate specificity). We will present and analyse separately the findings from such studies given that these test characteristic estimates are unlikely to reflect diagnostic test performance in clinical practice.

We will include both single-gate (studies which 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.

We will include studies based on patients, samples, and spiked samples.

When interpreting the results, we will make sure that the limitations of different study designs are carefully considered, using quality assessment and analysis. We will ensure that the interpretation is in line with the strength of the evidence and will explicitly report when the evidence is only available from weak study designs.

Participants

Studies recruiting people presenting with suspicion of SARS-CoV-2 infection are eligible for all reviews. In addition, review #3 will include studies in people known to have previously had SARS-CoV-2 infection to evaluate tests for past infection.

Reviews #1, #2, #3, and #5 will also include studies that recruit populations where tests are being used to screen for disease (for example, contact tracing or community screening).

For the initial version of these reviews, we will include studies that recruited people either known to have SARS-CoV-2 infection or known not to have SARS-CoV-2 infection.

We will include studies based on serum banks created from known cases of COVID-19 and controls.

Laboratory studies based on spiked and clean sample sets will be included in early versions of this review as we anticipate that for many tests this may be the only evidence available.

Studies must include a minimum of 10 samples or 10 participants.

Index tests

We will create a list of eligible index tests that have obtained regulatory approval for inclusion in reviews #1, #2, and #3. As the situation is constantly changing, we have not listed the eligible tests here, but rather the sources we will use to identify tests that have regulatory approval on a stated date.

We will search research organization and regulatory websites to identify tests with regulatory approval. The two main resources which we will use are:

- **World Health Organization (WHO)** COVID-19 listing in International Medical Device Regulators Forum (IMDRF) jurisdictions (www.who.int/diagnostics_laboratory/EUL/en/), which includes listings of US Food and Drug Administration (FDA), Health Canada, Japan, Australia (Therapeutic Goods Administration), Singapore (Health Sciences Authority), Brazil (Agência Nacional de Vigilância Sanitária), South Korea (Ministry of Food and Drug Safety), China (National Medical Products Administration), and Russia (Roszdravnadzor);
- **FIND** SARS-COV-2 Diagnostic pipeline (www.finddx.org/covid-19/pipeline/), which overlaps with the WHO list, but in addition includes CE-IVD and IVD India.

The list has been created by review of websites and contact with commercial test manufacturers.

In addition, we will check against key national websites, including US FDA (www.fda.gov/medical-devices/emergency-situations-medical-devices/emergency-use-authorizations#coronavirus2019); China FDA (subsites.chinadaily.com.cn/nmpa/2020 03/27/c_465663.htm?bsh_bid=5496527208); National Institute for Health Research (NIHR) Innovation Observatory (www.io.nihr.ac.uk/) (and other resources from organization, for example, www.rapidmicrobiology.com/test-method/testing-for-the-wuhan-coronavirus-a-k-a-covid-19-sars-cov-2-and-2019-ncov).

For review #4 on signs (including pulse-oximetry levels) and symptoms, we will collate evidence on all signs and symptoms reported in the identified studies. We will include combinations of signs and symptoms, but not when they are combined with laboratory, imaging, or other types of index tests.

For review #5 on routine biomarkers, we will collate evidence on all routine biomarker tests reported in the identified studies.

Target conditions

To be eligible studies will need to identify at least one of:

- current SARS-CoV-2 infection;
- COVID-19 pneumonia;
- past SARS-CoV-2 infection.

Reference standards

We anticipate that studies will use a range of reference standards within and across the reviews. 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 expect 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 neutralization test (PRNT) or ELISA tests;
- information available at a subsequent time point;
- WHO and other case definitions.

This list is not exhaustive, and we will record all reference standards encountered. We will recruit methodological and clinical experts to produce 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.

Search methods for identification of studies

Electronic searches

Electronic searches will be undertaken from two different sources. Both of these searches will aim to identify all articles related to COVID-19, and will not be restricted to those evaluating biomarkers or tests. Thus, there are no test terms, diagnosis terms, or methodological terms in the searches. Searches will be limited to 2019 and 2020.

Living search from the University of Bern

We will use the COVID-19 living search results of the Institute of Social and Preventive Medicine (ISPM) at the University of Bern (www.ispm.unibe.ch). 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 1.

Cochrane COVID-19 register searches

We will also include searches undertaken by Cochrane of MEDLINE, which are being run to develop the Cochrane COVID-19 register (https://covid-19.cochrane.org/). These include searches of trials registers at CT.GOV and the WHO International Clinical Trials Registry Platform (ICTRP), as well as PubMed.

Search strategies are designed for maximum sensitivity, to retrieve all human studies on COVID-19. There are no language limits. Strategies may be revised to account for changes to the COVID-19 study register's eligibility criteria, changes to database interfaces, and search performance assessments. See Appendix 2.

Searching other resources

We will check repositories of COVID-19 publications against these search results including:

- the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre; eppi.ioe.ac.uk/COVID19_MAP/covid_map_v4.html);
- Meta-evidence (meta-evidence.co.uk/the-role-of-evidence-synthesis-in-covid19/).

From these websites, we will search company and product websites for studies about test accuracy.

We will contact companies to request further information about studies.

We will contact research groups that are completing test evaluations (for example, UK Public Health England-funded studies, FIND studies (www.finddx.org/). We will make appeals in each version of the review for researchers to supply details of unpublished studies.

We will impose no language restrictions.

Data collection and analysis

Selection of studies

Two review authors will independently screen studies. We will resolve disagreements by discussion with a third experienced review author for initial titles and abstract screening, and through discussion between three review authors for eligibility assessments.

Data extraction and management

Two review authors will independently perform data extraction. We will resolve disagreements by discussion between three review authors.

We will contact study authors where we need to check details and obtain missing information.

We will encourage study authors to submit unpublished studies in the FIND register or other online resources.

Assessment of methodological quality

Two review authors will independently assess risk of bias and applicability concerns using the QUADAS-2 checklist (Table 1). We will resolve disagreements by discussion between three review authors.

QUADAS-2 checklist

This is a generic protocol, addressing different tests, for different situations, settings, and populations. Therefore, the QUADAS-2 operationalization will be split into the five main categories: laboratory-based molecular tests, point-of-care tests, antibody tests, signs and symptoms, and routinely available laboratory tests.

Statistical analysis and data synthesis

We will present results of estimated sensitivity and specificity using paired forest plots in Review Manager 2014, and summarized in tables as appropriate.

For reviews #1, #2, and #3, we anticipate that most of the results will be presented without meta-analysis, due to the small numbers of studies currently available on each of the new molecular, antigen, and antibody tests. As there are currently different approaches being used by different manufacturers, we anticipate that there will be differences in findings for different tests and we will report data without aggregation to make this clear. We will also disaggregate data by study design, reporting results from artificial laboratory samples, sample sets without patient data, and samples obtained from patients in three different groups.

Where pooling is possible, we will estimate mean sensitivity and specificity using hierarchical models where tests either report binary results (expected for all COVID-19 tests in reviews #1, #2, and #3; symptoms and signs in review #4) or at commonly reported thresholds for routinely available tests in review #5. 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. Meta-analysis will be undertaken 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 2010).

Investigations of heterogeneity

Sources of heterogeneity that will be investigated if adequate data are available are listed in the Secondary objectives, either using stratification (where we believe it is inappropriate to combine studies) or through meta-regression models.

For reviews #1, #2, and #3 on COVID-19 tests, we will stratify analyses by test brand and study methods. We will also consider stratification by reference standard and study sample type dependent on the sampling and verification approaches that have been identified. For review #3, we will undertake separate analyses based on whether the recruited sample has current symptoms (that is, test is detecting current active disease) or known previous disease (having no symptoms for seven days).

We will stratify reviews #4 and #5 by reference standard and study design.

Sensitivity analyses

We will undertake sensitivity analyses considering the impact of:

- unpublished studies;
- studies identified only from industry Instructions for Use (IFU) documentation;
- studies using sample banks or spiked samples;
- studies with inadequate reference standards;
- for previous infection (review #3), increasing lengths of time since symptoms cleared.

Assessment of reporting bias

We will 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.

Summary of findings

We will aim to list key findings in 'Summary of findings' tables 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, preprints, and new test approvals weekly, and, dependent on the number of new and important studies found, we will consider updating each review with each search if resources allow.

ACKNOWLEDGEMENTS

The editorial base of the Cochrane Infectious Diseases Group 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.

The authors thank Prof Danielle Van der Windt who was the Cochrane Diagnostic Test Accuracy (DTA) Contact Editor for this review; the clinical and methodological referees; the Cochrane DTA Editorial Team; and Anne Lawson who copy-edited the protocol.

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

REFERENCES

Additional references

Macaskill 2010

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, 2010. Available from https://methods.cochrane.org/sdt/handbook-dta-reviews.

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.

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.

ADDITIONAL TABLES

Table 1. QUADAS-2 checklist

Index test(s)	Review #1. Laboratory based molecular tests	Review #2. Point-of-care tests	Review #3. Antibody tests	Review #4. Signs and symptoms	Review #5. Routine laboratory tests
Patients (setting, intended use of index test, presentation, prior testing)	Considered to be the "gold standard" for acute infection. May have been used with different samples, in different settings, for case-finding or confirmation of infection in patients with suspected COVID-19.	In patients with suspected COVID-19 or contact tracing. Point-of-care: case-finding in the general population, care homes for elderly people, emergency departments.	In patients with signs and symptoms suspected of COVID-19 and for case finding; also in patients with past exposure to SARS-CoV-2.	General practice, primary care, emergency care. In patients presenting with suspected COVID-19. No prior testing. Signs and symptoms often used for triage or referral.	Mainly meant for situations where a laboratory was close; emergency care, hospital, ICU. COVID triage centres. In patients presenting with suspected COVID-19.
Reference standard and target condition	The focus will be on the diagnosis of COVID-19 pneumonia or infection with SARS-CoV-2. For this protocol, 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.				

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

	UNCLEAR: if the selection procedure was not clear or not reported.				
Did the study avoid inappropriate exclusions?	Studies may have excluded patients, or selected patients in such a way that they avoided including those who were difficult to diagnosis 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-to-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 postfactum or comorbidities (cardiovascular disease, diabetes, immunosuppression). UNCLEAR: if the exclusion criteria were not reported.				
Did the study avoid inappropriate inclusions?	Some laboratory studies may have intentionally included groups of patients in whom the accuracy was likely to differ, such as those with particularly low or high viral loads, or who had other diseases, such that the sample over-represented these groups. This needs to be addressed on a case-to-case basis. Artificial spiked samples are a clear example. 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 RT-PCR was assessed in a case-control design; to screen contacts or for stopping contact isolation. Studies done in sample banks and spiked samples. LOW: any other situation: these tests may be used in different settings and for different purposes. UNCLEAR: if a description about the participants was lacking.	HIGH: if accuracy of tests was assessed in a case-control design; if not used to diagnose early acute infection; to screen contacts or for stopping contact isolation. Studies done in sample banks and spiked samples. LOW: any other situation: these tests may have been used in different settings and for different purposes. UNCLEAR: if a description about the participants was lacking.	HIGH: if accuracy of tests was assessed in a case-control design; when patients were tested too early in the disease phase for detection of past infection. Studies done in sample banks and spiked samples. LOW: any other situation: these tests may be used in different settings and for different purposes. UNCLEAR: if a description about the participants was lacking.	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.	HIGH: if accuracy of laboratory tests was assessed in a case-control design, or in an already highly selected group of participants. LOW: any situation where generic laboratory tests were among the first tests to be done on the included participants. UNCLEAR: if a description about the participants was lacking.
INDEX TESTS					

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

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 (in review # 4, Signs and symptoms). UNCLEAR: if threshold selection was not clearly reported.				
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?	HIGH: if tests were built in-house. If tests were undertaken in a different setting, or using samples, equipment, or personnel not available in practice.	HIGH: if tests were built in-house. If tests were undertaken in a different setting, or using samples, equipment or personnel not available in practice.	HIGH: if tests were built in-house. If tests were undertaken in a different setting, or using samples, equipment. or personnel not available in practice.	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.	This will probably be answered 'LOW' in all cases, except when tests used a threshold that was much higher or lower than in practice, or undertaken in a different setting, or using samples, equipment, or 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 knowl-	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.				

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

edge of the results of the index test?	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. For testing for previous infection, a time interval of at least two weeks is required since resolution of symptoms before the index test was undertaken. NO: if there was more than 24 hours between the index test and the reference standard or if patients were otherwise reported to be assessed with the index versus reference standard test at moments of different severity. UNCLEAR: if the time interval was not reported.
Did all patients receive a reference standard?	YES: if all patients received a reference standard (clearly no partial verification). NO: if only (part of) the index test positives or index test negatives received the complete reference standard. UNCLEAR: if it was not reported.
Did all patients receive the same reference standard?	YES: if all patients 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.

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

Were all patients included in the analysis?	YES: if all included patients were included in the analyses as well. NO: if after the inclusion/exclusion process, patients 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: real-time polymerase chain reaction; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2; WHO: World Health Organization.

APPENDICES

Appendix 1. Living search from the University of Bern

The following information is taken from the university of Bern website (see: https://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 retrieve 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 2. Cochrane COVID-19 register searches

Source	Strategy
CT.gov	COVID-19[a]
WHO ICTRP	Health topic: 2019-nCov / COVID-19

(Continued)

PubMed	(("2019 nCoV"[tiab] OR 2019nCoV[tiab] OR "2019 novel coronavirus"[tiab] OR "COVID 19"[tiab] OR COVID19[tiab] OR "new coronavirus"[tiab] OR "novel coronavirus"[tiab] OR "novel coro-na virus"[tiab] OR "SARS CoV-2"[tiab] OR (Wuhan[tiab] AND (coronavirus[tiab] OR "corona virus"[tiab])) OR "COVID-19"[Supplementary Concept] OR "severe acute respiratory syndrome coronavirus 2"[Supplementary Concept]) NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])) NOT (editorial[pt] OR comment[pt] OR letter[pt] OR newspaper article[pt])
Embase[b]	(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-

[a]Automatic term mapping links results for 2019-nCoV, 2019 novel coronavirus, SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.
[b]The published Cochrane COVID-19 register currently does not display Embase records. However, Embase records obtained through Martha Knuth for the Centers for Disease Control and Prevention (CDC), Stephen B Thacker CDC Library, COVID-19 Research Articles Downloadable Database (www.cdc.gov/library/researchguides/2019novelcoronavirus/researcharticles.html) were included in the version we received from the Cochrane Central Executive Office. Embase records were deduplicated against PubMed records by Robin Featherstone from the Cochrane Central Executive Office.

Appendix 3. 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 < two months: ≥ 60; aged two to 11 months: ≥ 50; aged one to five years: ≥ 40. While the diagnosis is made on clinical grounds; chest imaging may identify or exclude some pulmonary complications.

ARDS: onset within one week of a known clinical insult or new or worsening respiratory symptoms.

Chest imaging (that is, X-ray, computer 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 $SpO_2 \leq 97\%$ to calculate OSI or SpO_2/FiO_2 ratio:

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

HISTORY

Protocol first published: Issue 4, 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.

DECLARATIONS OF INTEREST

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.

Ann Van den Bruel: 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. 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.

SOURCES OF SUPPORT

Internal sources

- Liverpool School of Tropical Medicine, UK

External sources

- Department for International Development, UK

  Project number: 300342-104
- National Institute for Health Research (NIHR), UK