Cochrane Library

Cochrane Database of Systematic Reviews

Universal screening for SARS-CoV-2 infection: a rapid review (Review)

Viswanathan M, Kahwati L, Jahn B, Giger K, Dobrescu AI, Hill C, Klerings I, Meixner J, Persad E, Teufer B, Gartlehner G

Viswanathan M, Kahwati L, Jahn B, Giger K, Dobrescu AI, Hill C, Klerings I, Meixner J, Persad E, Teufer B, Gartlehner G.
Universal screening for SARS-CoV-2 infection: a rapid review.
Cochrane Database of Systematic Reviews 2020, Issue 9. Art. No.: CD013718.
DOI: 10.1002/14651858.CD013718.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Rapid Review]

Universal screening for SARS-CoV-2 infection: a rapid review

Meera Viswanathan[1], Leila Kahwati[1], Beate Jahn[2], Kayla Giger[3], Andreea Iulia Dobrescu[4], Christine Hill[1], Irma Klerings[4], Jana Meixner[4], Emma Persad[4], Birgit Teufer[4], Gerald Gartlehner[4]

[1]RTI International, Research Triangle Park, North Carolina, USA. [2]Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria. [3]Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [4]Cochrane Austria, Department for Evidence-based Medicine and Evaluation, Danube University Krems, Krems, Austria

Contact address: Meera Viswanathan, viswanathan@rti.org.

Editorial group: Cochrane Public Health Group.
Publication status and date: New, published in Issue 9, 2020.

Citation: Viswanathan M, Kahwati L, Jahn B, Giger K, Dobrescu AI, Hill C, Klerings I, Meixner J, Persad E, Teufer B, Gartlehner G. Universal screening for SARS-CoV-2 infection: a rapid review. *Cochrane Database of Systematic Reviews* 2020, Issue 9. Art. No.: CD013718. DOI: 10.1002/14651858.CD013718.

ABSTRACT

Background

Coronavirus disease 2019 (COVID-19) is caused by the novel betacoronavirus, severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). Most people infected with SARS-CoV-2 have mild disease with unspecific symptoms, but about 5% become critically ill with respiratory failure, septic shock and multiple organ failure. An unknown proportion of infected individuals never experience COVID-19 symptoms although they are infectious, that is, they remain asymptomatic. Those who develop the disease, go through a presymptomatic period during which they are infectious. Universal screening for SARS-CoV-2 infections to detect individuals who are infected before they present clinically, could therefore be an important measure to contain the spread of the disease.

Objectives

We conducted a rapid review to assess (1) the effectiveness of universal screening for SARS-CoV-2 infection compared with no screening and (2) the accuracy of universal screening in people who have not presented to clinical care for symptoms of COVID-19.

Search methods

An information specialist searched Ovid MEDLINE and the Centers for Disease Control (CDC) COVID-19 Research Articles Downloadable Database up to 26 May 2020. We searched Embase.com, the CENTRAL, and the Cochrane Covid-19 Study Register on 14 April 2020. We searched LitCovid to 4 April 2020. The World Health Organization (WHO) provided records from daily searches in Chinese databases and in PubMed up to 15 April 2020. We also searched three model repositories (Covid-Analytics, Models of Infectious Disease Agent Study [MIDAS], and Society for Medical Decision Making) on 8 April 2020.

Selection criteria

Trials, observational studies, or mathematical modelling studies assessing screening effectiveness or screening accuracy among general populations in which the prevalence of SARS-CoV2 is unknown.

Data collection and analysis

After pilot testing review forms, one review author screened titles and abstracts. Two review authors independently screened the full text of studies and resolved any disagreements by discussion with a third review author. Abstracts excluded by a first review author were dually reviewed by a second review author prior to exclusion. One review author independently extracted data, which was checked by a second review author for completeness and accuracy. Two review authors independently rated the quality of included studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for diagnostic accuracy studies and a modified form designed originally for

economic evaluations for modelling studies. We resolved differences by consensus. We synthesized the evidence in narrative and tabular formats. We rated the certainty of evidence for days to outbreak, transmission, cases missed and detected, diagnostic accuracy (i.e. true positives, false positives, true negatives, false negatives) using the GRADE approach.

Main results

We included 22 publications. Two modelling studies reported on effectiveness of universal screening. Twenty studies (17 cohort studies and 3 modelling studies) reported on screening test accuracy.

Effectiveness of screening

We included two modelling studies. One study suggests that symptom screening at travel hubs, such as airports, may slightly slow but not stop the importation of infected cases (assuming 10 or 100 infected travellers per week reduced the delay in a local outbreak to 8 days or 1 day, respectively). We assessed risk of bias as minor or no concerns, and certainty of evidence was low, downgraded for very serious indirectness. The second modelling study provides very low-certainty evidence that screening of healthcare workers in emergency departments using laboratory tests may reduce transmission to patients and other healthcare workers (assuming a transmission constant of 1.2 new infections per 10,000 people, weekly screening reduced infections by 5.1% within 30 days). The certainty of evidence was very low, downgraded for high risk of bias (major concerns) and indirectness. No modelling studies reported on harms of screening.

Screening test accuracy

All 17 cohort studies compared an index screening strategy to a reference reverse transcriptase polymerase chain reaction (RT-PCR) test. All but one study reported on the accuracy of single point-in-time screening and varied widely in prevalence of SARS-CoV-2, settings, and methods of measurement.

We assessed the overall risk of bias as unclear in 16 out of 17 studies, mainly due to limited information on the index test and reference standard. We rated one study as being at high risk of bias due to the inclusion of two separate populations with likely different prevalences. For several screening strategies, the estimates of sensitivity came from small samples.

For single point-in-time strategies, for symptom assessment, the sensitivity from 12 cohorts (524 people) ranged from 0.00 to 0.60 (very low-certainty evidence) and the specificity from 12 cohorts (16,165 people) ranged from 0.66 to 1.00 (low-certainty evidence). For screening using direct temperature measurement (3 cohorts, 822 people), international travel history (2 cohorts, 13,080 people), or exposure to known infected people (3 cohorts, 13,205 people) or suspected infected people (2 cohorts, 954 people), sensitivity ranged from 0.00 to 0.23 (very low- to low-certainty evidence) and specificity ranged from 0.90 to 1.00 (low- to moderate-certainty evidence). For symptom assessment plus direct temperature measurement (2 cohorts, 779 people), sensitivity ranged from 0.12 to 0.69 (very low-certainty evidence) and specificity from 0.90 to 1.00 (low-certainty evidence). For rapid PCR test (1 cohort, 21 people), sensitivity was 0.80 (95% confidence interval (CI) 0.44 to 0.96; very low-certainty evidence) and specificity was 0.73 (95% CI 0.39 to 0.94; very low-certainty evidence). One cohort (76 people) reported on repeated screening with symptom assessment and demonstrates a sensitivity of 0.44 (95% CI 0.29 to 0.59; very low-certainty evidence) and specificity of 0.62 (95% CI 0.42 to 0.79; low-certainty evidence).

Three modelling studies evaluated the accuracy of screening at airports. The main outcomes measured were cases missed or detected by entry or exit screening, or both, at airports. One study suggests very low sensitivity at 0.30 (95% CI 0.1 to 0.53), missing 70% of infected travellers. Another study described an unrealistic scenario to achieve a 90% detection rate, requiring 0% asymptomatic infections. The final study provides very uncertain evidence due to low methodological quality.

Authors' conclusions

The evidence base for the effectiveness of screening comes from two mathematical modelling studies and is limited by their assumptions. Low-certainty evidence suggests that screening at travel hubs may slightly slow the importation of infected cases. This review highlights the uncertainty and variation in accuracy of screening strategies. A high proportion of infected individuals may be missed and go on to infect others, and some healthy individuals may be falsely identified as positive, requiring confirmatory testing and potentially leading to the unnecessary isolation of these individuals. Further studies need to evaluate the utility of rapid laboratory tests, combined screening, and repeated screening. More research is also needed on reference standards with greater accuracy than RT-PCR.

Given the poor sensitivity of existing approaches, our findings point to the need for greater emphasis on other ways that may prevent transmission such as face coverings, physical distancing, quarantine, and adequate personal protective equipment for frontline workers.

PLAIN LANGUAGE SUMMARY

How effective is screening for COVID-19?

Why does screening matter?

Screening aims to identify a condition in people who may not be showing any symptoms. Some people may have the COVID-19 virus but appear healthy or have only mild symptoms. It is important to identify infected people so they can stay away from others and seek

appropriate care. Incorrectly identifying COVID-19 in healthy people could lead to unnecessary self-isolation and further tests. Incorrectly identifying no infection in infected people could spread the virus.

Screening for COVID-19 can include temperature checks, or asking about international travel or contact with COVID-19 cases, or rapid tests. Screening can occur over the telephone, online, or in person, in homes, clinics, workplaces, airports or schools.

What did the review study?

We wanted to identify:

· the benefits and negative effects of screening apparently healthy people for COVID-19 infection;

· whether screening can identify those with and without the virus correctly.

To answer these questions rapidly, we shortened some steps of the normal Cochrane Review process. We are confident these changes do not affect our overall conclusions.

What did we do?

We looked for studies that screened people who had not sought care for potential COVID-19 symptoms.

This review includes evidence up to May 2020.

Key results

We found 22 studies; 17 assessed people (cohort studies) and five were computer-generated models (modelling studies). Studies took place in USA, Europe, and Asia.

Benefits and negative effects
Two modelling studies reported on the benefits and negative effects of screening. One suggested that asking about symptoms at airports may slightly slow, but not stop, the importation of infected people.

Another model reported that weekly or biweekly screening of healthcare workers may reduce transmission to patients and other healthcare workers in emergency departments.

No studies reported on negative effects of screening.

Identification of infected people
Seventeen cohort studies and three modelling studies reported on whether screening can correctly identify those with and without the virus. Studies varied widely in the baseline level of COVID-19, settings, and methods. All cohort studies compared screening strategies to a 'gold standard' test called RT-PCR.

Cohort studies
All screening strategies (17 studies, 17,574 people), incorrectly identified:

· between 20 and 100 out of 100 infected people as healthy;

· between 0 and 38 people out of 100 healthy people as infected

Asking about symptoms (13 studies, 16,762 people), incorrectly identified:

· between 40 to 100 out of 100 infected people as healthy

· between 0 to 34 out of 100 healthy people as infected

Temperature measurements, *asking about international travel*, *exposure to known infected people* and *exposure to known or suspected infected people* (6 studies, 14,741 people), incorrectly identified**:**

· between 77 and 100 out of 100 infected people as healthy

· between 0 and 10 out of 100 healthy people as infected

Asking about symptoms plus temperature measurement (2 studies, 779 people), incorrectly identified:

· between 31 and 88 out of 100 infected people as healthy

· between 0 to 10 people out of 100 healthy people as infected

There was insufficient evidence from two small studies on rapid laboratory tests and repeated symptom assessment to tell how accurate they were in identifying healthy and infected people.

Modelling studies
Three studies modelled entry and exit screening in airports. One study missed 70% of infected travellers. Another detected 90% of infections, but used an unrealistic scenario. The third used very unreliable methods so we cannot use evidence from this study.

How confident are we in the results of the studies?

Our confidence in these findings is limited because most studies did not describe their screening methods clearly, some found very few cases of infections and the types of participants and settings varied greatly, making it difficult to judge whether the results apply broadly.

Authors' conclusions

One-time screening in apparently healthy people is likely to miss people who are infected. We are unsure whether combined screenings, repeated symptom assessment, or rapid laboratory tests are useful.

As more people become infected, screening will identify more cases. However, because screening can miss people who are infected, public health measures such as face coverings, physical distancing, and quarantine for those who are apparently healthy, continue to be very important.

BACKGROUND

Description of the condition

Coronavirus disease 2019 (COVID-19) is of zoonotic origin and caused by the novel betacoronavirus, severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2; WHO 2020a). The virus is genetically similar to the coronaviruses that caused severe acute respiratory syndrome (SARS) and the Middle East respiratory syndrome (MERS), but SARS-CoV-2 appears to have greater transmissibility and lower pathogenicity than SARS and MERS (WHO 2020a). The virus spreads primarily through droplets of saliva or discharge from the nose during close unprotected contact between an infector and infectee (WHO 2020b). On 30 January 2020, the World Health Organization (WHO) declared the outbreak a global health emergency, and on 11 March 2020, WHO declared a pandemic (WHO 2020c).

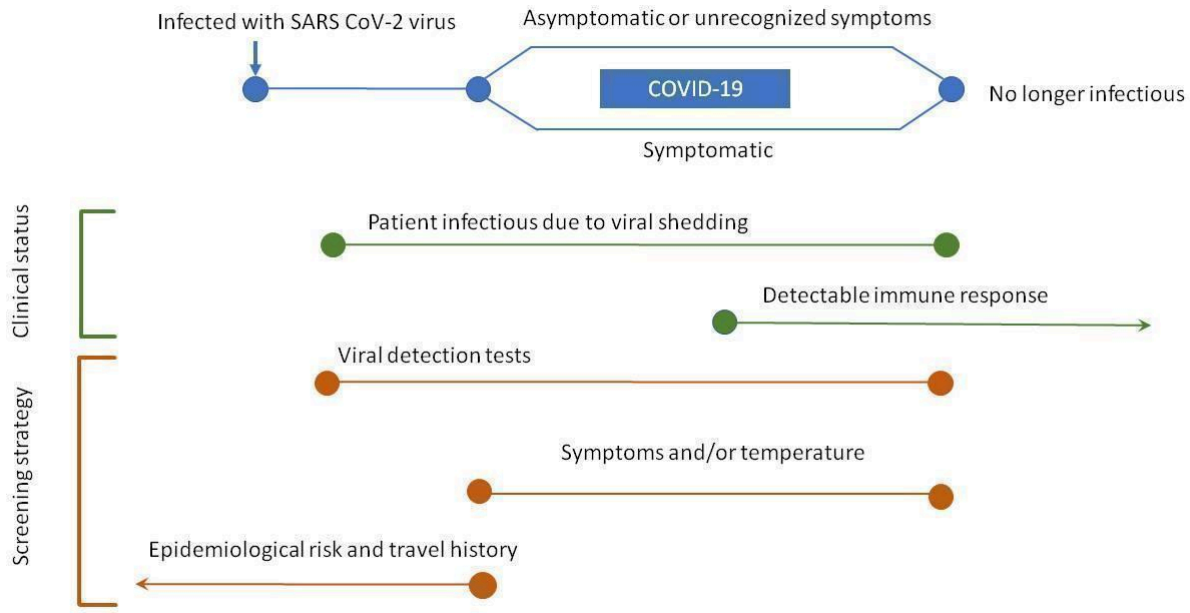
Most people infected with SARS-CoV-2 have mild disease with unspecific symptoms (Wu 2020). It is thought that about 5% become critically ill with respiratory failure, septic shock, multiple organ failure, or a combination of these (Wu 2020). Estimates ranging from 6% to 41% of infected individuals, however, have an asymptomatic course of disease and never develop symptoms. These individuals may carry similar viral loads to symptomatic individuals (Lee 2020). Those who develop the disease may go through a presymptomatic period during which they are already infectious because of viral shedding in the upper respiratory tract (Wolfel 2020).

Universal screening for SARS-CoV-2 infections, (i.e. trying to detect individuals who are infectious before they present clinically or who remain asymptomatic), could therefore be an important measure to contain the spread of the disease. In the city of Wuhan, China, for example, the government initiated door-to-door symptom screening of residents with support from thousands of community workers (Pan 2020). Other forms of commonly used screening measures are body temperature checks, asking individuals about international travel, or inquiring about exposure to known or suspected COVID-19 cases. In fact, screening using temperature checks is already being used routinely to admit people into community or workplace settings (CDC 2020a; Cripps 2020; Emirates 2020).

Description of the screening test

Screening is a strategy to detect an unrecognized disease or symptoms in individuals who are asymptomatic or may not relate specific symptoms to a target disease. Screening is not diagnostic and usually requires further tests to confirm or rule out a disease. In the case of screening for SARS-CoV-2 infections, the goal of screening is to detect infected individuals to reduce onward transmission. Figure 1 depicts the role of screening in identifying individuals with SARS-CoV-2.

Figure 1. Role of screening in identifying people with SARS-CoV-2



Screening strategies may include clinical symptom assessment (e.g. asking people about the presence of symptoms such as fever, dry cough, shortness of breath, gastrointestinal symptoms, loss of smell or taste), epidemiological risk assessment (history of travel from hot spots, contact with people known or suspected of having SARS-CoV-2 infection, occupational risk), body temperature checks, feasible point-of-care laboratory testing (e.g. rapid reverse transcription polymerase chain reaction (RT-PCR)), or a combination of these approaches.

The reference standards for evaluating the accuracy of screening strategies may include viral detection tests, such as standard RT-PCR and clustered regularly interspaced short palindromic repeats

(CRISPR), or a combination of these approaches. RT-PCR tests require nasopharyngeal or oropharyngeal swabs; proper specimen collection is crucial to accurate results (CDC 2020b).

The setting of screening may be over the telephone, online, or in person. In-person screening may occur in a variety of settings including healthcare clinics, emergency rooms, pharmacies, workplaces, arrival or departure screening at travel hubs such as airports and train and bus stations, roadside checks, schools, religious sites, neighbourhood door-to-door, or other community settings.

How the screening test might work

The purpose and trade-off between the sensitivity and specificity of universal screening strategies may change as the pandemic progresses. In the early stages of the spread of the virus, screening seeks to identify infected people, ideally with a high degree of accuracy, to contain spread. Without adequate sensitivity, containment efforts may fail due to false reassurance resulting from a negative screening test, which may be more likely to be a false negative. Poor specificity also carries harms: because the management strategy for positive screening tests may involve quarantining or self-isolation, and restriction of mobility until confirmatory testing can be completed. In the early stages of a pandemic, where prevention of transmission is the primary desired societal outcome, the harms from poor sensitivity may outweigh those of poor specificity.

When containment is no longer possible because of widespread community transmission or when resources are limited, widespread screening of the general population may be challenging and of limited utility, but widespread screening of specific populations (e.g. healthcare workers) may help support the healthcare system by both identifying positive cases (to avoid transmitting the virus in healthcare settings) and negative cases (to avoid sidelining non-infected workers) with a high degree of accuracy.

After sustained and prolonged community transmission, universal screening based on symptom assessment, direct temperature measurement, or feasible point of care testing may be used to identify people without prior infection with a high degree of accuracy, shield them from potential exposure, and support re-entry of those who have been infected with SARS-CoV-2 (with or without symptoms) into the workforce, schools, and community.

OBJECTIVES

To assess the effectiveness and accuracy of different screening strategies for SARS-CoV-2 in populations without known or suspected COVID-19 disease.

Specifically, the review addressed the following two questions.

1. What is the effectiveness of universal screening, that is, screening among people who have not presented to clinical care for symptoms related to COVID-19, for SARS-CoV-2 infection compared with (a) no screening or (b) screening in selected populations based on occupation, geographic setting, or community characteristics? Outcomes include incident cases, missed cases, successfully detected cases, averted cases, reduced transmission, mortality, false alarm, and false reassurance.
2. What is the accuracy of universal screening strategies among people who have not presented to clinical care for symptoms related to COVID-19 for SARS-CoV-2 infection? Outcomes include sensitivity, specificity, positive predictive value, negative predictive value, area under the receiver operating characteristic curve.

METHODS

We registered the protocol of this rapid review in the Open Science Framework on 8 April 2020 (Viswanathan 2020). We amended the protocol on 12 June 2020 and registered the amendment in Open Science Framework (osf.io/uqmxv). We revised the title from 'Screening asymptomatic persons for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)' to 'Universal screening for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)'. This change was intended to clarify the scope of the review and the specific outcomes being considered under the review. We also switched the order of the questions to be reviewed, to focus on effectiveness first, and then accuracy, and clarified eligible study designs for each question. We removed antibody tests from the list of eligible screening approaches because these are unlikely to be used for mass screening to identify active infections. Finally, we added false alarm and false reassurance to the effectiveness question to address potential harms from screening.

We adhered to the PRISMA statement throughout this manuscript (Moher 2009), and to PRISMA for diagnostic test accuracy (McInnes 2018). To conduct this rapid review, we employed abbreviated systematic review methods that adhered to guidance from the Cochrane Rapid Reviews Methods Group (Garrity 2020). Compared with the methods of a systematic review, the review team applied the following methodological shortcuts for this rapid review:

- no dual screening of included abstracts; if an abstract was included by one review author, the review team retrieved the full-text article;
- no dual, independent rating of the certainty of evidence; one review author conducted the ratings, a second review author checked for plausibility and correctness of data.

Criteria for considering studies for this review

Table 1 presents prespecified eligibility criteria for effectiveness, and Table 2 for accuracy.

Study design

To address our questions of interest, we included randomized and nonrandomized studies. In addition, we also included modelling studies, because we expected few empirical studies on screening effectiveness.

- Randomized controlled trials (RCTs)
- Quasi-RCTs
- Non-RCTs
- Prospective cohort studies
- Retrospective cohort studies
- Case-control studies (other than diagnostic case-control studies)
- Cross-sectional studies
- Controlled before-and-after studies

- Modelling studies

We excluded:

- case series;
- case reports;
- diagnostic case-control studies;
- systematic reviews (used for reference list checking).

For accuracy, we included any design on patients that provided information on test accuracy, including modelling studies. We excluded diagnostic case-control studies because they typically included participants who presented to healthcare for evaluation of suspected COVID-19 symptoms, which does not reflect the population of interest for this review.

Minimum study duration

We did not use a minimum duration of follow-up for effectiveness outcomes. The maximum duration between index test and reference standard was less than one day for diagnostic test accuracy studies.

Language

- English
- Chinese (English-language abstracts or, if available, English summaries provided by a Chinese WHO Collaborating Centre)

Types of participants

We included populations without known SARS-CoV-2 infection and who had not presented to healthcare for symptoms. Special populations of interest were healthcare workers without known SARS-CoV-2 infection or who had not presented to healthcare for symptoms.

We excluded studies focusing on diagnostic testing among people with presumptive COVID-19 symptoms and testing associated with contact tracing because of known exposure to individuals infected with SARS-CoV-2 (except for healthcare workers). We included studies with a subset of relevant participants if more than 80% met the inclusion criteria.

Types of interventions

Screening involves the use of assessments or laboratory tests to identify infections in populations without recognized symptoms. We included screening in unselected populations and screening in populations at potentially higher risk because of occupation, congregate living setting, or travel status. We excluded case finding in symptomatic populations.

We included the following screening strategies for effectiveness and accuracy:

- clinical symptom assessment (e.g. directly asking individuals about fever, dry cough, shortness of breath, gastrointestinal symptoms, loss of smell or taste or using existing medical records to retrospectively assess the presence of symptoms);
- epidemiological risk assessment (e.g. asking about history of travel from hot spots, asking about contact with people known or suspected of having SARS-CoV-2 infection);
- body temperature measurement;
- feasible point-of-care laboratory tests (e.g. rapid RT-PCR tests, CRISPR tests);
- a combination of the above approaches.

We included studies of repeat screening strategies, which involves rescreening among people confirmed to be negative for SARS-CoV-2 in an initial round of screening.

Types of comparators or reference standards

- For the effectiveness key question, we included studies comparing universal screening with no screening, or universal screening with screening in selected populations based on occupation, geographic setting, or community characteristics.
- For the accuracy key question, we included studies comparing results from index screening strategies with confirmatory reference standards such as standard RT-PCR, CRISPR, or a combination of these approaches.

Outcomes

For the effectiveness key question, outcomes of interest were:

- incident cases;
- missed cases;
- successfully detected cases;
- averted cases;
- reduced transmission;
- mortality;
- false alarm; and
- false reassurance.

For the accuracy key question, outcomes of interest were:

- sensitivity;
- specificity;
- positive predictive value;
- negative predictive value; or
- area under the receiver operating characteristic curve.

We included studies in the review even if only partial results were available (e.g. some results were missing for sensitivity or specificity). We sent out queries to study authors for unclear outcome reporting.

Search methods for identification of studies

We conducted the latest update searches in May 2020. An experienced information specialist (IK) designed and conducted systematic searches of the literature published in the following databases without language restrictions.

- Ovid MEDLINE(R) ALL from 1 November 2019 to 26 May 2020 (search date: 27 May 2020);
- Centers for Disease Control (CDC) COVID-19 Research Articles Downloadable Database from inception to 26 May 2020 (search date: 27 May 2020);
- Embase.com (Elsevier) from 1 November 2019 to 14 April 2020 (search date: 14 April 2020);
- Cochrane Central Register of Controlled Trials (CENTRAL; 2020, issue 4) in the Cochrane Library, from 1 November 2019 to 12 April 2020 (search date: 14 April 2020);

- Cochrane Covid-19 Study Register (covid-19.cochrane.org/) from inception to 14 April 2020 (search date: 14 April 2020)
- LitCovid from inception to 4 April 2020 (search date: 4 April 2020).

In addition, we searched the CDC COVID-19 Research Articles Downloadable Database and the Cochrane Covid-19 Study Register (covid-19.cochrane.org/) from inception to 13 April 2020. We searched LitCovid (www.ncbi.nlm.nih.gov/research/coronavirus/), from inception to 4 April 2020. We did not use some databases in the May 2020 update searches because they had not contributed unique relevant references to the previous search results. These searches included preprints in addition to published citations.

See Appendix 1; Appendix 2; Appendix 3; Appendix 4; Appendix 5; Appendix 6; and Appendix 7 for the full search strategies.

In addition, WHO provided a list of citations and abstracts prepared by the WHO Collaborating Centre for Guideline Implementation and Knowledge Translation, Lanzhou, China. They conducted manual searches of selected Chinese journals as well as searches of the China National Knowledge Infrastructure. For articles published in Chinese only, the Collaborating Centre prepared an English translation of the abstracts. If no abstract was available, the Collaborating Centre provided a 'Brief summary' in English. This list was updated on a daily basis. The last date of the search considered for this review was 15 April 2020. Additionally, the WHO provided us with a collection of abstracts from PubMed on a daily basis up to 15 April 2020.

In addition, review authors screened reference lists of relevant publications and included studies for additional citations.

Finally, we also searched three model repositories (Covid-Analytics (app.smartsheet.com/b/publish?EQBCT=1a3bc6acad99475f99acfd55a04a1564), Models of Infectious Disease Agent Study (MIDAS, www.covidanalytics.io/), and Society for Medical Decision Making (midasnetwork.us/covid-19/)) on 8 April 2020.

Data collection and analysis

Selection of studies

A team of experienced review authors screened all titles and abstracts based on predefined inclusion and exclusion criteria (Table 1; Table 2). Using a standardized title and abstract form, we conducted a pilot exercise using the same 50 abstracts for the entire screening team to calibrate the team and test the review form. Abstracts were advanced to full-text review after inclusion by one review author. Abstracts excluded by a first review author were dually reviewed by a senior review author (MV, GG, LCK) prior to exclusion. The review author team retrieved the full texts of all available included abstracts. Two review authors, including one senior review author (MV, GG, LCK, BJ), screened all full-text publications independently. They resolved disagreements by consensus or by involving a third, senior review author. Reasons for exclusion were recorded for all studies that were excluded during full text review.

The team conducted literature screening using DistillerSR (Evidence Partners, Ottawa, Canada, www.evidencepartners.com/).

Inclusion of non-English language studies

The review team screened English-language abstracts of Chinese publications. Collaborators of the Cochrane China Network translated full texts of relevant abstracts into English.

Data extraction and management

One review author extracted data from the included studies into standardized forms; a second senior review author (MV, GG, LCK, BJ) checked the data extraction for completeness and correctness. For diagnostic test accuracy studies, data items were author, year, setting, number of participants, proportion of female participants, comorbidities, screening strategy, reference test, and results (Appendix 8).

For modelling studies, the data items were author, year, research question, type of model, screening population, setting, main model parameters (e.g. incubation period, test accuracy, detectability of symptoms, dynamics of epidemic), screening strategy, and results (Appendix 9).

Assessment of methodological quality of included studies

Two senior review authors (MV, GG, LCK, BJ) independently rated the methodological quality of included studies. For diagnostic test accuracy studies, we used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool (Whiting 2011). We did not tailor any of the standard QUADAS-2 signalling questions, but rather considered each signalling question with respect to the context of the review. For example, for index tests assessing symptoms, we considered whether the study had an assessment that listed all symptoms, over what time period those symptoms were assessed, and specific criteria for determining a positive test based on symptoms. For the criteria related to whether the reference standard is likely to correctly classify the target condition, we considered any studies that used RT-PCR based on single specimen collection as unclear on this domain given the evolving evidence related to false negatives. We rated each of the QUADAS domains as having low, unclear, or high risk of bias and then assigned an overall study 'Risk of bias' rating of low, unclear, or high. We assigned studies with any domain rated as high as having an overall high risk of bias rating; similarly studies with any domain rated as unclear were assigned as having an unclear overall risk of bias rating. For these studies, we also rated the applicability concerns for participant selection, index test, and reference standard as low, unclear, or high. Because no validated 'Risk of bias' checklist for transmission models is available, we adapted a tool for the methodological quality of economic evaluations (Evers 2005). We rated the methodological quality of studies using three categories:

- no or minor concerns;
- moderate concerns; or
- major concerns.

We resolved disagreements between review authors by discussion or by consulting a third review author.

Contacting study authors

In cases of missing or contradictory outcome data, we contacted the study authors by email to ask for clarification or additional data.

Data synthesis

We synthesized results narratively and in tabular form. Because of the heterogeneity of available primary studies, we did not report quantitative analyses. We produced a paired forest plot for sensitivity and specificity with 95% CI.

Assessment of the certainty of the evidence

To rate the certainty of evidence, we followed GRADE guidance (Brozek 2020; Schünemann 2008). We rated the certainty of evidence for the following outcomes: days to outbreak, cases missed or detected, diagnostic accuracy (i.e. true positives, false positives, true negatives, false negatives). One senior review author rated the certainty of evidence using GRADEpro GDT. A second senior review author checked for plausibility and correctness of data.

RESULTS

Results of the searches

We identified 4371 deduplicated citations through our systematic searches, last conducted on 26 May 2020. Additionally, we identified 4407 abstracts from Chinese database searches and 2189 search results from the WHO; from these we pulled eight for further review (full-text for articles in English, and translation of background and methods for articles in Chinese). We excluded all of them (diagnostic case control: four; wrong population: three; wrong or no screening test: one). We also included two eligible publications from the handsearches (Gudbjartsson 2020b; Sutton 2020). In addition, we found one companion article in handsearch (Gudbjartsson 2020b), which was an update of a preprint publication (Gudbjartsson 2020a). The systematic searches and searches from Chinese databases and from the WHO may have had some overlap, but their formats did not permit merging to allow the identification of duplicates. We included 22 studies. (See Characteristics of included studies; Figure 2).

Figure 2. Study flow diagram

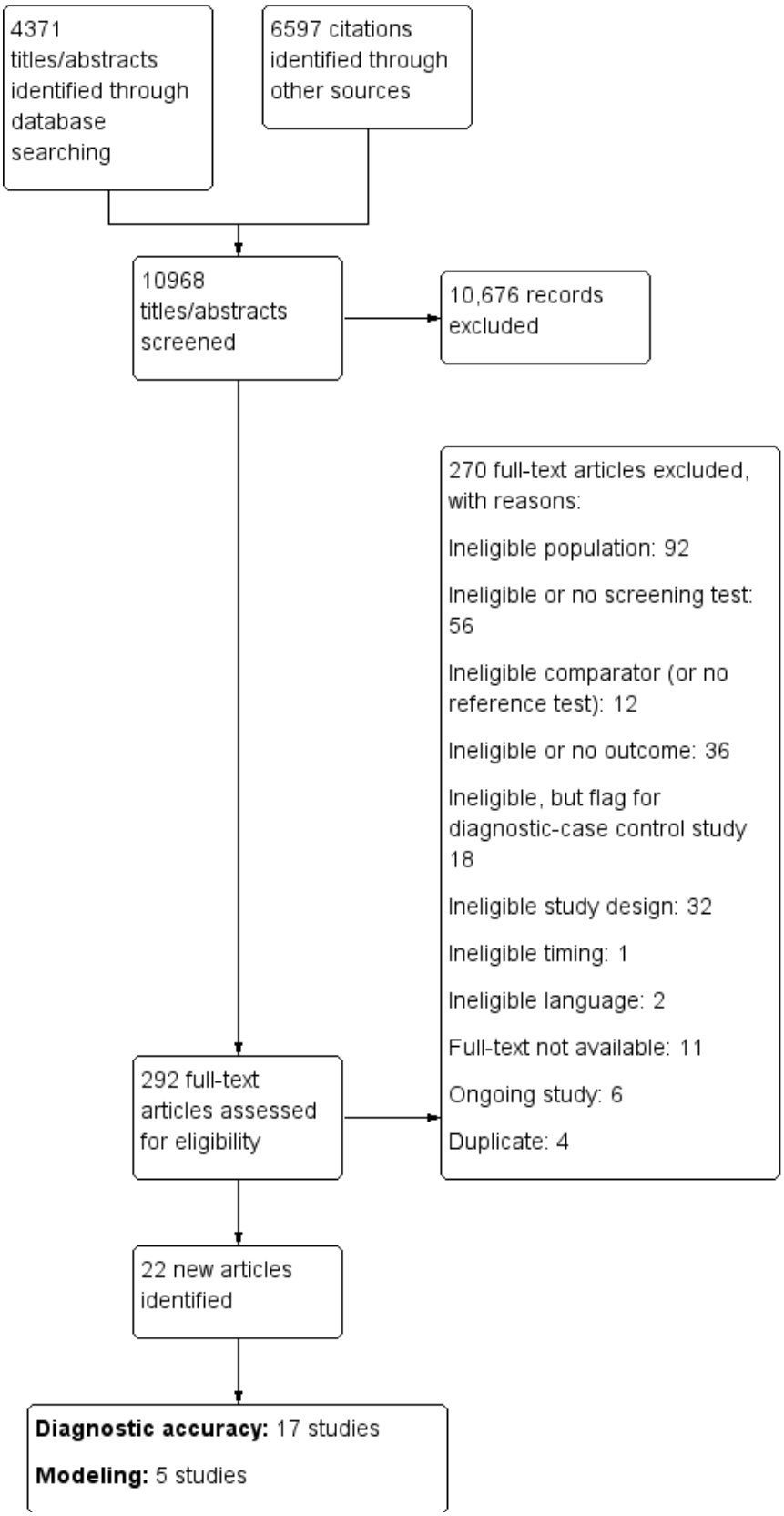
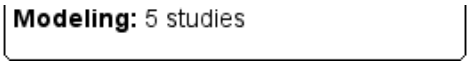


Figure 2. (Continued)



Included studies

See Characteristics of included studies for detailed study characteristics for accuracy and modelling studies.

Effectiveness

We found two modelling studies that evaluated the effectiveness of two screening strategies (Clifford 2020; Zhang 2020). Clifford 2020 addressed symptom screening of air travellers, Zhuang 2020 addressed periodic laboratory testing of asymptomatic healthcare workers in emergency departments. Neither study further specified the screening tests. Clifford 2020 assumed a sensitivity of 86%, Zhang 2020 of 100%. Clifford 2020, assessing symptom screening of air travellers, also considered the effectiveness of screening in combination with sensitizing travellers about the importance of self-isolation and rapid care-seeking in the case of symptoms.

Screening intervals were specified by a 12-hour flight time between arrival and departure airports for the screening of air travellers (Clifford 2020), and weekly or biweekly for the study on laboratory screening of healthcare workers (Zhang 2020). To determine the time for potential symptomatic detections, Clifford 2020 assumed an incubation period of 5.2 days, and Zhang 2020 of 4 days.

Main outcome measures in Clifford 2020 were the delay of a COVID-19 outbreak in a previously unaffected country, and the reduction of infections in healthcare workers and patients at an emergency department in Zhang 2020.

We did not find any evidence on potential harms such as false alarm and false reassurance.

Accuracy: primary studies

We identified 17 eligible publications reporting on diagnostic accuracy (Arons 2020; Baggett 2020; Barrett 2020; Graham 2020; Gudbjartsson 2020b; Guery 2020; Hoehl 2020; Kimball 2020; Lavezzo 2020; Lytras 2020; Nishiura 2020; Olalla 2020; Osterdahl 2020; Roxby 2020; Samuels 2020; Sriwijitalai 2020; Sutton 2020).

The 17 publications include instances of multiple papers on one cohort and multiple cohorts in one publication. Two publications report on multiple rounds of screening in a single cohort of nursing home residents in King County, Washington State, USA. One reported on the first point-prevalence screening (Kimball 2020), and a second publication reported on the second point-prevalence screening and combined results of the two point-prevalence screenings (Arons 2020). One publication reported on two cohorts (Gudbjartsson 2020b). The Gudbjartsson 2020b study, set in Iceland, reported on an open-screening period (13 March to 1 April 2020), where all people with mild or no symptoms were invited to screen for symptoms and have the RT-PCR test. This period of recruitment was succeeded by another recruitment (1 April to 4 April 2020) that sent out invitations to a random sample of citizens for symptom screening, along with the RT-PCR test.

Frequency of screening

Of the 17 cohorts, 16 reported on the accuracy of a single point-in-time screening strategy (Baggett 2020; Barrett 2020; Graham 2020; Gudbjartsson 2020b; Guery 2020; Hoehl 2020; Kimball 2020; Lavezzo 2020; Lytras 2020; Nishiura 2020; Olalla 2020; Osterdahl 2020; Roxby 2020; Samuels 2020; Sriwijitalai 2020; Sutton 2020). One reported on cumulative screenings (Arons 2020).

Type of screening strategy

Together, these studies reported on the accuracy of eight distinct screening strategies. Specifically:

- 12 publications reported on the accuracy of symptom screening in 13 cohorts (Baggett 2020; Barrett 2020; Graham 2020; Gudbjartsson 2020b; Guery 2020; Hoehl 2020; Kimball 2020; Lavezzo 2020; Lytras 2020; Olalla 2020; Roxby 2020; Samuels 2020);
- one publication reported repeated screening for symptoms (Arons 2020);
- three publications reported on the accuracy of direct temperature measurement (Barrett 2020; Hoehl 2020; Samuels 2020);
- one publication, comprising two cohorts, reported on screening by asking about international travel history (Gudbjartsson 2020b);
- two publications reported on screening by asking about exposure to known infected people in three cohorts (Gudbjartsson 2020b; Hoehl 2020);
- two publications reported on the accuracy on screening by asking about exposure to known or suspected infections (Barrett 2020; Hoehl 2020);
- three publications reported on the accuracy of screening using a symptom assessment and direct temperature measurement (Nishiura 2020; Sutton 2020 Sriwijitalai 2020);
- one publication reported on screening with a rapid, point-of-care test (reverse transcriptase loop-mediated isothermal amplification (RT-LAMP)) (Osterdahl 2020).

All compared index tests to RT-PCR as the reference standard.

Settings

Studies were set in France, Germany, Greece, Italy, Iceland, Japan, Spain, Thailand, UK, and USA.

With the exception of the three population-based samples (Gudbjartsson 2020b; Lavezzo 2020), other studies were conducted in specific populations and settings. Five publications reported on four cohorts in nursing homes or senior communities (Arons 2020; Graham 2020; Kimball 2020; Osterdahl 2020; Roxby 2020), three reported on international travellers (two were specifically of evacuees from Hubei, China; Hoehl 2020; Nishiura 2020; Sriwijitalai 2020), two predominantly or wholly comprised healthcare workers (Barrett 2020; Guery 2020), two reported on people experiencing homelessness in congregate shelters (Baggett 2020; Samuels 2020),

and one was of pregnant women admitted for delivery in a New York City hospital (Sutton 2020).

Prevalence

The prevalence of positive cases varied by setting and may also represent the community prevalence at the time of sample collection. Population-based samples reported the lowest prevalence. The Iceland screening study (Gudbjartsson 2020b), reported prevalence ranging from 0.6% to 0.8% for their open and random screening arms, and prevalence in the study of the Italian municipality was 2.6% (Lavezzo 2020). Prevalence in all other studies was higher on average. Prevalence in travel cohorts ranged from 1.6% (Hoehl 2020) to 5.1% (Lytras 2020). Prevalence in the senior independent and assisted living facility was 3.5% (Roxby 2020). Prevalence among healthcare workers ranged from 0.4% (Olalla 2020) to 4.9% (Barrett 2020). The prevalence among the sample of consecutive pregnant women in New York was 15% (Sutton 2020). The prevalence in homeless shelters ranged from 11.7% (Samuels 2020) to 36.0% (Baggett 2020). The prevalence in the nursing homes was the highest on average and ranged from 30.3% (USA, Kimball, 2020) to 47.6% (UK, Osterdahl, 2020) in single point-in-time screenings. A repeat screening in a nursing home found that as many as 62% identified positive (Arons 2020).

Timing of screening

One study did not specify the period of testing, but others noted testing dates ranging from 1 February 2020, to 19 to 24 April 2020.

Accuracy: modelling studies

We found three modelling studies evaluating the accuracy of screening at airports (Gostic 2020; Quilty 2020; Wells 2020). Main outcome measures were cases missed or detected by entry or exit screening, or both, at airports (Gostic 2020; Quilty 2020; Wells 2020). In the evaluated screening strategies, studies evaluated symptom screening with infrared thermal scanner alone (Quilty 2020), or in combination with self-reporting of exposure risk via questionnaires (Gostic 2020). One study calculated the probability of case detection based on estimates of presymptomatic, infected people upon arrival at destination airports and the distribution of countries that asked travellers about travel history (Wells 2020).

Screening intervals, specified by the flight time between arrival and exit airports ranged from 12 hours (Quilty 2020), to 24 hours (Gostic 2020). To determine the time for potential symptomatic detections, models considered an incubation period between 4.5 and 6.5 days and the time until hospitalization that would prevent people from travelling ranging from 7.5 to 13.5 days. Assumed sensitivity of symptom screening ranged from 70% (Gostic 2020), to 86% (Quilty 2020), in the base-case analyses. Screening for exposure risk considered awareness and the probability that people would report correctly on exposure risk.

Excluded studies

We excluded 270 studies at the full-text stage, and list the reasons for exclusion in Figure 2. For a full list of excluded studies, please contact the corresponding author.

Risk of bias in included studies

We have listed the criteria for assessing the methodological quality of modelling studies in Appendix 10. The methodological quality of the five modelling studies was mixed. Detailed assessments are in Appendix 11 and Appendix 12. For two studies, we had no or only minor concerns regarding the methodological quality (Clifford 2020; Gostic 2020); for one study, we had moderate methodological concerns (Quilty 2020), and major concerns for two studies (Wells 2020; Zhang 2020). Methodological concerns pertained to unclear definitions of outcomes and research objectives, lack of sensitivity analyses, or the lack of assessing competing alternatives.

We assessed one diagnostic test accuracy study as having an high risk of bias overall (Barrett 2020); we assessed the others as having an unclear risk of bias overall. Detailed 'Risk of bias' assessments are in Appendix 13; Appendix 14; Appendix 15; Appendix 16. Although these studies generally had no concerns about risk of bias with respect to participant selection and flow and timing of testing, specification around the index screening strategy and threshold for a ‘positive’ index test was not well-specified. For example, studies that assessed symptoms were not always clear about the symptoms that were assessed, the time period over which people were asked about symptoms (e.g. symptoms on day of testing, or symptoms over the prior week or two weeks), and whether the symptoms were assessed relative to baseline symptoms (for example in the case of people with chronic cough). With respect to the reference standard, it is unclear whether the RT-PCR (the reference standard in all studies) can accurately classify infection. Although the laboratory assay appears to have high analytic validity; the use of inadequate or improper sampling techniques may reduce its validity for use as a reference standard (Tang 2020) if only based on one collected swab. Further, validity may also be affected by mismatches between primers and probes and the target sequence, amplification inhibitors in the sample, sampling timing with respect to viral load, and inappropriate sample handling (Tahamtan 2020). Lastly, no studies reported whether results of index and reference tests were interpreted without knowledge of the other result, however, the flow of testing in these studies suggests:

- it is unlikely that the same people conducted both tests; and
- the reference standard results would not have been available at the time of the index test.

Effects of interventions/results of the synthesis

Effectiveness

Two modelling studies provided data on the effectiveness of screening strategies (Clifford 2020; Zhang 2020). Table 3 provides a study-level summary of the evidence from eligible studies. One study addressed symptom screening of air travellers (Clifford 2020); the other study addressed laboratory screening of asymptomatic healthcare workers in emergency departments (Zhang 2020). Table 4 and Table 5 present our assessment of the certainty of the evidence.

The first modelling study assessed whether symptom screening aimed at air travellers can delay a SARS-CoV-2 outbreak in a previously unaffected country (Clifford 2020). Study authors hypothesized that the effectiveness of symptom screening is based on a combination of detection (assumed sensitivity of 46% for arrival and departure screening and 42% for departure screening only) and sensitization. Sensitization was defined as an increase in self-awareness of symptoms and an increased likelihood of self-quarantine in those who develop symptoms.

Assuming a best-case scenario (only one infected traveller per week and 50% of travellers are sensitized), the model showed that symptom screening may delay an outbreak in an unaffected country by 83 days (50% of simulations showed a delay of at least of 83 days; 75% of at least 36 days, and 97.5% of at least 8 days; Clifford 2020). However, assuming 10 or 100 infected travellers per week reduced the delay to 8 or 1 days, respectively. Under the best-case scenario, departure screening alone led to a delay of the outbreak by 76 days (75%: 33 days, 97.5%: 7 days).

The second modelling study (available as a preprint publication only) assessed the effectiveness of routine laboratory screening of asymptomatic healthcare workers in emergency departments with an unspecified test (Zhang 2020). For this study, we had major methodological concerns because of a lack of sensitivity analyses, a lack of information on recovery rates, and model assumptions that seemed unrealistic (e.g. 100% sensitivity of screening test).

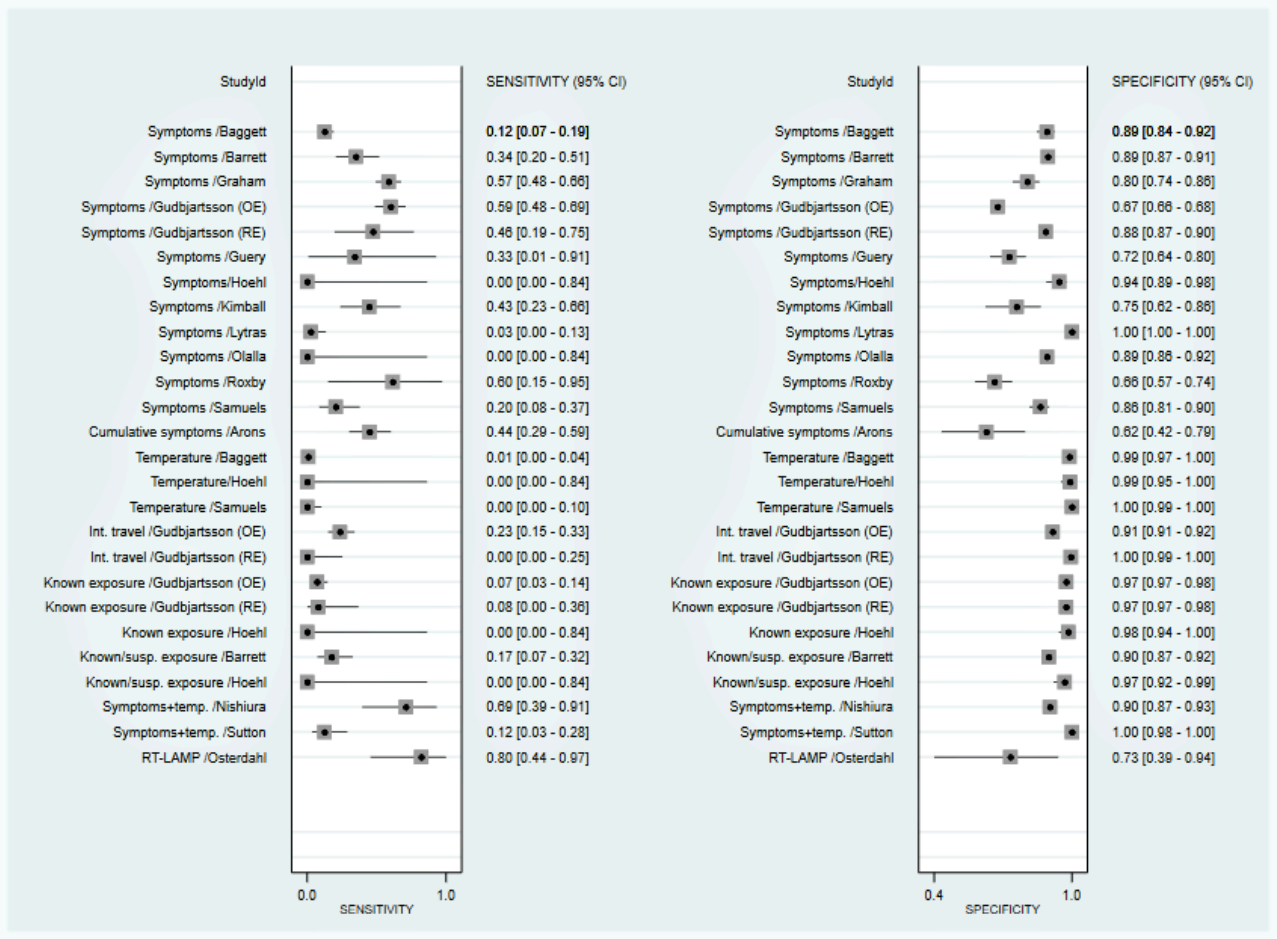
Assuming different transmission constants in the population, study authors assessed the reduction of transmissions with weekly and biweekly testing of healthcare personnel. Assuming a transmission constant of 1.2 new infections per 10,000 people, weekly screening reduced patient and healthcare worker infections in emergency rooms by 5.1% within 30 days. Biweekly screening reduced infections by 2.3% (Zhang 2020).

Taking a higher transmission constant into consideration (3.7 new infections per 10,000 people), the reduction of infections for weekly screening was 21.1% (for healthcare workers and patients) and 9.8% for patients (9.7% for healthcare workers) for biweekly screening within 30 days.

Accuracy

Table 6 presents strategy-specific summary results. Figure 3 offers a paired forest plot across all screening approaches. Table 7 presents a study-level summary of the evidence from eligible studies, and Table 8, Table 9, Table 10, Table 11, Table 12, Table 13, Table 14 and Table 15 present our assessment of the certainty of the evidence. The selected prevalence of 1%, 15%, and 30% in Tables 8 through 15 represented low, mid, and high ranges of prevalence that we found in our eligible studies. Overall, studies reported sensitivities ranging from 0 to 0.80. Specificity was higher than sensitivity in all studies, ranging from 0.62 to 1.00. All studies were rated as having uncertain or high risk of bias. The certainty for estimates of sensitivity ranged from very low to low and for specificity from low to moderate.

Figure 3. Paired forest plot of sensitivity and specificity for included studies*



*Excludes studies without complete results (Lavezzo 2020; Sriwijitalai 2020)

We describe the results from primary studies for each screening strategy below. We thereafter present results from three modelling studies following hypothetical cohorts of travellers.

Single screening: screening with a symptom assessment

Twelve publications, comprising 13 cohorts with data on 16,762 participants, presented results on accuracy of screening for symptoms (Baggett 2020; Barrett 2020; Graham 2020; Gudbjartsson 2020b; Guery 2020; Hoehl 2020; Kimball 2020; Lavezzo 2020; Lytras 2020; Olalla 2020; Roxby 2020; Roxby 2020). All used RT-PCR as the reference standard. Only four specified site of sample collection (Graham 2020; Guery 2020; Kimball 2020; Samuels 2020). All used nasopharyngeal swabs; Kimball 2020 also used oropharyngeal swabs among some residents. These studies varied substantially in baseline prevalence, ranging from 0.6% for population screening from a random sample in Iceland (Gudbjartsson 2020b), to 40.5% in a nursing home in the UK (Graham 2020). The source populations also varied substantially and included nursing home staff and residents (prevalence ranged from 30.3% to 40.33%), homeless shelters (11.7% to 36.2%), assisted living (3.5%), healthcare workers (0.4% to 4.9%), international travellers (1.6% to 5.1%), and the general population (0.6% to 2.6%). All cohorts used RT-PCR as the reference standard. The methods and specific symptoms included in each study also varied. Measurement approaches varied from history taking with prespecified symptoms on assessment forms or reviewing case reports and the clinical team to asking for self-reported symptoms from patients without further specification. Although some studies distinguished between chronic ongoing symptoms and new symptoms, others did not. Some studies specified a 14-day look-back period, others did not. Although studies did not always specify a threshold, those that did set it as any symptom. Although the heterogeneity in sensitivity (ranging from 0.00 to 0.60) may be explained by these wide variations in prevalence, settings, and methods of measurement, we did not find clear patterns of association between any single factor and sensitivity. It is also unclear whether studies generally captured symptoms early, when participants were likely to be asymptomatic. Specificity was higher than sensitivity, ranging from 0.66 to 1.00.

Repeated screening: screening with symptom assessment

Arons 2020 reported the accuracy of cumulative screenings in 76 nursing home residents (Kimball 2020 reported the results of the first screening). RT-PCR was the reference standard. The study collected nasopharyngeal swabs on all residents and oropharyngeal swabs on most residents. Of the 76 nursing home residents tested in week one, 49 had tested negative and were retested in week two. Over the two-week period, 62% tested positive. The sensitivity and specificity of the cumulative screening were 0.44 (95% CI 0.29 to 0.59) and 0.62 (95% CI 0.42 to 0.79) respectively. Arons 2020 also reported that 9 of 24 participants testing positive in the second screening at week two had symptoms (sensitivity: 0.38).

Single screening: screening with direct temperature measurement

Three studies (822 people), reported on the accuracy of temperature checks (Baggett 2020; Hoehl 2020; Samuels 2020). Hoehl 2020 studied evacuees from Wuhan, and Baggett 2020 and Samuels 2020 were conducted in homeless shelters. All cohorts used RT-PCR as the reference standard. Only Samuels 2020 specified site of sample collection; the study used nasopharyngeal swabs. Prevalence ranged from 1.7% to 36.0%. Sensitivity ranged from 0.00 to 0.01 and specificity from 0.99 to 1.00 in the three cohorts.

Single screening: screening by asking about international travel history

Gudbjartsson 2020b reported results from two cohorts (13,080 people) on the accuracy of screening for travel history. Prevalence ranged from 0.6% to 0.8%. RT-PCR was the reference standard; the site of sample collection was not specified. Sensitivity ranged from 0.0 to 0.23 and specificity ranged from 0.91 to 1.00 in the two cohorts.

Single screening: screening by asking about exposure to known infected people

Two publications including three cohorts (13,205 people), reported on exposure to known or suspected infected people (Gudbjartsson 2020b; Hoehl 2020). Gudbjartsson 2020b reported results from two population-based cohorts (13,080 people) on the accuracy of screening for exposure to known infected people. All cohorts used RT-PCR as the reference standard; the site of sample collection was not specified. Prevalence ranged from 0.6% to 0.8%. Sensitivity ranged from 0.07 to 0.08 and specificity was 0.97 in the two cohorts. A third study (Hoehl 2020), reported a 1.6% prevalence in evacuees returning from Wuhan and reported a sensitivity of 0.00 and specificity of 0.98.

Single screening: screening for exposure to known or suspected infected people

Two studies (654 people) reported on exposure to known or suspected infected people (Barrett 2020; Hoehl 2020). Hoehl 2020 reported exposure to a known or suspected infected person in a cohort of evacuees from Wuhan. In a study primarily comprising healthcare workers in a university, Barrett 2020 also assessed recent exposure to known or suspected infected people, but specified exposure as outside the workplace. RT-PCR was the reference standard; the site of sample collection was not specified. Prevalence ranged from 1.6% to 4.9%. Sensitivity ranged from 0.0 to 0.17 and specificity from 0.90 to 0.97 in the two cohorts.

Single screening: screening for symptoms and temperature

Regarding the accuracy of screening for symptoms combined with universal temperature checks, two of three studies (779 people) reported complete accuracy results (Nishiura 2020; Sutton 2020). Nishiura 2020 reported on evacuees from China, and Sutton 2020 reported on pregnant women in New York. RT-PCR was the reference standard; only Sutton 2020 reported the site of sample collection (nasopharyngeal). Baseline prevalence was 2.3% among evacuees and 15.5% in pregnant women. The range of sensitivity varied widely, from 0.12 to 0.69. Specificity was higher, ranging from 0.90 to 1.00.

A third study did not provide adequate information to assess accuracy but reported that only 1 of 12 symptomatic cases tested positive for SARS-CoV2 (Sriwijitalai 2020).

Single screening: screening with RT-LAMP

Osterdahl 2020 reported that RT-LAMP had a sensitivity of 0.80 (95% CI 0.44 to 0.96) and a specificity of 0.73 (95% CI 0.39 to 0.94) based on screening of 21 high-dependency care home residents. RT-PCR was the reference standard, with specimens collected via pharyngeal and deep nasal swabs on two consecutive days.

Modelling studies

Three modelling studies followed hypothetical cohorts of infected travellers who underwent screening at the departure and arrival airports (Gostic 2020; Quilty 2020; Wells 2020).

Gostic 2020 simulated screening using thermal image scanning and travel history questionnaires at the departure and arrival airports. The model assumed a mean incubation period of 5.5 days, a flight duration of 24 hours, a sensitivity of thermal image scanning of 70%, and that 25% of travellers truthfully self-report exposure risks. Even in a best-case scenario, assuming that only 5% of infected travellers would be asymptomatic, the sensitivity of the screening strategy was 0.30 (95% CI 0.1 to 0.53), missing 70% of infected travellers. When the proportion of asymptomatic infected travellers was increased to 25% and 50%, the sensitivity decreased to 0.27 and 0.20, respectively.

Quilty 2020 modelled arrival and departure screening with thermal image scanners at airports with more optimistic assumptions about the sensitivity of thermal image scanning than Gostic 2020 (86% sensitivity). Assuming that 17% of infected travellers would be asymptomatic, the sensitivity to detect infected travellers was 0.54 (95% CI 0.64 to 0.42). In a scenario analysis, study authors assessed how parameters would need to be changed to achieve a detection rate of 90% of infected travellers. Results showed that a 90% detection rate would only be possible with an unrealistic scenario of 0% asymptomatic infections, 100% sensitivity of arrival screening and an incubation period that is around 10-fold shorter than the period from symptom onset to severe disease (e.g. hospitalization) (Quilty 2020).

A low methodological quality study by (Wells 2020), found that screening travellers detected 82 (95% CI 72 to 95) cases that would have been imported from mainland China by other countries up to 15 February 2020. The study does not specify the denominator (total imported cases in the destination country), so the implication of this outcome is unclear.

DISCUSSION

Summary of main findings and certainty of the evidence

In emerging stages of a pandemic, the primary focus is on containment of the disease. Because a substantial proportion of individuals infected with SARS-CoV-2 never develop COVID-19 (Gudbjartsson 2020b), universal screening, that is, screening among those not seeking care for symptoms, could be an important strategy to identify positive cases and contain the spread of the disease. Screening may also be targeted based on occupation or other risk factors. For example, screening of all healthcare workers in hospitals and nursing homes may protect vulnerable patients and residents.

Regarding effectiveness of screening for SARS-CoV-2 infection, we did not find any primary studies addressing this question. A modelling study showed that screening at travel hubs, such as airports, may slightly slow but not stop the importation of infected cases (Clifford 2020). We rated the certainty of evidence on days to an outbreak as low. Symptom screening upon departure from or arrival at airports and sensitization of travellers regarding symptoms of COVID-19 may delay a major outbreak, but only in the very early stages, when prevalence among travellers is likely to be very low. These results are mainly driven by the assumed incubation period, proportion of infected but asymptomatic individuals, and test sensitivity. In this model, however, the assumed sensitivity of symptom screening was higher than sensitivities reported in screening accuracy studies included in this review. We did not find comprehensive transmission models demonstrating long-term impact of screening interventions in a situation where many countries are affected but disease control measures including travel restrictions are slowly loosened (Tang 2020; WHO 2020d; Xiao 2020a; Xiao 2020b; Xie 2020).

A low-quality modelling study reported that weekly or biweekly screening of healthcare workers with an unspecified laboratory test can substantially reduce transmission of the disease to patients and other healthcare workers in emergency rooms (Zhang 2020). We rated the certainty of this evidence as very low because of unrealistic assumptions of the model. For example, the model assumed a test sensitivity of 100% and that transmission patterns in emergency rooms are the same as on cruise ships.

Regarding the accuracy of screening, overall the evidence indicates that current screening strategies are not sensitive, meaning they cannot accurately identify people with infections among populations not already known to have symptoms. The timing of screening relative to course of infection was unclear in most studies; low sensitivity may be related to screening generally occurring shortly after exposure, when many participants are asymptomatic. Regarding test accuracy for single point-in-time screening, we found evidence on seven strategies: symptom assessment, direct temperature measurement, travel history, assessment of exposure to known infected people, assessment of exposure to known or suspected infected people, symptom assessment plus direct temperature measurement, and a rapid point-of-care PCR test. We did not find any evidence on one-time or repeated screening with standard RT-PCR as the screening test.

Studies reported sensitivities ranging from 0 to 0.80. These point estimates were generally accompanied by wide CIs, reflecting the uncertainty of these estimates. Variations in prevalence, settings, and methods of measurement may be associated with heterogeneity in sensitivity, although we did not find clear patterns of association with any single factor.

Point estimates for specificity were higher than for sensitivity in all studies, ranging from 0.62 to 1.00, but may still be too low, resulting in many false positives that would have to undergo confirmatory testing and possible isolation while waiting on confirmatory results. As for sensitivity, point estimates had wide CIs. The certainty of evidence was very low to low for sensitivity and low to moderate for specificity for assessing symptoms, directly measuring temperature, or asking about international travel history, or exposure to known or suspected infected people. The certainty of evidence was very low for sensitivity and ranged from very low to low for specificity for a rapid PCR test (specifically RT-LAMP), combined approaches (specifically, symptoms plus exposure), and repeated screening to assess symptoms.

No single strategy paired high sensitivity with high specificity. The highest reported sensitivity came from a single study of RT-LAMP (Osterdahl 2020), (sensitivity = 0.80,and specificity = 0.73), but these results are from a study of uncertain bias with only 21 participants, leading to very low certainty in the results.

Screening with symptom assessment had low sensitivity (point estimates ranged from 0.00 to 0.60) and moderate specificity (point estimates ranged from 0.66 to 1.00). The value of symptom screening is limited in the context of poor sensitivity. A major consideration limiting the utility of screening based on symptoms is the long latency period of the virus before symptoms manifest (current estimates are an average of five days; WHO 2020d). The evidence on repeated screening for symptoms was very limited, leading to uncertainty in our conclusions.

Screening using direct temperature measurement similarly has low sensitivity (point estimates 0.00 to 0.01). Despite moderate-certainty evidence that temperature measurement alone can be highly specific (i.e. a high temperature is more likely to reflect an infection that would require confirmatory follow-up testing than a false positive), the utility of universal screening with direct temperature measurement may be limited given poor sensitivity (i.e. a normal temperature misses identifying many true infections) because the natural history of this infection includes a relatively long pre-symptomatic duration and sometimes even has an asymptomatic course of disease.

Screening for travel history had poor sensitivity, with point estimates ranging from 0 to 0.23. Specificity for travel history ranged from 0.91 to 1.00. Poor sensitivity accompanied by changes in travel policy across many countries as the pandemic grows severely limits the utility of a screening approach using travel history.

Screening for knowledge of exposure to known infected people also had poor sensitivity, ranging from 0.00 to 0.08. Specificity for knowledge of exposure to known infected individuals was 0.97 to 0.98. A broader approach, asking about screening for knowledge of known or suspected infected people also had similar sensitivity and specificity estimates. Sensitivity ranged from 0.0 to 0.17 and specificity from 0.90 to 0.97. As with screening for travel history, poor sensitivity, widespread community transmission, and low testing prevalence in some settings can reduce the utility of questions regarding known or suspected exposure to infected people as a screening strategy.

We found only one study of a rapid viral detection test (Osterdahl 2020); we anticipate that many evaluations of new, rapid tests will be published in the coming weeks and months.

Notably, however, the screening strategies described all have negative predictive values of 0.95 or higher when prevalence is 5% or lower, that is, they offer reassurance to those who test negative that they are unlikely to be infected. When community prevalence is known to be low, screening approaches may result in few false negatives among those with negative tests (although they would miss significant proportions of infected people). A key stumbling block to using community prevalence as a signpost for whether to screen is that community prevalence data may be flawed or missing.

In studies with higher prevalence, the negative predictive value of these screening approaches is lower. With low sensitivity, however, many of those infected would be missed.

Limitations of the evidence base

We found only modelling studies on the effectiveness question with unrealistically positive assumptions. No studies reported on harms. Most of the articles we included were not designed specifically as test accuracy studies; rather they were designed as point prevalence studies or reports of public health practice in action. Thus, detailed information about the way in which index tests were implemented and thresholds for what was considered a 'positive' index test were not always reported.

Another limitation of this evidence concerns the validity of RT-PCR as the reference standard for identifying SARS-CoV-2 infection. Although RT-PCR has very good analytic validity under laboratory conditions, concerns surround the site, method, and timing of specimen collection which can decrease the clinical validity in practice due to false negatives (Tang 2020). These issues may explain in part reports that infected people are retesting as positive after negative tests (Xiao 2020a; Xiao 2020b).

Overall completeness and applicability of the evidence

We note that the studies identified thus far have limited applicability to widespread population screening with the exception of the study conducted among a randomly sampled Icelandic population (Gudbjartsson 2020b), and in an Italian community (Lavezzo 2020). We anticipate that future population-based screening studies will add substantially to our understanding of the effectiveness and accuracy of screening for SARS-CoV-2.

Potential biases in the review process

Because of time constraints, we conducted a rapid review and abbreviated certain methodological steps of the review process. Specifically, we dually screened only excluded titles and abstracts, and did not independently rate the certainty of evidence.

Most importantly, we limited literature searches to English and Chinese languages, which could have missed studies in languages of countries that have been severely affected by the COVID-19 pandemic (e.g. Italy or Spain). Overall, however, we believe that the impact of these methodological shortcuts is likely to be minor and does not change the conclusions of our review.

Because of the dynamic nature of this field we also searched prepublication databases and included nine studies (Barrett 2020; Clifford 2020; Graham 2020; Gudbjartsson 2020b; Lavezzo 2020; Olalla 2020; Osterdahl 2020; Samuels 2020; Zhang 2020) that had not gone through peer review at the time of identification (some were eventually published in peer-reviewed journals). Other preprints awaiting peer review may change during the peer-review process.

AUTHORS' CONCLUSIONS

Implications for practice

Twenty-two studies provided evidence on the effectiveness and accuracy of single point-in-time and repeated screening. The evidence suggests that one-time screening approaches with a symptom assessment, direct temperature measurement, travel

history, assessment for exposure to known or suspected infected people, or combined symptoms assessment with temperature measurement may miss between 40% and 100% people who are infected, although the certainty of our conclusions ranges from very low to moderate. The limited accuracy of RT-PCR (Woloshin 2020), is also a constraint in confirming the results of various screening approaches. The limited utility of current screening approaches does not justify inaction: our findings point to the need for greater emphasis on other ways to prevent transmission, such as face coverings, physical distancing (Chu 2020), quarantine (Nussbaumer-Streit 2020), and adequate personal protective equipment for frontline workers (Verbeek 2020). When general screening for symptoms or temperature is to be deployed, community prevalence (when known to be accurate) may offer some insight on when screening may be most useful. In very low-prevalence settings, screening for symptoms or temperature, despite low overall accuracy, may result in few false negatives and many true negatives. Repeated screenings, in such settings, may result in more cases being identified eventually and reduced harm from false reassurance. In high-prevalence settings such as congregate living facilities (e.g. nursing homes, homeless shelters, dormitories, prisons), however, universal testing with RT-PCT may be a preferred strategy to screening.

Implications for future research

Regarding the effectiveness of screening, primary studies on the effectiveness of screening versus no screening or universal screening versus targeted screening could provide valuable evidence; evidence on harms of screening including false alarm and false reassurance would also be valuable. Future modelling studies need to base their assumptions on systematic reviews of the best available evidence regarding accuracy of tests and characteristics of SARS-CoV-2 infections.

Regarding the accuracy of screening, although many screening strategies evaluated in this review have inherent constraints on accuracy arising from the long latency period of the virus, emerging information on previously unreported symptoms and risk factors, and lack of clarity on the correct threshold for determining positive cases, future strategies may address one or more of these constraints. Screening strategies may be improved and then combined in ways to address individual deficiencies in current testing approaches. For example, if comprehensive and improved symptom assessment with or without other epidemiologic risk factors can identify positive cases accurately, it may identify cases that PCR tests might miss because of imperfect PCR specimen collection. This approach would require screening strategies combined with rapid or standard PCR tests, to identify a greater proportion of positive cases than either approach singly. Repeated rounds of combined screening (for symptoms, elevated temperatures, exposure to infected people, and other factors), and rapid testing strategies, particularly in somewhat or wholly captive populations ranging from workplaces, schools, and dormitories to nursing homes and prisons, may also help identify cases early and head off outbreaks. Further research is needed on more comprehensive symptom and risk assessment, rapid laboratory tests, and combinations of approaches. Finally, the limited accuracy of RT-PCR as a reference standard points to the need for evaluation of accuracy again for other more reliable reference standards as they are developed.

ACKNOWLEDGEMENTS

We thank Douglas Salzwedel for the peer review of the search strategy, and Roger Chou, Yemisi Takwoingi, Rob Scholten, Jennifer Lin, Jon Deeks and Ian Marshall for peer reviewing different versions of this manuscript. We also want to thank Xuan Hu for translating the Chinese language studies into English. We thank Loraine Monroe for her help with document preparation and Ravi Devarajan for his help with graphics.

The editorial process was managed by Cochrane's Editorial Service in collaboration with Cochrane Public Health, and we thank Helen Wakeford, Sarah Hodgkinson, Lisa Bero, Andy Anglemyer, Luke Wolfenden, Paul Garner and Jodie Doyle for their editorial comments. Thank you to Jane Cracknell for transferring the manuscript from Word to RevMan and to Denise Mitchell for copy-editing the manuscript.

REFERENCES

References to studies included in this review

Arons 2020 *{published data only}*

Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *New England Journal of Medicine* 2020;**24**(382):2081-90.

Baggett 2020 *{published data only}*

Baggett TP, Keyes H, Sporn N, Gaeta JM. Prevalence of SARS-CoV-2 infection in residents of a large homeless shelter in Boston. *JAMA* 2020;**323**(21):2191-2. [DOI: 10.1001/jama.2020.6887]

Barrett 2020 *{published data only}*

Barrett ES, Horton DB, Roy J, Gennaro ML, Brooks A, Tischfield J. Prevalence of SARS-CoV-2 infection in previously undiagnosed health care workers at the onset of the U.S. COVID-19 epidemic. *medRxiv [Preprint]* 2020. [DOI: /doi.org/10.1101/2020.04.20.20072470] [medRxiv 2020.04.20.20072470]

Clifford 2020 *{published data only}*

Clifford SJ, Pearson CA, Klepac P, Van Zandvoort K, Quilty BJ, Eggo RM, et al. Effectiveness of interventions targeting air travellers for delaying local outbreaks of SARS-CoV-2. *Journal of Travel Medicine* 2020;**27**(5):taaa068.

Gostic 2020 *{published data only}*

Gostic K, Gomez AC, Mummah RO, Kucharski AJ, Lloyd-Smith JO. Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19. *eLife* 2020;**9**:e55570. [DOI: 10.7554/eLife.55570] [PMID: 32091395]

Graham 2020 *{published data only}*

Graham NS, Junghans C, Downes R, Sendall C, Lai H, McKirdy A, et al. SARS-CoV-2 infection, clinical features and outcome of COVID-19 in United Kingdom nursing homes. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.05.19.20105460]

Gudbjartsson 2020a *{published data only}*

Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Early spread of SARS-Cov-2 in the Icelandic population. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.03.26.20044446]

Gudbjartsson 2020b *{published data only}*

* Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GI, et al. Spread of SARS-CoV-2 in the Icelandic population. *New England Journal of Medicine* 2020;**382**(24):2302-15. [DOI: 10.1056/NEJMoa2006100] [PMID: 32289214]

Guery 2020 *{published data only}*

Guery R, Cécile D, Noëlle B, Virginie N, Louise C, François R, et al. Limited effectiveness of systematic screening by nasopharyngeal RT-PCR of medicalized nursing home staff after a first case of COVID-19 in a resident. *Medecine et Maladies Infectieuses* 2020;**S0399-077X**(20):30126-8. [DOI: 10.1016/j.medmal.2020.04.020]

Hoehl 2020 *{published data only}*

Hoehl S, Rabenau H, Berger A, Kortenbusch M, Cinatl J, Bojkova D et al. Evidence of SARS-CoV-2 infection in returning travelers from Wuhan, China. *New England Journal of Medicine* 2020;**382**(13):1278-80. [DOI: 10.1056/NEJMc2001899] [PMID: 32069388]

Kimball 2020 *{published data only}*

Kimball A, Hatfield KM, Arons M, James A, Taylor J, Spicer K, et al, CDC Covid- Investigation Team. Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility - King County, Washington, March 2020. *MMWR - Morbidity & Mortality Weekly Report* 2020;**69**(13):377-81. [DOI: 10.15585/mmwr.mm6913e1] [PMID: 32240128]

Lavezzo 2020 *{published data only}*

Lavezzo E, Franchin E, Ciavarella C, Cuomo-Dannenburg G, Barzon L, Del Vecchio C, et al. Suppression of COVID-19 outbreak in the municipality of Vo, Italy. *medRxiv [Preprint]* 2020. [10.1101/2020.04.17.20053157]

Lytras 2020 *{published data only}*

Lytras T, Dellis G, Flountzi A, Hatzianastasiou S, Nikolopoulou G, Tsekou K, et al. High prevalence of SARS-CoV-2 infection in repatriation flights to Greece from three European countries. *Journal of Travel Medicine* 2020;**27**(3):taaa054. [DOI: 10.1093/jtm/taaa054] [PMID: 32297940]

Nishiura 2020 *{published data only}*

Nishiura H, Kobayashi T, Suzuki A, Jung SM, Hayashi K, Kinoshita R, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *International Journal of Infectious Diseases* 2020;**94**:154-5. [DOI: 10.1016/j.ijid.2020.03.020] [PMID: 32179137]

Olalla 2020 *{published data only}*

Olalla J, Correa AM, Martin-Escalante MD, Hortas ML, Martin-Sendarrubias MJ, Fuentes V, et al. Search for asymptomatic carriers of SARS-CoV-2 in healthcare workers during the pandemic: a Spanish experience. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.05.18.20103283]

Osterdahl 2020 *{published data only}*

Osterdahl MF, Lee KA, Ni Lochlainn M, Wilson S, Douthwaite S, Horsfall R, et al. Detecting SARS-CoV-2 at point of care: preliminary data comparing loop-mediated isothermal amplification (LAMP) to PCR. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.04.01.20047357]

Quilty 2020 *{published data only}*

Quilty BJ, Clifford S, Flasche S, Eggo RM, CMMID nCoV working group. Effectiveness of airport screening at detecting travellers infected with novel coronavirus (2019-nCoV). *Euro Surveillance* 2020;**25**(5):2000080. [DOI: 10.2807/1560-7917.ES.2020.25.5.2000080] [PMID: 32046816]

Roxby 2020 *{published data only}*

Roxby AC, Greninger AL, Hatfield KM, Lynch JB, Dellit TH, James A. Detection of SARS-CoV-2 among residents and staff members of an independent and assisted living community for older adults — Seattle, Washington, 2020. *MMWR - Morbidity & Mortality Weekly Report* 2020;**69**(14):416-8. [DOI: 10.15585/mmwr.mm6914e2] [PMID: 32271726]

Samuels 2020 *{published data only}*

Samuels EA, Karb R, Vanjani R, Trimbur MC, Napoli A. Congregate shelter characteristics and prevalence of asymptomatic SARS-CoV-2. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.05.21.20108985]

Sriwijitalai 2020 *{published data only}*

Sriwijitalai W, Wiwanitkit V. Positive screening for Wuhan novel coronavirus infection at international airport: what's the final diagnosis for positive cases. *International Journal of Preventive Medicine* 2020;**11**:30. [DOI: 10.4103/ijpvm.IJPVM_42_20] [PMID: 32175070]

Sutton 2020 *{published data only}*

Sutton D, Fuchs K, D'Alton M, Goffman D. Universal screening for SARS-CoV-2 in women admitted for delivery. *New England Journal of Medicine* 2020;**382**(22):2163-4. [DOI: 10.1056/NEJMc2009316] [PMID: 32283004]

Wells 2020 *{published data only}*

Wells CR, Sah P, Moghadas SM, Pandey A, Shoukat A, Wang Y, et al. Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. *Proceedings of the National Academy of Sciences of the United States of America* 2020;**117**(13):7504-9. [DOI: 10.1073/pnas.2002616117] [PMID: 32170017]

Zhang 2020 *{published data only}*

Zhang Y, Cheng S-R. Periodic COVID-19 testing in emergency department staff. *medRxiv [Preprint]* 2020. [DOI: 10.1101/2020.04.28.20084053]

Additional references

Brozek 2020

Brozek J, Canelo-Aybar C, Akl EA, Bowena JM, Bucher J, Chiuh WA. GRADE Guidelines 30: The GRADE approach to assessing the certainty of evidence from models – an overview in the context of health decision-making. *Journal of Clinical Epidemiology* [in press].

CDC 2020a

Centers for Disease Control and Prevention. Coronavirus Disease 2019 (COVID-19). Screening clients at entry to homeless shelters. www.cdc.gov/coronavirus/2019-ncov/community/homeless-shelters/screening-clients-respiratory-infection-symptoms.html (accessed prior to 4 September 2020).

CDC 2020b

Centers for Disease Control and Prevention. Coronavirus disease 2019 (COVID-19). Specimen collection. Interim guidelines for collecting, handling, and testing clinical specimens from persons for coronavirus disease 2019 (COVID-19). www.cdc.gov/coronavirus/2019-ncov/lab/guidelines-clinical-specimens.html (accessed 22 May 2020).

Chu 2020

Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schunemann HJ. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet Digital Health* 2020;**395**(10242):1973-87.

Cripps 2020

Cripps K. Airline passengers under COVID-19 blood tests before boarding. www.cnn.com/travel/article/emirates-passengers-blood-test-covid-19/index.html (accessed 16 April 2020).

Emirates 2020

Emirates. Emirates to implement thermal screening on all US bound flights. www.emirates.com/media-centre/emirates-to-implement-thermal-screening-on-all-us-bound-flights/ (accessed prior to 4 September 2020).

Evers 2005

Evers S, Goossens M, De Vet H, Van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: consensus on health economic criteria. *International Journal of Technology Assessment in Health Care* 2005;**21**(2):240-5. [PMID: 15921065]

Garrity 2020

Garritty C, Gartlehner G, Kamel C, King VJ, Nussbaumer-Streit B, Stevens A, et al. Cochrane Rapid Reviews. Interim Guidance from the Cochrane Rapid Reviews Methods Group. March 2020. covidreviews.cochrane.org/sites/covidreviews.cochrane.org/files/public/uploads/cochrane_rr_-_guidance-23mar2020-final.pdf.

GRADEpro GDT [Computer program]

McMaster University (developed by Evidence Prime) GRADEpro GDT. Version before 1 September 2020. Hamilton (ON): McMaster University (developed by Evidence Prime). Available at gradepro.org.

Lee 2020

Lee S, Kim T, Lee E, Lee C, Kim H, Rhee H, et al. Clinical course and molecular viral shedding among asymptomatic and symptomatic patients with SARS-CoV-2 infection in a community treatment center in the Republic of Korea. *JAMA Internal Medicine* 2020 August 6 [Epub ahead of print]. [DOI: 10.1001/jamainternmed.2020.3862]

McInnes 2018

McInnes MD, Moher D, Thombs BD, McGrath TA, Bossuyt PM, PRISMA-DTA Group. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018;**319**(4):388-96. [DOI: 10.1001/jama.2017.19163] [PMID: 29362800]

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and

meta-analyses: the PRISMA statement. *PLoS Medicine* 2009;**6**(7):e1000097. [DOI: 10.1371/journal.pmed.100009] [PMID: 19621072]

Nussbaumer-Streit 2020

Nussbaumer-Streit B, Mayr V, Dobrescu A, Chapman A, Persad E, Klerings I, et al. Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. *Cochrane Database of Systematic Reviews* 2020, Issue 4. Art. No: CD013574. [DOI: 10.1002/14651858.CD013574]

Pan 2020

Pan A, Liu L, Wang C, Guo H, Hao X, Wang Q, et al. Association of Public Health Interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. *JAMA* 2020;**323**(19):1-9. [DOI: 10.1001/jama.2020.6130] [PMID: 32275295]

Schünemann 2008

Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al, GRADE Working Group. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;**336**(7653):1106-10. [DOI: 10.1136/bmj.39500.677199.AE] [PMID: 18483053]

Tahamtan 2020

Tahamtan A, Ardebili A. Real-time RT-PCR in COVID-19 detection: issues affecting the results. *Expert Review of Molecular Diagnostics* 2020;**20**(5):453-4. [DOI: 10.1080/14737159.2020.1757437.] [PMID: 32297805]

Tang 2020

Tang Y-W, Schmitz JE, Persing DH, Stratton CW. Laboratory diagnosis of COVID-19: current issues and challenges. *Journal of Clinical Microbiology* 2020;**58**(6):e00512-20. [DOI: 10.1128/JCM.00512-20] [PMID: 32245835]

Verbeek 2020

Verbeek JH, Rajamaki B, Ijaz S, Sauni R, Toomey E, Blackwood B, et al. Personal protective equipment for preventing highly infectious diseases due to exposure to contaminated body fluids in healthcare staff. *Cochrane Database of Systematic Reviews* 2020, Issue 5. Art. No: CD011621. [DOI: 10.1002/14651858.CD011621.pub5]

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al, QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Anals of Internal Medicine* 2011;**155**(8):529-36. [DOI: 10.7326/0003-4819-155-8-201110180-00009] [PMID: 22007046]

WHO 2020a

World Health Organization. Report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf (accessed prior to 4 September 2020).

WHO 2020b

World Health Organization. Transmission of SARS-CoV-2: implications for infection prevention precautions. https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions (accessed prior to 4 September 2020).

WHO 2020c

World Health Organization. WHO announces COVID-19 outbreak a pandemic. www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic (accessed prior to 4 September 2020).

WHO 2020d

World Health Organization. Q&A on coronaviruses (COVID-19). who.int/news-room/q-a-detail/q-a-coronaviruses (accessed prior to 4 September 2020).

Wolfel 2020

Wolfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Muller MA, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* 2020;**581**(7809):465-9. [DOI: 10.1038/s41586-020-2196-x] [PMID: 32235945]

Woloshin 2020

Woloshin S, Patel N, Kesselheim AS. False negative tests for SARS-CoV-2 infection - challenges and implications. *New England Journal of Medicine* 2020;**381**(6):e38. [DOI: 10.1056/NEJMp2015897] [PMID: 32502334.]

Wu 2020

Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 2020;**323**(13):1239-42. [DOI: 110.1001/jama.2020.2648] [PMID: 32091533]

Xiao 2020a

Xiao AT, Tong YX, Gao C, Zhu L, Zhang YJ, Zhang S. Dynamic profile of RT-PCR findings from 301 COVID-19 patients in Wuhan, China: a descriptive study. *Journal of Clinical Virology* 2020;**127**:104346. [DOI: 10.1016/j.jcv.2020.104346] [PMID: 32361324]

Xiao 2020b

Xiao AT, Tong YX, Zhang S. False-negative of RT-PCR and prolonged nucleic acid conversion in COVID-19: rather than recurrence. *Journal of Medical Virology* 2020. [DOI: 10.1002/jmv.25855] [PMID: 32270882]

Xie 2020

Xie C, Lu J, Wu D, Zhang L, Zhao H, Rao B, et al. False negative rate of COVID-19 is eliminated by using nasal swab test. *Travel Medicine and Infectious Disease* 2020 April 11 [Epub ahead of print]. [DOI: 10.1016/j.tmaid.2020.101668] [PMID: 32283215]

Zhuang 2020

Zhuang Q, Liu S, Zhang X, Jiang W, Wang K, Wang S, et al. Surveillance and taxonomic analysis of the coronavirus dominant in pigeons in China. *Transboundary & Emerging Diseases* 2020;**00**:1-10. [DOI: 10.1111/tbed.13541] [PMID: 32163661]

References to other published versions of this review

Viswanathan 2020

Viswanathan M, Gartlehner G, Kahwati L, Jahn B, Giger K, Dobrescu A, et al. Sreening Asymptomatic Persons for Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) (Protocol). *[https://osf.io/cf639]* 8 April 2020.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

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

Arons 2020

Study characteristics	
Study design	DTA study
Objectives	To report on findings from symptom-based screening for SARS-CoV-2 infection
Population and setting	King County, Washington State, USA Skilled nursing facility in King County, Washington Total N: 76 Mean (SD) age: negative: 75.1 (10.9), positive: 80.7 (8.4) N (%) female: negative: 32 (60.4%), positive 16 (69.6%)
Screening strategies	Symptom assessment including typical and atypical signs and symptoms of COVID-19. Typical COVID-19 signs and symptoms include fever, cough, and shortness of breath; potential atypical symptoms assessed included sore throat, chills, increased confusion, rhinorrhoea or nasal congestion, myalgia, dizziness, malaise, headache, nausea, and diarrhea. Residents were categorized as asymptomatic (no symptoms or only stable chronic symptoms) or symptomatic (at least 1 new or worsened typical or atypical symptom of COVID-19) on the day of testing or during the preceding 14 days. Residents with positive test results and were asymptomatic at time of testing were reevaluated 1 week later to ascertain whether any symptoms had developed in the interim. Those who developed new symptoms were recategorized as presymptomatic
Notes	**Reference test** RT-PCR. Nasopharyngeal swabs were obtained from all 76 residents who agreed to testing and were present in the facility at the time; oropharyngeal swabs were also collected from most residents, depending upon their co-operation. The Washington State Public Health Laboratory performed one-step real-time RT-PCR assay on all specimens using the SARS-CoV-2 CDC assay protocol, which determines the presence of the virus through identification of 2 genetic markers, the N1 and N2 nucleocapsid protein gene regions

Baggett 2020

Study characteristics	
Study design	DTA study
Objectives	To screen homeless shelter residents for SARS-CoV-2 infection
Population and setting	Boston, USA Homeless shelter in Boston, MA, USA

Baggett 2020 *(Continued)*

	Total N: 408 Mean (SD) age: 51.6 (12.8) N (%) female: 115 (28.2%)
Screening strategies	Symptom assessment (cough, fever, or 'other' symptoms) Temperature (at least 100 °F or 37.8 °C) None
Notes	**Reference test** RT-PCR

Barrett 2020

Study characteristics	
Study design	DTA study
Objectives	To determine the baseline prevalence of SARS-CoV-2 infection in a cohort of previously undiagnosed healthcare workers and a comparison group of non-healthcare workers
Population and setting	New Jersey, USA University setting, including both healthcare and non-healthcare facilities Total N: 829 Mean (SD) age: mean age: NR. 428 (51.6%) between 20-39 years; 315 (38.0%) between 40-59 years; 86 (10.4%) ≥ 60 years N (%) female: 530 (63.9%)
Screening strategies	Symptom screening including fever, cough, shortness of breath, vomiting, diarrhoea, or change in smell or taste Recent exposure
Notes	**Reference test** RT-PCR

Clifford 2020

Study characteristics	
Study design	Transmission model
Objectives	To evaluate if interventions aimed at air travellers can delay establishment of a SARS-CoV-2 outbreak in a previously unaffected country
Population and setting	0.1/1/10/100 infected travellers per week % asymptomatic: not explicitly modelled

Clifford 2020 *(Continued)*

	Awareness of exposure: modelled as a result of sensitization
Screening strategies	No screening, departure-only, and departure-and-arrival screening in combination with traveller sensitization: 0%, and 50% effectiveness Syndromic screening; sensitivity: 86% NA; sensitivity: NA Follow-up interventions: sensitized travellers leading to increased likelihood of self-isolation and rapid care-seeking Time between screenings: 12 h Additional measures: traveller sensitization: 0%, 25%, 50%, sensitized travellers enter destination but cause fewer secondary cases because of an increased likelihood of self-isolation and rapid care-seeking
Notes	**Model parameters** Incubation period: mean 5.2 days Time from symptom to hospitalization: mean 9.2 days How dynamics of epidemic were considered: 'Upon arrival, all infected carry the risk of potential secondary infections, determined by the average number of those infections, R0 (1.4-3.9), and its dispersion (k = 0.54). R0 impacts outbreak threshold. travellers' sensitization leads to reduced R0 **Model code available** Code: github.com/cmmid/screening_outbreak_delay/

Gostic 2020

Study characteristics	
Study design	Mathematical model
Objectives	To assess the expected effectiveness of screening for COVID-19 • deriving the probability that an infected individual would be detected or missed (individual analysis) • following population of 30 infected travellers (individually sampled) (population analysis)
Population and setting	• Infected individual • Cohort of 30 infected individuals drawn from a 1000-candidate parameter set % asymptomatic: 5%/25%/50% Awareness of exposure: 20%
Screening strategies	Departure screening (test 1 followed by test 2) and/or arrival screening (test 1 followed by test 2) Screening for symptoms, infrared thermal scanners for fever; sensitivity: 70% Questionnaire-based screening; sensitivity: 25%, (0%-100% individual analysis) Follow-up interventions: none Time between screenings: 24-h flight Additional measures: NA
Notes	**Model parameters**

Gostic 2020 *(Continued)*

	Incubation period: 4.5 days/5.5 days/6.5 days (individual analysis) 5.5 days (gamma distribution with mean as above, and SD = 2.25) Time from symptom to hospitalization: 3-7 days time from symptom onset to patient isolation How dynamics of epidemic were considered: "- In a growing epidemic, the fraction of cases who are recently exposed increases with R0 (R0 (1.5 to 4.0)); - In a stable epidemic times since exposure follow a uniform distribution across the time period between exposure and isolation" No effect in individual-level analysis **Model code available** Code: https://github.com/kgostic/traveller_screening/releases/tag/v2.1, online app: https://lloyd-smithlab.shinyapps.io/travelScreeningModel/https://github.com/kgostic/traveller_screening/releases/tag/v2.1, online app: https://lloyd-smithlab.shinyapps.io/travelScreeningModel/https://github.com/kgostic/traveller_screening/releases/tag/v2.1, online app: https://lloyd-smithlab.shinyapps.io/travelScreeningModel/

Graham 2020

Study characteristics	
Study design	DTA study
Objectives	To understand SARS-CoV-2 infection and transmission in UK nursing homes in order to develop preventive strategies for protecting the frail elderly residents
Population and setting	London, UK 4 nursing homes affected by COVID-19 outbreaks in central London Total N: 313 residents screened (394 residents in total) Mean (SD) age: mean 83 years (entire population) N (%) female: 246 (62,4%) (numbers only for entire population of 394 residents!)
Screening strategies	Assessment of typical (new fever, cough and/or breathlessness) and atypical (newly altered mental status or behaviour, anorexia, diarrhoea or vomiting) symptoms of SARS-CoV-2-infection during the previous 2 weeks
Notes	**Reference test** RT-PCR (nasal/pharyngeal swab)

Gudbjartsson 2020a

Study characteristics	
Study design	
Objectives	
Population and setting	
Screening strategies	

Gudbjartsson 2020a *(Continued)*

Notes	

Gudbjartsson 2020b

Study characteristics	
Study design	DTA study
Objectives	To report findings from different strategies of screening for SARS-CoV-2 infection
Population and setting	Iceland Population recruited through online registration for testing between 13 March and 1 April, which was available to anyone including those without symptoms. In addition to open registration, randomly chosen Icelanders were sent text message invitations to participate in testing between 1 April and 4 April Total N: 13,080 Mean (SD) age: 39 (open enrolment); 45 (randomized invitation to enrolment) N (%) female: 7212 (55%)
Screening strategies	Recent international travel history Symptoms reported (fever, coughing, body ache, headache, sore throat, runny nose, fatigue, loss of smell, other) Known contact with infected person
Notes	**Reference test** Quantitative real-time PCR (qRT-PCR)

Guery 2020

Study characteristics	
Study design	DTA study
Objectives	To test all residents and staff members upon occurrence of first confirmed case of COVID-19 in a medicalized nursing home
Population and setting	Nantes, France Staff members at a medicalized nursing home associated with a university hospital Total N: 136 Mean (SD) age: median (IQR): 39 (27 to 48.5) N (%) female: 112 (82%)
Screening strategies	Symptom screening (included rhinitis, headache, cough, major fatigue, myalgia, diarrhea, ageusia/dysgeusia, chest pain, abdominal pain, odynophagia, dyspnea, anosmia, fever)
Notes	**Reference test**

Guery 2020 *(Continued)*

	Nasopharyngeal RT-PCR

Hoehl 2020

Study characteristics	
Study design	DTA study
Objectives	To report on findings from symptom-based screening for SARS-CoV-2 infection
Population and setting	Expatriates from Hubei, China to Germany For screening for symptoms and potential exposure, the study looked at travellers returning home from China to Germany. For temperature screening, the sample was described as those who were asymptomatic on the aeroplane and had no known or suspected contact with infected people. Total N: 125 Mean (SD) age: 5 months to 68 years N (%) female: NR
Screening strategies	• symptoms and temperature check • Temperature check • Exposure to known infected person • Exposure to known or suspected infection people
Notes	**Reference test** RT-PCR (confirmed by two commercial sets: LightMix Modular SARS and Wuhan CoV E-gene, and LightMix Modular Wuhan CoV RdRP-gene, both produced by TIB MOLBIOL)

Kimball 2020

Study characteristics	
Study design	DTA study
Objectives	To report on finding from symptom-based screening for SARS-CoV-2 infection
Population and setting	King County, Washington State, USA Skilled nursing facility in King County, Washington Total N: 76 Mean (SD) age: negative: 75.1 (10.9), positive: 80.7 (8.4) N (%) female: negative: 32 (60.4%), positive 16 (69.6%)
Screening strategies	Symptom assessment including typical and atypical signs and symptoms of COVID-19. Typical COVID-19 signs and symptoms include fever, cough, and shortness of breath; potential atypical symptoms assessed included sore throat, chills, increased confusion, rhinorrhoea or nasal congestion, myalgia, dizziness, malaise, headache, nausea, and diarrhoea. Residents were categorized as asymptomatic (no symptoms or only stable chronic symptoms) or symptomatic (at least one new or worsened typical or atypical symptom of COVID-19) on the day of testing or during the preceding 14 days. Residents with

Kimball 2020 *(Continued)*

	positive test results and were asymptomatic at time of testing were reevaluated 1 week later to ascertain whether any symptoms had developed in the interim. Those who developed new symptoms were recategorized as presymptomatic
Notes	**Reference test** RT-PCR. Nasopharyngeal swabs were obtained from all 76 residents who agreed to testing and were present in the facility at the time; oropharyngeal swabs were also collected from most residents, depending upon their cooperation. The Washington State Public Health Laboratory performed one-step real-time RT-PCR assay on all specimens using the SARS-CoV-2 CDC assay protocol, which determines the presence of the virus through identification of 2 genetic markers, the N1 and N2 nucleocapsid protein gene regions

Lavezzo 2020

Study characteristics	
Study design	DTA study
Objectives	
Population and setting	Vo', Italy Municipality of Vo', Italy (population 3275 individuals) Total N: 2812 Mean (SD) age: NR N (%) female: NR
Screening strategies	Symptom assessment, including fever, cough, headache, sore throat, discomfort, conjunctivitis, diarrhoea
Notes	**Reference test** RT-PCR

Lytras 2020

Study characteristics	
Study design	DTA study
Objectives	To estimate the prevalence of SARS-CoV-2 infection among repatriated citizens
Population and setting	Athens, Greece Airport in Athen's Greece Total N: 783 Mean (SD) age: median 27 (IQR 22-40) N (%) female: NR

Lytras 2020 *(Continued)*

Screening strategies	Symptom screening at time of arrival
Notes	**Reference test** RT-PCR

Nishiura 2020

Study characteristics	
Study design	DTA study
Objectives	To estimate the asymptomatic ratio to improve understanding of COVID-19 transmission and the spectrum of disease it causes
Population and setting	Japan Information was used on Japanese nationals who were evacuated from Wuhan, China on chartered flights Total N: 565 Mean (SD) age: NR N (%) female: NR
Screening strategies	Temperature screening before disembarkation and face-to-face interviews eliciting information on symptoms including fever, cough, and other non-specific symptoms consistent with COVID-19
Notes	**Reference test** RT-PCR

Olalla 2020

Study characteristics	
Study design	DTA study
Objectives	Determine the percentage of healthcare workers carrying SARS-CoV-2 infection in high exposure areas of the hospital
Population and setting	Andalusia, Spain Hospital, specifically with higher exposure risk, including the ED, ICU, internal medicine, pulmonology, and other medicine specialties Total N: 498 Mean (SD) age: 41.5 (NR) N (%) female: 354 (reported: 80%, calculated: 71.1%)
Screening strategies	Symptoms on the day of testing (fever, cough, shortness of breath, runny nose, sore throat, diarrhoea, vomiting, headache, myalgia, general malaise)

Olalla 2020 *(Continued)*

	Symptoms in the past 14 days (fever, cough, shortness of breath, runny nose, sore throat, diarrhoea, vomiting, headache, myalgia, general malaise)
Notes	**Reference test** RT-PCR

Osterdahl 2020

Study characteristics	
Study design	DTA study
Objectives	To present 'real world data' evaluating RT-LAMP compared to the current gold standard of RT-PCR
Population and setting	National Health Service high-dependency care home, UK High-dependency care home where an outbreak was suspected Total N: 21 Mean (SD) age: median 76, interquartile range 61-81 years N (%) female: 17 (70%)
Screening strategies	Point-of-care testing with RT-LAMP using swabs from the nose and throat The RT-LAMP method employed was the MicrosensDx RapiPrep SARS-CoV-2 research use test. The method used magnetic bead capture to maximize the yield of target nucleic acid during sample preparation from the dry swab, which is followed by optimised RT-LAMP to amplify and detect the SARS-CoV-2 genome, targeting the ORF1a gene. The assay runs under isothermal conditions, which can yield results in 25 minutes on average
Notes	**Reference test** Tandem RT-PCR based on 2 swabs (pharyngeal and deep nasal swabs) collected over 2 days, at least 1 positive test was considered 'positive'

Quilty 2020

Study characteristics	
Study design	Mathematical model
Objectives	To evaluate the effectiveness of exit and arrival screening for detecting travellers entering Europe with 2019-nCoV infection
Population and setting	100 2019-nCoV-infected travellers planning to board a flight, (quote:) "starting travel is randomly and uniformly distributed between the time of infection" and twice the expected time to severe disease % asymptomatic: 17% Awareness of exposure: NA
Screening strategies	Arrival test (test 1) and departure test (test 1) Thermal imaging scanner; sensitivity: 86%

Quilty 2020 *(Continued)*

	NA; sensitivity: NA Follow-up interventions: none Time between screenings: 12-h flight + 1 h Additional measures: NA
Notes	**Model parameters** Incubation period: mean 5.2 days, variance 4.1 days Time from symptom to hospitalization: mean 9.1 days, variance 14.7 days How dynamics of epidemic were considered: NA **Model code no longer available online:** app cmmid-lshtm.shinyapps.io/traveller_screening/

Roxby 2020

Study characteristics	
Study design	DTA study
Objectives	To report on findings from symptom-based screening for SARS-CoV-2 infection
Population and setting	Seattle, Washington, USA Senior independent and assisted-living community consisting of 83 apartments (45 independent living and 38 assisted living) Total N: 142 Mean (SD) age: residents: 86 years (range 69-102 years), staff: 40 years (range 16-70 years) N (%) female: residents: 61 (77%), staff: 45 (72%)
Screening strategies	Questionnare assessing fever, cough, and other symptoms during the preceding 14 days
Notes	**Reference test** Real time RT-PCR

Samuels 2020

Study characteristics	
Study design	DTA study
Objectives	To describe the varying prevalence of asymptomatic SARS-CoV-2 infection in congregate shelters and associated shelter characteristics and practices
Population and setting	Providence Metro Region, Rhode Island, USA Five shelters in the Providence Metro Region, from April 19-April 24, 2020 Total N: 299

Samuels 2020 *(Continued)*

	Mean (SD) age: mean 47.9 (range 18-85) N (%) female: 59 (20%)
Screening strategies	Symptom (described as COVID-19 symptoms but not further described) Temperature (> 100.4 °F/38 °C)
Notes	**Reference test** Nasopharyngeal swab for RT-PCR

Sriwijitalai 2020

Study characteristics	
Study design	DTA study
Objectives	To report on findings from temperature scanning and clinical history-based screening for SARS-CoV-2 infection
Population and setting	Thailand Flights (58) in Thailand connect to Wuhan China. Screening was done at international airports. Total N: NR Mean (SD) age: NR N (%) female: NR
Screening strategies	Body temperature scanning and clinical history taking, no additional details provided
Notes	**Reference test** Molecular diagnostic test, unspecified

Sutton 2020

Study characteristics	
Study design	DTA study
Objectives	To report findings of universal screening for SARS-CoV-2 infection
Population and setting	New York City, USA Women presenting for delivery at New York–Presbyterian Allen Hospital and Columbia University Irving Medical Center Obstetric hospital wing Total N: 214 Mean (SD) age: NR N (%) female: 100%

Sutton 2020 *(Continued)*

Screening strategies	Symptom and temperature screen
Notes	**Reference test** RT-PCR, nasopharyngeal swabs

Wells 2020

Study characteristics	
Study design	Statistical prediction model
Objectives	To evaluate the impact of travel restrictions and border control measures and the role of the airport travel network
Population and setting	People travelling from China % asymptomatic: not specified Awareness of exposure: not specified
Screening strategies	Border control and travel restrictions Not specified; sensitivity: not specified Health questionnaire for self-identification of exposure risk; sensitivity: NA Follow-up interventions: NA Time between screenings: not specified Additional measures: travel lockdown, quarantine
Notes	**Model parameters** Incubation period: mean 5.2, 12 days for some individuals Time from symptom to hospitalization: in scenario assumed cases travel up to the point of hospitalization (between 9.1 days and 12.5 days after symptom onset) How dynamics of epidemic were considered: (quote) "Prediction model used reported data from current COVID-19 epidemics to reproduce number of exported cases to given countries" **Model code available** Code: github.com/WellsRC/Coronavirus-2019

Zhang 2020

Study characteristics	
Study design	Transmission model with Markov chain to track infection timeline of healthcare workers
Objectives	To provide a quantitative analysis and model for predicting the impact of periodic COVID-19 testing for all asymptomatic ED staff as a possible strategy to mitigate disease transmission in the healthcare setting

Zhang 2020 *(Continued)*

Population and setting	Healthcare workers (307 at Harborview Medical Center, 0.41 infected at simulation start) % asymptomatic: 0.13% healthcare workers at simulation start (presymptomatic only) Awareness of exposure: NA
Screening strategies	• No screening • Weekly screening • Biweekly screening of all asymptomatic healthcare workers not specified; sensitivity: 100% for all tests; sensitivity: NA Follow-up interventions: healthcare workers removed from workforce Time between screenings: depending on screening strategy 1) one week, 3) two Additional measures: healthcare workers removed from workforce
Notes	**Model parameters** Incubation period: 4 days without symptoms Time from symptom to hospitalization: NA How dynamics of epidemic were considered: ED with 157.56 visits from patients from King Country, WA per day with 0.21 infected patients per day at start of simulation. A transmission constant (1.219e-4, 3.660e-4, 4.067e-5 new infections/person) is applied to simulate dynamic over time. **Model code available**: no

COVID-19: coronavirus disease 2019; **DTA:** diagnostic test accuracy; **ED:** emergency department; **ICU:** intensive care unit; **IQR:** interquartile range; **N:** number; **NA:** not applicable; **NR:** not reported; **PCR:** polymerase chain reaction; **R0:** reproduction number; **RT-LAMP:** reverse transcriptase loop-mediated isothermal amplification; **RT-PCR:** reverse transcription polymerase chain reaction; **SARS-CoV-2:** severe acute respiratory syndrome coronavirus-2; **SD:** standard deviation

ADDITIONAL TABLES

Table 1. Study eligibility criteria for effectiveness of universal screening

Criteria	Inclusion	Exclusion
Study design	RCTs	Case series
	Quasi-RCTs	Case reports
	Non-RCTs	Diagnostic case-control studies
	Prospective cohort studies	Systematic reviews (used for reference list checking)
	Retrospective cohort studies	
	Case-control studies	
	Cross-sectional studies	
	Controlled before-and-after studies	
	Modelling studies	
	Designs with data on screening accuracy	
Minimum duration	All	None

Table 1. Study eligibility criteria for effectiveness of universal screening *(Continued)*

Population	General populations without known SARS-CoV-2 infection Special populations of interest are healthcare workers without known SARS-CoV-2 infection Studies with a subset of relevant participants if more than 80% meet review inclusion criteria	Studies focusing only on diagnostic testing among people with suspected or presumptive COVID-19 symptoms or known exposure to SARS-CoV-2 (except for healthcare workers) Studies with a subset of relevant participants with < 80% that meet review inclusion criteria
Intervention(s)	Screening using the above approaches across the entire population or selected segments of the population based on occupation, geographic setting, or community characteristics	Other interventions
Comparator(s)	No screening, screening in selected populations based on occupation, geographic setting, or community characteristics	Other comparators
Outcome(s)	Incident cases, missed cases, successfully detected cases, averted cases, reduced transmission, mortality, false alarm, false reassurance	Other outcomes

COVID-19: coronavirus disease 2019; **CRISPR:** clustered regularly interspaced short palindromic repeats; **PCR:** polymerase chain reaction; **RCT:** randomized controlled trial; **RT-PCR:** reverse transcription polymerase chain reaction; **SARS-CoV-2:** severe acute respiratory syndrome coronavirus-2

Table 2. Study eligibility criteria for accuracy of universal screening

Criteria	Inclusion	Exclusion
Study design	Any design with data on screening accuracy modelling studies	Case series Case reports Diagnostic case-control studies Systematic reviews (used for reference list checking)
Minimum duration	Maximum duration between index test and reference standard is < 1 day for diagnostic test accuracy studies	Duration between index and reference test ≥ 1 day
Population	General populations without known SARS-CoV-2 infection Special populations of interest are healthcare workers without known SARS-CoV-2 infection. Studies with a subset of relevant participants if more than 80% meet review inclusion criteria	Studies focusing only on diagnostic testing among people with suspected or presumptive COVID-19 symptoms or known exposure to SARS-CoV-2 (except for healthcare workers). Studies with a subset of relevant participants with < 80% that meet review inclusion criteria
Target condition	Infection with SARS-CoV-2	Other conditions

Table 2. Study eligibility criteria for accuracy of universal screening *(Continued)*

Index tests	Clinical symptoms (e.g. fever, dry cough, shortness of breath, gastrointestinal symptoms, loss of smell or taste), epidemiological risk assessment (history of travel from hot spots, contact with people known or suspected of having SARS-CoV-2 infection), body temperature checks, feasible point-of-care laboratory tests (e.g. rapid PCR test, CRISPR), or a combination of these approaches	Other index tests, including rapid antibody immunoassay tests
Reference standard	Confirmatory diagnostic tests (RT-PCR or antibody immunoassay tests, CRISPR, or a combination of these approaches)	Other reference standards
Outcome(s)	Sensitivity, specificity, positive predictive value, negative predictive value, area under the receiver operating characteristic curve	Other outcomes

COVID-19: coronavirus disease 2019; **CRISPR:** clustered regularly interspaced short palindromic repeats; **PCR:** polymerase chain reaction; **RCT:** randomized controlled trial; **RT-PCR:** reverse transcription polymerase chain reaction; **SARS-CoV-2:** severe acute respiratory syndrome coronavirus-2

Table 3. Study characteristics and results from screening studies assessing effectiveness

Study	Type of model	Objectives	Data source, setting, n	Interventions	Quality
Clifford 2020	Transmission model, R_0 = 1.4-3.9; time from symptom to hospitalization 9.2 days (mean); incubation period 5.2 days (mean)	To evaluate if interventions aimed at air travellers can delay a SARS-CoV-2 outbreak in a previously unaffected country	NR 0.1/1/10/100 infected travellers per week n = NR	• no screening • departure-only screening • departure-and-arrival screening in combination with traveller sensitization Interventions were considered in combination with traveller sensitization	Minor or no concerns
Zhang 2020	Transmission model with Markov chain R_0 = NR; incubation period 4 days; time from symptom to hospitalization NR	To provide a quantitative analysis and model for predicting the impact of periodic COVID-19 testing for all asymptomatic ED staff as a possible strategy to mitigate disease transmission in the healthcare setting	Healthcare workers from Harborview Medical Center, USA n = 307	• no screening • weekly screening • biweekly screening of all asymptomatic healthcare workers	Major concerns

COVID-19: coronavirus disease 2019; **ED:** emergency department; **NR:** not reported; **R0:** reproduction number; **SARS-CoV-2:** severe acute respiratory syndrome coronavirus-2

Table 4. Certainty of evidence: does screening for symptoms of air travellers reduce days to outbreak of SARS-CoV-2 in unaffected populations

Certainty assessment							Impact	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication bias**		
1 (Clifford 2020)	Modelling study	Not serious	Not serious	Very serious[a]	Not serious	None	Assuming that travellers are sensitized to symptoms by the process of screening and then self-quarantine with symptoms, and also assuming only one infected traveller per week, travel screening can avert an outbreak by 83 days (50% of simulations; 75%: 36 days, 97.5%: 8 days) However, relaxing either or both assumptions reduces days averted to > 1-8 days	⨁⨁◯◯ Low

SARS-CoV-2: severe acute respiratory syndrome coronavirus-2

[a]Assumes sensitivity of symptom screening is 86% which is substantially higher than results from diagnostic accuracy studies; also assumes 100% adherence to self-quarantine.

Table 5. Certainty of evidence: does laboratory screening of healthcare workers reduce SARS-CoV-2 transmission to patients and other healthcare workers?

Certainty assessment							Impact	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication bias**		
1 (Zhang 2020)	Modelling study	Very serious[a]	Not serious	Very serious[b]	Not serious	None	Assuming a transmission rate of 1.2 new infections per 10,000 people, weekly screening could reduce patient and healthcare worker infections in emergency rooms by 5.1% within 30 days. Biweekly screening would reduce infections by 2.3%. Assuming a transmission rate of 3.7 new infections per 10,000 people, the reduction of infections for weekly screening was 21.1% (for healthcare workers and patients) and 9.8% for patients (9.7% for healthcare workers) for biweekly screening within 30 days	⨁◯◯◯ Very low

SARS-CoV-2: severe acute respiratory syndrome coronavirus-2

[a]No sensitivity analyses, assumes that sensitivity of screening is 100%; no information on recovery rates.
[b]Assumes that an emergency room has the same infection spread as a cruise ship (Diamond Princess).

Table 6. Summary of results from screening test accuracy studies[a]

Screening strategy	№ of publications and cohorts	№ of participants	Prevalence	Sensitivity	GRADE certainty for sensitivity	Specificity	GRADE certainty for specificity
Symptom assessment	13/12[b]	16,762	0.4% to 40.3%	0.00 to 0.60	Very low	0.66 to 1.00	Low
Repeated symptom assessment	1/1	76	62%	0.44	Very low	0.62	Low
Direct temperature measurement	3/3	822	1.7% to 36.0%	0.00 to 0.01	Low	0.99 to 1.00	Moderate
Asking about international travel history	1/2	13080	0.6% to 0.8%	0.00 to 0.23	Very low	0.91 to 1.00	Low
Asking about exposure to known infected people	2/3	13205	0.6% to 1.6%	0.00 to 0.08	Low	0.97 to 0.98	Moderate
Asking about exposure to known or suspected infected people	2/2	654	1.6% to 4.9%	0.00 to 0.17	Low	0.90 to 0.98	Moderate
Symptom assessment and direct temperature measurement	2/2	779	2.3% to 15.5%	0.12 to 0.69	Very low	0.90 to 1.00	Low
RT-LAMP	1/1	21	47.6%	0.80	Very low	0.73	Very low

RT-LAMP: reverse transcriptase loop mediated isothermal amplification (a rapid, point of care viral detection method)

[a]all studies compared against the RT-PCR as a reference standard, only data from primary research studies included in this table.
[b]Estimates include the high risk-of-bias study (Barrett 2020), and the study without specificity estimates (Lavezzo 2020). Excluding these studies does not change the range of estimates. Sensitivity in Barrett 2020. is 0.34 and specificity is 0.89. Sensitivity in Lavezzo 2020 is 0.59.

Table 7. Study characteristics and results from screening test accuracy studies[a]

Screening strategy	Study	Population	Time period	Screening test characteristics	Prevalence (%)	TP	TN	FP	FN	Sensitivity	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)

Table 7. Study characteristics and results from screening test accuracy studies[a] *(Continued)*

					thresh-old				(95% CI)			
Single screen: symptoms	Baggett 2020	People experiencing homelessness in Boston, Massachusetts, USA, congregate shelters	2-3 April 2020	Residents were asked about history of cough and shortness of breath and were given the option to report other symptoms	Any symptoms	147/408 (36.0%)	18	232	29	129	0.12 (0.07 to 0.18)	0.89 (0.84 to 0.92)	0.38 (0.26 to 0.52)	0.64 (0.63 to 0.66)
	Barrett 2020	University workers, predominantly (67%) healthcare workers in USA	24 March-7 April 2020	Questionnaire on recent symptoms including fever, cough, shortness of breath, vomiting, diarrhoea, or change in smell or taste	Any COVID-19 symptoms in last week	41/829 (4.9%)	14	704	84	27	0.34 (0.20 to 0.51)	0.89 (0.87 to 0.91)	0.14 (0.09 to 0.21)	0.96 (0.95 to 0.97)
	Graham 2020	Residents and staff in nursing homes in the UK	Initiated on 15 April 2020	Case note review and information from the medical and nursing team on 5 key symptoms: typical (new fever, cough and/or breathlessness) and atypical (newly altered mental status or behaviour, anorexia, diarrhoea or vomiting) features of COVID-19	Any symptoms	126/313 (40.3%)	72	150	37	54	0.57 (0.48 to 0.66)	0.80 (0.74 to 0.86)	0.66 (0.58 to 0.73)	0.74 (0.69 to 0.77)
	Gudbjartsson 2020b	Open invitation to screen to people living in Iceland	13 March-1 April 2020	Questionnaire regarding symptoms including fever, cough, body aches, headaches, sore throat, rhinorrhoea, other (excludes fatigue and loss of smell or taste)	Any symptoms	87/10,797 (0.8%)	51	7182	3528	36	0.59 (0.48 to 0.69)	0.67 (0.66 to 0.68)	0.01 (0.01 to 0.02)	0.99 (0.99 to 1.00)
	Gudbjartsson 2020b	Random-sample population screening of people living in Iceland	1-4 April 2020	Questionnaire regarding symptoms including fever, cough, body aches, headaches, sore throat, rhinorrhoea, fatigue, loss of smell or taste, other	Any symptoms	13/2283 (0.6%)	6	2005	265	7	0.46 (0.19 to 0.75)	0.88 (0.87 to 0.90)	0.02 (0.01 to 0.04)	0.99 (0.99 to 1.00)
	Guery 2020	Staff members, health workers, and administrative personnel in a nursing home in France	16-17 April 2020	Standardized symptom assessment form (rhinitis, headache cough, major fatigue, myalgia, diarrhoea, ageusia/dysgeusia, chest pain, abdominal pain, odynophagia, dyspnea, anosmia, fever)	Any symptom	3/136 (2.2%)	1	96	37	2	0.33 (0.008 to 0.91)	0.72 (0.64 to 0.80)	0.03 (0.005 to 0.12)	0.98 (0.96 to 0.99)

Table 7. Study characteristics and results from screening test accuracy studies[a] *(Continued)*

Hoehl 2020	Evacuees to Frankfurt, Germany from Hubei, China	1 February 2020	Evaluation of self-reported fever, fatigue, sore throat, cough, runny nose, muscle aches, and diarrhoea; evaluation of signs of infection in the nose and throat	2/125 (1.6%) symptom	0	116	7	2	0.00 (0.00 to 0.84)	0.94 (0.87 to 0.98)	0	0.98 (0.98 to 0.98)
Kimball 2020	Long-term care skilled nursing facility in King County, Washington, USA	13 March 2020	Symptom assessment tool completed by facility nursing staff members by reviewing screening records of residents for the preceding 14 days and by clinician interview of residents for typical or atypical symptoms on the day of testing or during the preceding 14 days. Typical symptoms include subjective fever of temperature > 100.4 °F (38 °C), cough, and shortness of breath; atypical symptoms include sore throat, chills, increased confusion, rhinorrhoea or nasal congestion, myalgia, dizziness, malaise, headache, nausea, and diarrhoea	23/76 (30.3%) 1 new or worsened typical or atypical symptom of COVID-19) vs no symptoms or only stable chronic symptoms	10	40	13	13	0.44 (0.23 to 0.66)	0.75 (0.62 to 0.86)	0.44 (0.28 to 0.60)	0.75 (0.68 to 0.82)
Lavezzo 2020	Municipality of Vo', Italy	After 21 February (date NR)	Collected information on clinical symptoms, including fever, cough, headache, sore throat, discomfort, conjunctivitis, and diarrhoea	NR/2812 (2.6%)	43	NR	NR	30	0.59	NA	NA	NA

Table 7. Study characteristics and results from screening test accuracy studies[a] *(Continued)*

	Lytras 2020	Travellers to Greece from UK, Spain and Turkey	20-25 March 2020	Passengers asked to fill out paper form with clinical, demographic, and contact information	40/783 (5.1%)	1	743	0	39	0.03 (0.006 to 0.13)	1.00 (0.995 to 1.00)	0.95 (0.95 to 0.95)	0.95 (0.93 to 0.96)
	Olalla 2020	Doctors, nurses, nursing assistants, security guards, administrative and cleaning staff in high COVID exposure areas of the hospital in Spain	15-24 April 2020	Symptoms on the same day of the extraction and similar symptoms in the previous 14 days (fever, cough, shortness of breath, runny nose, sore throat, diarrhoea, vomiting, headache, myalgia and feeling of general malaise	2/498 (0.4%) symptoms	0	441	55	2	0.00 (0.00 to 0.84)	0.89 (0.86 to 0.91)	0	0.996 (0.995 to 0.996)
	Roxby 2020	Senior independent and assisted-living community consisting of 83 apartments in Seattle, Washington, USA	10-11 March 2020	Questionnaire assessing fever, cough and other symptoms during the preceding14 days	5/142 (3.5%)	3	90	47	2	0.60 (0.15 to 0.95)	0.66 (0.57 to 0.74)	0.06 (0.03 to 0.12)	0.98 (0.94 to 0.99)
	Samuels 2020	People experiencing homelessness in Providence, Rhode Island, USA, congregate shelters	19-24 April 2020	COVID-19 symptoms (fever, cough, shortness of breath, body aches, gastrointestinal upset, loss of taste or smell	25/299 (11.7%) symptoms	7	227	37	28	0.20 (0.08 to 0.37)	0.86 (0.81 to 0.90)	0.16 (0.08 to 0.28)	0.89 (0.87 to 0.91)
Cumulative repeated screening: symptoms	Arons 2020	Long-term care skilled nursing facility in King County, Washington, USA, repeated screening of 49 tested negative in first point-prevalence screening 1 week earlier (76 participants; Kimball 2020)	19-20 March 2020	Symptom assessment tool completed by facility nursing staff members by reviewing screening records of residents for the preceding 14 days and by clinician interview of residents for typical or atypical symptoms on the day of testing or during the preceding 14 days, Typical symptoms include subjective fever of temperature > 100.4 °F (38 °C), cough, and shortness of breath; atypical symptoms include sore throat, chills, increased confusion, rhinorrhoea or nasal congestion, myalgia, dizziness, malaise, headache, nausea, and diarrhoea	48/76 (63.1%) 1 new or worsened typical or atypical symptom of COV-	21	18[b]	11	27	0.44 (0.29 to 0.59)	0.62 (0.42 to 0.79)	0.65 (0.52 to 0.77)	0.40 (0.31 to 0.49)

Table 7. Study characteristics and results from screening test accuracy studies[a] *(Continued)*

					ID-19) vs no symptoms or only stable chronic symptoms								
Single screen: temperature	Baggett 2020	People experiencing homelessness in Boston, Massachusetts, USA, congregate shelters	2-3 April 2020	Oral thermometers	147/408 (36.0%)	1	258	3	146	0.007 (0.00 to 0.037)	0.99 (0.97 to 1.00)	0.25 (0.03 to 0.76)	0.64 (0.63 to 0.64)
	Hoehl 2020	Evacuees to Frankfurt, Germany from Hubei, China	1 February 2020	NR	2/15 (13.3%)	0	112	1	2	0.00 (0.00 to 0.84)	0.99 (0.95 to 1.00)	0	0.98 (0.98 to 0.98)
	Samuels 2020	People experiencing homelessness in Providence, Rhode Island, USA, congregate shelters	19-24 April 2020	NR	35/299 (11.7%)	0	264	0	35	0.00 (0 to 0.10)	1.00 (0.99 to 1.00)	NA	0.88 (0.88 to 0.88)
Single screen: International travel	Gudbjartsson 2020b	Open invitation to screen to people living in Iceland	13 March-1 April 2020	Questionnaire regarding recent travels	47/10,792 (0.4%)	20	9791	919	67	0.23 (0.15 to 0.33)	0.91 (0.91 to 0.92)	0.02 (0.01 to 0.03)	0.99 (0.99 to 0.99)

Table 7. Study characteristics and results from screening test accuracy studies[a] *(Continued)*

history	Gudbjartsson 2020b	Random-sample population screening of people living in Iceland	1-4 April 2020	Questionnaire regarding recent travels	Any international travel	13/2283 (0.6%)	0	2259	11	13	0.00 (0.00 to 0.25)	0.99 (0.99 to 1.00)	0	0.99 (0.99 to 0.99)
Single screen: exposure to known infected people	Gudbjartsson 2020b	Open invitation to screen to people living in Iceland	13 March-1 April 2020	Questionnaire regarding known contact with infected people	Known contact with infected people	87/10,797 (0.8%)	6	10435	275	81	0.07 (0.036 to 0.14)	0.97 (0.97 to 0.1.00)	0.02 (0.01 to 0.05)	0.99 (0.99 to 0.99)
	Gudbjartsson 2020b	Random-sample population screening of people living in Iceland	1-4 April 2020	Questionnaire regarding known contact with infected people	Known contact with infected people	13/2283 (0.6%)	1	2211	59	12	0.08 (0.00 to 0.36)	0.97 (0.97 to 0.98)	0.02 (0.00 to 0.10)	0.99 (0.99 to 0.99)
	Hoehl 2020	Evacuees to Frankfurt, Germany from Hubei, China	1 February 2020	Evaluated for known contact with people with SARS-CoV-2	Any exposure	2/125 (1.6%)	0	121	2	2	0.00 (0.00 to 0.84)	0.98 (0.84 to 1.00)	0	0.98 (0.98 to 0.98)
Single screen: recent exposure to known or	Barrett 2020	University workers, predominantly (67%) healthcare workers in USA	24 March-7 April 2020	Questionnaire regarding potential exposure	Any exposure outside of work to some-	41/829 (4.9%)	7	707	81	34	0.17 (0.07 to 0.32)	0.90 (0.87 to 0.92)	0.08 (0.04 to 0.15)	0.95 (0.95 to 0.96)

Table 7. Study characteristics and results from screening test accuracy studies[a] *(Continued)*

suspected infected people					one with COVID-19 or new fever, cough, or shortness of breath								
	Hoehl 2020	Evacuees to Frankfurt, Germany from Hubei, China	1 February 2020	Evaluated for known contact with people with SARS-CoV-2 or accompanied people with SARS-CoV-2 infection or other health problems	2/125 (1.6%)	0	119	4	2	0.00 (0.00 to 0.84)	0.97 (0.92 to 0.99)	0	0.98 (0.98 to 0.98)
Single screen: RT-LAMP	Osterdahl 2020	National Health Service high-dependency care home, UK: cumulative results from single test for 21 participants in days 1 and 3, and repeated testing for 15 negative participants on day 4	16-20 March 2020	Deep nasal swabs transported in dry sterile tubes, magnetic bead capture to maximize the yield of target nucleic acid, followed by RT-LAMP	NR/21 (47.6%)	8	8	3	2	0.80 (0.44 to 0.96)	0.73 (0.39 to 0.94)	0.73 (0.39 to 0.94)	0.80 (0.44 to 0.98)
Single screen: symptoms plus temperature check	Nishiura 2020	Evacuees to Japan from Wuhan, China	By 6 February 2020 (start date NR)	Temperature screening before disembarkation and face-to-face interviews eliciting information on symptoms including fever, cough, and other nonspecific symptoms consistent with COVID-19	NR/565 (2.3%)	9	498	54	4	0.69 (0.39 to 0.91)	0.90 (95% CI. 0.87 to 0.93)	0.14 (0.10 to 0.21)	0.99 (0.98 to 1.00)
	Sriwijitalai 2020	Travellers arriving at international airports in Thailand	First month of screening	Body temperature scanning and clinical history taking	NR	1	NR	11	NR	NA	NA	0.08	NA

Table 7. Study characteristics and results from screening test accuracy studies[a] *(Continued)*

		(time NR)										
Sutton 2020	Pregnant women admitted for delivery at New York–Presbyterian Allen Hospital and Columbia University Irving Medical Center, USA	22 March and 4 April 2020	Fever or other symptoms of COVID-19 on admission	33/214 (15.4%)	4	181	0	29	0.12 (0.03 to 0.28)	1.00 (0.98 to 1.00)	1.00	0.86 (0.85 to 0.88)

CI: confidence interval; **COVID-19:** coronavirus disease 2019; **FN:** false negative; **FP:** false positive; **NA:** not applicable; **NPV:** negative predictive value; **NR:** not reported; **PPV:** positive predictive value; **RT-LAMP:** reverse transcriptase loop-mediated isothermal amplification; **SARS-CoV-2:** severe acute respiratory syndrome coronavirus-2; **TN:** true negative; **TP:** true positive

[a]All studies compared against the RT-PCR as a reference standard.

[b]This value is based on reported numbers, which would result in the total N summing to 77 rather than 76. We retained the reported number. If the true negative is 17 rather than 18, the specificity would be 0.61 instead of 0.62.

Table 8. Certainty of the evidence: should symptom assessment be used to screen for SARS-CoV2 in people not presenting to clinical care for symptoms of COVID-19?

Outcome	№ of studies № of people	Study design	Factors that may decrease certainty of evidence					Estimated test outcome per 1000 people tested			Test accuracy certainty of evidence
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 15%	Pre-test probability of 30%	
True positives (people correctly classified as having SARS-CoV2)	13 studies (Baggett 2020; Barrett 2020; Graham 2020; Gudbjartsson 2020b; Guery 2020; Hoehl 2020; Kimball 2020; Lavezzo 2020; Lytras 2020; Olalla 2020; Roxby 2020; Samuels 2020) 597 people	Cross-sectional studies	Serious[a]	Serious[b]	Serious[c]	Not serious	None	0 to 6	0 to 90	0 to 180	⨁◯◯◯ Very low
False negatives (people incorrectly classified as not having SARS-CoV2)	4 to 10		60 to 150	120 to 300							

Table 8. Certainty of the evidence: should symptom assessment be used to screen for SARS-CoV2 in people not presenting to clinical care for symptoms of COVID-19? *(Continued)*

True negatives (people correctly classified as not having SARS-CoV2)	12 studies (Baggett 2020; Barrett 2020; Graham 2020; Gudbjartsson 2020b; Guery 2020; Hoehl 2020; Kimball 2020; Lytras 2020; Olalla 2020; Roxby 2020; Samuels 2020) 16,762 people	Cross-sectional studies	Serious[a]	Serious[b]	Not serious	Not serious	None	653 to 990	561 to 850	462 to 700	⊕⊕◯◯ Low
False positives (people incorrectly classified as having SARS-CoV2)	0 to 337							0 to 289	0 to 238		
Sensitivity	0.00 to 0.60			**Specificity**	0.66 to 1.00		**Prevalences**	1%, 15%, 30%			

[a]Uncertain risk of bias.
[b]Uncertain applicability of index test.
[c]Wide range of sensitivity estimates.

Table 9. Certainty of the evidence: should repeated symptom assessment be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19?

Outcome	№ of studies № of people	Study design	Factors that may decrease certainty of evidence					Estimated test outcome per 1000 people tested			Test accuracy certainty of evidence
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 15%	Pre-test probability of 30%	
True positives (people with SARS-CoV-2)	1 study 48 people (Arons 2020)	Cohort type accuracy study	Serious[a]	Serious[b]	not serious	Serious[c]	none	4 (3 to 6)	22 (14 to 30)	66 (44 to 89)	⊕◯◯◯ Very low
False negatives (people incorrectly classified as not having SARS-CoV-2)	6 (4 to 7)	28 (20 to 36)	84 (61 to 106)								

Table 9. Certainty of the evidence: should repeated symptom assessment be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19? *(Continued)*

True negatives (people without SARS-CoV-2)	1 study 29 patients (Arons 2020)	cohort type accuracy study	Serious[a]	Serious[b]	not serious	Serious[c]	none	614 (416 to 782)	589 (399 to 751)	527 (357 to 672)	⊕⊕◯◯ Low
False positives (people incorrectly classified as having SARS-CoV-2)	376 (208 to 574)	361 (199 to 551)	323 (178 to 493)								
Sensitivity	0.44 (95% CI, 0.29 to 0.59)		**Specificity**	0.62 (95% CI, 0.42 to 0.79)		**Prevalences**	1%, 15%, 30%				

[a]Unclear risk of bias.
[b]Uncertain applicability of index test.
[c]Few events.

Table 10. Certainty of the evidence: should direct temperature measurements be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19?

Outcome	№ of studies № of people	Study design	Factors that may decrease certainty of evidence					Estimated test outcome per 1000 people tested			Test accuracy certainty of evidence
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 15%	Pre-test probability of 30%	
True positives (people with SARS-Cov-2)	3 studies 184 people (Baggett 2020; Hoehl 2020; Samuels 2020)	Cross-sectional	Serious[a]	Not serious	Not serious	Serious[b]	none	0 to 0	0 to 1	0 to 2	⊕⊕◯◯ Low
False negatives (people incorrectly classified as not having SARS-Cov-2)	10 to 10	149 to 150	298 to 300								
True negatives (people without SARS-Cov-2)	3 studies 638 people (Baggett 2020; Hoehl 2020; Samuels 2020)	Cross-sectional	Serious[a]	Not serious	Not serious	Not serious	None	980 to 990	842 to 850	693 to 700	⊕⊕⊕◯ Moderate

Table 10. Certainty of the evidence: should direct temperature measurements be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19? *(Continued)*

False positives (people incorrectly classified as having SARS-Cov-2)	0 to 10	0 to 8	0 to 7		
Sensitivity	0.00 to 0.01	**Specificity**	0.99 to 1.00	**Prevalences**	1%, 15%, 30%

[a]Unclear risk of bias.
[b]Few events.

Table 11. Certainty of the evidence: should screening for international travel history be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19?

Outcome	№ of studies № of people	Study design	Factors that may decrease certainty of evidence					Estimated test outcome per 1000 people tested			Test accuracy certainty of evidence
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 15%	Pre-test probability of 30%	
True positives (people with SARS-CoV2)	2 studies 100 people (Gudbjartsson 2020b)	Cross-sectional	Serious[a]	Serious[b]	Not serious	Serious[c]	None	0 to 2	0 to 35	0 to 69	⊕◯◯◯ Very low
False negatives (people incorrectly classified as not having SARS-CoV2)	8 to 10	115 to 150	231 to 300								
True negatives (people without SARS-CoV2)	2 studies 12980 people (Gudbjartsson 2020b)	Cross-sectional	Serious[a]	Serious[b]	Not serious	Not serious	None	901 to 980	774 to 842	637 to 693	⊕⊕◯◯ Low
False positives (people incorrectly classified as having SARS-CoV2)	5 to 89	4 to 76	3 to 63								
Sensitivity	0.00 to 0.23	**Specificity**	0.91 to 1.00		**Prevalences**	1%, 15%, 30%					

[a]Unclear risk of bias.
[b]Limited applicability because of dynamic trajectory of pandemic; implementation of travel restrictions and slowing of worldwide travel.
[c]Few events.

Table 12. Certainty of the evidence: should screening for exposure to known infected people be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19?

Outcome	№ of studies № of people	Study design	Factors that may decrease certainty of evidence					Estimated test outcome per 1000 people tested			Test accuracy certainty of evidence
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pretest probability of 1%	Pretest probability of 15%	Pretest probability of 30%	
True positives (people with SARS-CoV2)	3 studies 102 people (Gudbjartsson 2020b; Hoehl 2020)	Cross-sectional	Serious[a]	Not serious	Not serious	Serious[b]	None	0 to 1	0 to 12	0 to 24	⊕⊕◯◯ Low
False negatives (people incorrectly classified as not having SARS-CoV2)	9 to 10	138 to 150	276 to 300								
True negatives (people without SARS-CoV2)	3 studies 13,103 people (Gudbjartsson 2020b; Hoehl 2020)	Cross-sectional	Serious[a]	Not serious	Not serious	Not serious	None	960 to 970	825 to 833	679 to 686	⊕⊕⊕◯ Moderate
False positives (people incorrectly classified as having SARS-CoV2)	20 to 30	17 to 25	14 to 21								
Sensitivity	0.00 to 0.08	**Specificity**	0.97 to 0.98		**Prevalences**	1%, 15%, 30%					

[a]Unclear risk of bias
[b]Few events

Table 13. Certainty of the evidence: should screening for exposure to known or suspected infectious people be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19?

Outcome	№ of studies № of people	Study design	Factors that may decrease certainty of evidence					Estimated test outcome per 1000 people tested			Test accuracy certainty of evidence
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 15%	Pre-test probability of 30%	
True positives (people with SARS-CoV-2)	2 studies 43 people (Barrett 2020; Hoehl 2020)	Cross-sectional	Serious[a]	Not serious	Not serious	Serious[b]	None	0 to 2	0 to 26	0 to 51	⊕⊕◯◯ Low
False negatives (people incorrectly classified as not having SARS-CoV-2)	8 to 10	124 to 150	249 to 300								
True negatives (people without SARS-CoV-2)	2 studies 911 people (Barrett 2020; Hoehl 2020)	Cross-sectional	Serious[a]	Not serious	Not serious	Not serious	None	891 to 960	765 to 825	630 to 679	⊕⊕⊕◯ Moderate
False positives (people incorrectly classified as having SARS-CoV-2)	30 to 99	25 to 85	21 to 70								
Sensitivity	0.00 to 0.17	**Specificity**	0.90 to 0.97	**Prevalences**	1%, 15%, 30%						

[a]Uncertain risk of bias.
[b]Few events.

Table 14. Certainty of the evidence: should screening for symptoms and temperature be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19?

Outcome	№ of studies № of people	Study design	Factors that may decrease certainty of evidence	Estimated test outcome per 1000 people tested	Test accuracy certainty

Table 14. Certainty of the evidence: should screening for symptoms and temperature be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19? *(Continued)*

			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 15%	Pre-test probability of 30%	of evidence
True positives (people with SARS-CoV2)	2 studies 46 people (Nishiura 2020; Sutton 2020)	Cross-sectional	Serious[a]	Serious[b]	Serious	Serious[c]	None	1 to 7	18 to 104	36 to 207	⊕◯◯◯ Very low
False negatives (people incorrectly classified as not having SARS-CoV2)	3 to 9	46 to 132	93 to 264								
True negatives (people without SARS-CoV2)	2 studies 733 people (Nishiura 2020; Sutton 2020)	Cross-sectional	Serious[a]	Serious[b]	Not serious	Not serious	None	891 to 990	765 to 850	630 to 700	⊕⊕◯◯ Low
False positives (people incorrectly classified as having SARS-CoV2)	0 to 99	0 to 85	0 to 70								
Sensitivity	0.12 to 0.69	**Specificity**	0.90 to 1.00	**Prevalences**	1%, 15%, 30%						

[a]Uncertain risk of bias.
[b]Uncertain applicability of index test.
[c]Few positive cases, wide confidence intervals.

Table 15. Certainty of the evidence: should RT-LAMP be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19?

Outcome	№ of studies № of people	Study design	Factors that may decrease certainty of evidence					Estimated test outcome per 1000 people tested			Test accuracy certainty of evidence
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Pre-test probability of 1%	Pre-test probability of 15%	Pre-test probability of 30%	

Table 15. Certainty of the evidence: should RT-LAMP be used to screen for SARS-CoV-2 in people not presenting to clinical care for symptoms of COVID-19? *(Continued)*

True positives (people with SARS-CoV2)	1 study 10 people (Osterdahl 2020)	Cross-sectional	Serious[a]	Serious[b]	Not serious	Serious[c]	None	8 (4 to 10)	120 (66 to 144)	240 (132 to 288)	⊕◯◯◯ Very low
False negatives (people incorrectly classified as not having SARS-CoV2)	2 (0 to 6)	30 (6 to 84)	60 (12 to 168)								
True negatives (people without SARS-CoV2)	1 study 11 people (Osterdahl 2020)	Cross-sectional	Serious[a]	Serious[b]	Not serious	Serious[c]	None	723 (386 to 931)	620 (332 to 799)	511 (273 to 658)	⊕◯◯◯ Very low
False positives (people incorrectly classified as having SARS-CoV2)	267 (59 to 604)	230 (51 to 518)	189 (42 to 427)								
Sensitivity	0.80 (95% CI 0.44 to 0.96)		**Specificity**	0.73 (95% CI 0.39 to 0.94)		**Prevalences**	1%, 15%, 30%				

[a]Unclear risk of bias.
[b]Study population is not fully representative of the relevant range for population screening (all are from a high dependency care home).
[c]Small sample size, few positive tests, wide confidence intervals.

APPENDICES

Appendix 1. Search log

Search	Database name, time span, and host	Date searched	Hits
Original search	Ovid MEDLINE(R) ALL 1946 to 3 April 2020	4 Apr 2020	439
	Embase.com (Elsevier)	4 Apr 2020	558
	Cochrane Central Register of Controlled Trials - Issue 4 of 12, April 2020 (Cochrane Library/Wiley)	4 Apr 2020	58
	LitCovid www.ncbi.nlm.nih.gov/research/coronavirus/)	4 Apr 2020	200
	CDC COVID-19 Research Articles Downloadable Database (www.cdc.gov/library/researchguides/2019novelcoronavirus/researcharticles.html) updated 3 April 2020	4 Apr 2020	775
	Cochrane Covid-19 Study Register (covid-19.cochrane.org)	4 Apr 2020	646
		Total (before deduplication)	2676
		Total (after deduplication)	1530
Search	Database name, time span, and host	Date searched	Hits
Update 1	Ovid MEDLINE(R) ALL 1946 to 13 April 2020	14 Apr 2020	622
	Embase.com (Elsevier)	14 Apr 2020	647
	Cochrane Central Register of Controlled Trials - Issue 4 of 12, April 2020 (Cochrane Library/Wiley)	14 Apr 2020	70
	CDC COVID-19 Research Articles Downloadable Database (www.cdc.gov/library/researchguides/2019novelcoronavirus/researcharticles.html) updated 13 April 2020	14 Apr 2020	1288
	Cochrane Covid-19 Study Register (CRSWeb)	14 Apr 2020	288
		Total (before deduplication)	2915
		Total (after deduplication)	502
Search*	Database name, time span, and host	Date searched	Hits
Update 2	Ovid MEDLINE(R) ALL 1946 to 26 May 2020	27 May 2020	2078
	CDC COVID-19 Research Articles Downloadable Database (https://www.cdc.gov/library/researchguides/2019novelcoronavirus/researcharticles.html) Updated 26 May 2020	27 May 2020	3918

(Continued)

	Total (before deduplication)	5996
	Total (after deduplication)	2325

Appendix 2. Ovid MEDLINE(R) ALL 1946 to 26 May 2020

	#	Searches	Results
A. SARS-CoV-2/ Covid19	1	exp Coronavirus/	14646
	2	exp Coronavirus Infections/	13506
	3	(coronavir* or coronovir*).ti,ab,kf.	18359
	4	((corona* or corono*) adj1 (virus* or viral* or virinae*)).ti,ab,kf.	730
	5	(ncov or n-cov or 2019nCoV or nCoV2019 or CO?VID-19 or CO?VID19 or WNCoV or WNCoV or HCoV-19 or HCoV19 or 2019 novel* or SARS-CoV-2 or SARS-CoV-2 or SARSCoV2 or SARS-CoV2 or SARSCov19 or SARS-Cov19 or SARS-Cov-19 or SARS-Cov-19 or Ncovor or Ncorona* or Ncorono* or NcovWuhan* or NcovHubei* or NcovChina* or NcovChinese* or SARS2 or SARS-2 or SARS-coron?virus2 or SARS-coron?virus-2 or SARScoron?virus 2 or SARS coron?virus2).ti,ab,kf.	16543
	6	or/1-5	36005
B1. screening terms - general	7	mass screening/ or anonymous testing/ or mandatory testing/ or multiphasic screening/	104484
	8	early diagnosis/	26186
	9	Diagnostic Tests, Routine/	12060
	10	Point-of-Care Testing/	1732
	11	screen*.ti,ab,kf.	746945
	12	((early or rapid* or routin*) adj6 (detect* or diagnos* or identif* or test*)).ti,ab,kf.	413931
	13	(drive adj6 (diagnos* or detect* or test* or identif*)).ti,ab,kf.	2537
	14	((point of care or poc) adj6 (detect* or diagnos* or identif* or test*)).ti,ab,kf.	11304
	15	Risk Assessment/	261974
	16	(risk? adj1 assess*).ti,ab,kf.	77680
	17	or/7-16	1432454

(Continued)

B2. specific tests	18	nucleic acid amplification techniques/ or exp polymerase chain reaction/	454428
	19	(Polymerase Chain Reaction or PCR).ti,ab,kf.	632874
	20	nucleic acid amplification.ti,ab,kf.	3819
	21	exp Immunoassay/	486987
	22	(immunoassay? or Enzyme-Linked Immunosorbent Assay? or ELISA).ti,ab,kf.	271339
	23	((detect* or identif* or test* or immuno* or assay?) adj2 (antibod* or anti bod*)).ti,ab,kf.	89153
	24	(IgM adj2 IgG).ti,ab,kf.	22389
	25	Clustered Regularly Interspaced Short Palindromic Repeats/	3090
	26	(CRISPR or Clustered Regularly Interspaced Short Palindromic Repeats or DETECTR).ti,ab,kf.	17982
	27	Body Temperature/	47226
	28	((temperature or fever) adj5 (check* or measur*)).ti,ab,kf.	28807
	29	or/18-28	1508837
B3. diagnosis of asympto-matic carriers	30	exp Coronavirus Infections/di, dg [Diagnosis, Diagnostic Imaging]	2588
	31	(diagnos* or detect* or test* or identif*).ti,ab,kf.	8805888
	32	30 or 31	8806871
	33	(asymptomatic* or presymptomatic* or pre-symptomatic*).mp.	160866
	34	32 and 33	97337
All B	35	17 or 29 or 34	2851989
A+B	36	6 and 35	7303
humans	37	limit 36 to "humans only (removes records about animals)"	4953
publications 2019+, added to Medline since Nov 2019	38	limit 37 to yr="2019 -Current"	2238
	39	(201912* or 2020*).dt.	662433
	40	38 and 39	2078

Appendix 3. Embase.com (Elsevier) 14 April 2020

Embase.com		
14 April 2020		
No.	**Query**	**Results**
#1	'coronaviridae'/de OR 'coronavirinae'/exp	19503
#2	'coronavirus infection'/exp	11904
#3	coronavir*:ti,ab,kw OR coronovir*:ti,ab,kw	13560
#4	((corona* OR corono*) NEAR/1 (virus* OR viral* OR virinae*)):ti,ab,kw	588
#5	ncov:ti,ab,kw OR 'n cov':ti,ab,kw OR 2019ncov:ti,ab,kw OR ncov2019:ti,ab,kw OR 'covid 19':ti,ab,kw OR covid19:ti,ab,kw OR 'wn cov':ti,ab,kw OR wncov:ti,ab,kw OR 'hcov 19':ti,ab,kw OR hcov19:ti,ab,kw OR '2019 novel*':ti,ab,kw OR 'sars cov 2':ti,ab,kw OR 'sarscov 2':ti,ab,kw OR sarscov2:ti,ab,kw OR 'sars cov2':ti,ab,kw OR sarscov19:ti,ab,kw OR 'sars cov19':ti,ab,kw OR 'sarscov 19':ti,ab,kw OR 'sars cov 19':ti,ab,kw OR ncovor:ti,ab,kw OR ncorona*:ti,ab,kw OR ncorono*:ti,ab,kw OR ncovwuhan*:ti,ab,kw OR ncovhubei*:ti,ab,kw OR ncovchina*:ti,ab,kw OR ncovchinese*:ti,ab,kw OR sars2:ti,ab,kw OR 'sars 2':ti,ab,kw OR sarscoronavirus2:ti,ab,kw OR 'sars coronavirus 2':ti,ab,kw OR 'sarscoronavirus 2':ti,ab,kw OR 'sars coronavirus2':ti,ab,kw	2580
#6	#1 OR #2 OR #3 OR #4 OR #5	27741
#7	'screening'/de OR 'antibody screening'/de OR 'mass screening'/de OR 'anonymous testing'/de OR 'multiphasic screening'/de	237115
#8	'early diagnosis'/de	104470
#9	'point of care testing'/de	13071
#10	screen*:ti,ab,kw	1039964
#11	((early OR rapid* OR routin*) NEAR/6 (detect* OR diagnos* OR identif* OR test*)):ti,ab,kw	584338
#12	(drive NEAR/6 (diagnos* OR detect* OR test* OR identif*)):ti,ab,kw	3508
#13	(('point of care' OR poc) NEAR/6 (detect* OR diagnos* OR identif* OR test*)):ti,ab,kw	15444
#14	'risk assessment'/de	551658
#15	(risk$ NEAR/1 assess*):ti,ab,kw	105388
#16	#7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15	2140049
#17	'reverse transcription polymerase chain reaction'/exp OR 'polymerase chain reaction'/de OR 'nucleic acid analysis'/de	674545
#18	'polymerase chain reaction':ti,ab,kw OR pcr:ti,ab,kw	858787

(Continued)

#19	'nucleic acid amplification':ti,ab,kw	5391
#20	'immunoassay'/exp	573575
#21	immunoassay$:ti,ab,kw	90613
#22	'enzyme-linked immunosorbent assay$':ti,ab,kw OR elisa:ti,ab,kw	314528
#23	((detect* OR identif* OR test* OR immuno* OR assay$) NEAR/2 (antibod* OR 'anti bod*')):ti,ab,kw	112345
#24	(igm NEAR/2 igg):ti,ab,kw	30924
#25	'clustered regularly interspaced short palindromic repeat'/de	10521
#26	crispr:ti,ab,kw OR 'clustered regularly interspaced short palindromic repeats':ti,ab,kw OR detectr:ti,ab,kw	24376
#27	'body temperature measurement'/de OR 'body temperature monitoring'/de	2227
#28	((temperature OR fever) NEAR/5 (check* OR measur*)):ti,ab,kw	28804
#29	#17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28	1783204
#30	'coronavirus infection'/exp/dm_di	1777
#31	diagnos*:ti,ab,kw OR detect*:ti,ab,kw OR test*:ti,ab,kw OR identif*:ti,ab,kw	11462850
#32	#30 OR #31	11463559
#33	'asymptomatic disease'/de	13987
#34	asymptomatic*:ti,ab,kw OR presymptomatic*:ti,ab,kw OR 'pre symptomatic*':ti,ab,kw	226575
#35	#33 OR #34	231489
#36	#32 AND #35	147015
#37	#16 OR #29 OR #36	3801093
#38	#6 AND #37	8366
#39	('animal experiment'/de OR 'animal'/exp) NOT 'human'/exp	5593902
#40	#38 NOT #39 AND [1-11-2019]/sd NOT [15-4-2020]/sd AND [2019-2020]/py	647

Appendix 4. Cochrane Central Register of Controlled Trials - Issue 4 of 12, April 2020 (Cochrane Library/Wiley)

CENTRAL (Cochrane Library/Wiley)
14 April 2020

(Continued)

ID	Search	Hits
#1	[mh Coronavirus]	11
#2	[mh "Coronavirus Infections"]	38
#3	(coronavir* OR coronovir*):ti,ab,kw	144
#4	((corona* OR corono*) NEAR/1 (virus* OR viral* OR virinae*)):ti,ab,kw	14
#5	(ncov or n-cov or 2019nCoV or nCoV2019 or COVID-19 or COVID19 or WN-CoV or WN-CoV or HCoV-19 or HCoV19 or 2019 novel* or SARS-CoV-2 or SARSCoV-2 or SARS-CoV2 or SARS-CoV2 or SARSCov19 or SARS-Cov19 or SARSCov-19 or SARS-Cov-19 or Ncovor or Ncorona* or Ncorono* or NcovWuhan* or NcovHubei* or NcovChina* or NcovChinese* or SARS2 or SARS-2 or SARScoronavirus2 or SARS-coronavirus-2 or SARScoronavirus 2 or SARS coronavirus2):ti,ab,kw	369
#6	(#1-#5)	473
#7	[mh ^"mass screening"] OR [mh ^"anonymous testing"] OR [mh ^"mandatory testing"] OR [mh ^"multiphasic screening"]	3018
#8	[mh ^"early diagnosis"]	491
#9	[mh ^"Diagnostic Tests, Routine"]	220
#10	[mh ^"Point-of-Care Testing"]	60
#11	screen*:ti,ab,kw	67736
#12	((early OR rapid* OR routin*) NEAR/6 (detect* OR diagnos* OR identif* OR test*)):ti,ab,kw	18158
#13	(drive NEAR/6 (diagnos* OR detect* OR test* OR identif*)):ti,ab,kw	149
#14	(("point of care" OR poc) NEAR/6 (detect* OR diagnos* OR identif* OR test*)):ti,ab,kw	962
#15	[mh ^"Risk Assessment"]	8524
#16	(risk? NEAR/1 assess*):ti,ab,kw	23568
#17	(#7-#16)	103152
#18	[mh ^"nucleic acid amplification techniques"] OR [mh "polymerase chain reaction"]	2062
#19	("Polymerase Chain Reaction" OR PCR):ti,ab,kw	13213
#20	nucleic acid amplification:ti,ab,kw	164
#21	[mh Immunoassay]	4540
#22	(immunoassay? OR ELISA):ti,ab,kw	10976
#23	((detect* OR identif* OR test* OR immuno* OR assay?) NEAR/2 (antibod* OR (anti NEXT bod*))):ti,ab,kw	3894

(Continued)

#24	(IgM NEAR/2 IgG):ti,ab,kw	328
#25	[mh ^"Clustered Regularly Interspaced Short Palindromic Repeats"]	0
#26	(CRISPR OR "Clustered Regularly Interspaced Short Palindromic Repeats" OR DETECTR):ti,ab,kw	15
#27	[mh ^"Body Temperature"]	2168
#28	((temperature OR fever) NEAR/5 (check* OR measur*)):ti,ab,kw	2383
#29	{or #18-#28}	33818
#30	[mh "Coronavirus Infections"/DG] or [mh "Coronavirus Infections"/DI]	0
#31	(diagnos* OR detect* OR test* OR identif*):ti,ab,kw	585908
#32	(asymptomatic* OR presymptomatic* OR pre-symptomatic*):ti,ab,kw	10101
#33	(#30 or #31) and #32	5906
#34	#17 or #29 or #33	137335
#35	#6 and #34	121
#36	#35 with Publication Year from 2019 to 2020, in Trials	70

Appendix 5. CDC COVID-19 Research Articles Downloadable Database

(www.cdc.gov/library/researchguides/2019novelcoronavirus/researcharticles.html) updated 26 May 2020

Searched via Endnote x8		
Field	**Search terms**	**Result**
Title	screen	273
Title	diagnos	784
Title	detect	624
Title	identif	277
Any field	screening	1001
Any field	asymptomatic	881
Any field	presymptomatic	35
Any field	pre-symptomatic	43

(Continued)

Total (including duplicates)	3918

Appendix 6. Cochrane Covid-19 Study Register (CRSWeb)

Cochrane Covid-19 Study Register (CRSWeb)		
14 April 2020		
Query	**Search**	**Results**
1	screening AND COVID19:INREGISTER	98
2	(rapid OR early) NEAR1 (detect* OR identif* OR diagnos*) AND COVID19:INREGISTER	51
3	asymptomatic* OR presymptomatic* OR pre-symptomatic* AND COVID19:INREGISTER	72
4	(detect* OR identif* OR diagnos* OR screen* OR test*):ti AND COVID19:INREGISTER	132
5	#1 OR #2 OR #3 OR #4	288

Appendix 7. LitCovid 4 April 2020

(www.ncbi.nlm.nih.gov/research/coronavirus/)

LitCOVID	
4 April 2020	
Search	**Results**
screening	96
(rapid AND (testing OR diagnosis OR identification)	26
(early AND (testing OR diagnosis OR identification)	35
(asymptomatic AND (testing OR diagnosis OR identification)	27
(presymptomatic AND (testing OR diagnosis OR identification)	2
(pre-symptomatic AND (testing OR diagnosis OR identification)	14
Total (includes duplicates)	200

Appendix 8. Data extraction table for diagnostic accuracy studies

Refid
User
Level
Header → Not a diagnostic accuracy study
First author last name
Year published
RefID
Study objective
Study design → diagnostic test accuracy study
Geographic setting
Geographic prevalence
Setting type → healthcare settings (hospital, clinics, ED, etc.)
Setting type → healthcare settings (hospital, clinics, ED, etc.) (COMMENT)
Setting type → congregate living settings (dormitories, assisted living facilities, long-term care or nursing homes, prisons)
Setting type → congregate living settings (dormitories, assisted living facilities, long-term care or nursing homes, prisons) (COMMENT)
Setting type → transportation hubs (airports, train and bus stations)
Setting type → transportation hubs (airports, train and bus stations) (COMMENT)
Setting type → other (describe)
Setting type → other (describe) (COMMENT)
Setting type description (additional narrative description of setting)
Total N
Female
Age
Comorbidities
Symptom status
Symptom status_comment

(Continued)

Symptom status narrative (narrative description of the population)
Other population characteristics. Use this to record other characteristics reported by the study that you think might be relevant and that aren't captured by the above.
Index screening strategy 1 description
Index screening strategy 2 description
Index screening strategy 3 description
Reference test 1
Reference test 2
Reference test 3
Comparison 1
Total N analyzed_1
True positives_1
True negatives_1
False positives_1
False negatives_1
Comparison 2
Total N analyzed_2
True positives_2
True negatives_2
False positives_2
False negatives_2
Comparison 3
Total N analyzed_3
True positives_3
True negatives_3
False positives_3
False negatives_3
Calculated Sn

(Continued)

Calculated Sp
Study reported sensitivity
Study reported specificity
Study reported positive predictive value
Study reported negative predictive value
AUC
Study reported likelihood ratios
Other relevant results
Please check appropriate status → I have reviewed the reference list and there are no relevant citations to pull for review.
Please check appropriate status → I have reviewed the reference list and there are relevant citations to review.
Cut and paste any relevant citations from the reference list here.
Name of primary abstractor
For primary abstractors only, check the box below when you have completed the form and are ready to send it for second review. The form will remain in your 'reviewed by me' queue until you check this box. → I am ready to submit this form for second review author.
Name of second review author
Comments on 2nd review: Note: 2nd review author should fix typos and adjust text for consistency directly on the form. Use this box to record substantive issues that need discussion with the primary review author before the form can be finalized.
For second review author only, check the box when you have completed your second review. The form will remain in your 'reviewed by me' queue after you submit the form, but we will not know whether you have finished reviewing until you check this box. → I have finished second reviewing this study.

Appendix 9. Data extraction table for modelling studies

Refid
User
Level
First author last name
Year published
RefID
Research question

(Continued)

Model type
Population
% asymptomatic
Awareness of exposure
Incubation period
Time from symptoms to hospitalization
Time
How is dynamics of epidemic considered?
Screening strategy
Test 1
Test 2
Follow-up intervention
Time between screenings
Test 1 sensitivity
Test 2 sensitivity
Screening
Additional measures
Outcome 1
Outcome 2
Conclusion
Code available
Please check appropriate status → I have reviewed the reference list and there are no relevant citations to pull for review.
Please check appropriate status → I have reviewed the reference list and there are relevant citations to review.
Cut and paste any relevant citations from the reference list here.
Name of primary abstractor
For primary abstractors only, check the box below when you have completed the form and are ready to send it for second review. The form will remain in your 'reviewed by me' queue until you check this box. → I am ready to submit this form for second review author.
Name of second review author

(Continued)

For second review author only, check the box when you have completed your second review. The form will remain in your 'reviewed by me' queue after you submit the form, but we will not know whether you have finished reviewing until you check this box. → I have finished second reviewing this study.

Appendix 10. Assessment of methodological quality for modelling studies

Questions
1. Is the study population clearly described?
2. Are competing alternatives clearly described?
3. Is a well-defined research question posed in answerable form?
4. Is the chosen time horizon appropriate to include relevant consequences?
5. Are all important and relevant outcomes for each alternative identified?
6. Are all outcomes measured appropriately?
7. Are outcomes valued appropriately?
8. Is an analysis of outcomes of alternatives performed?
9. Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?
10. Do the conclusions follow from the data reported?
11. Does the study discuss the generalizability of the results to other settings and patient/client groups?
12. Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?
13. Are ethical and distributional issues discussed appropriately?
14. Overall

Adapted from list by Evers 2005

Appendix 11. Assessment of methodological quality for modelling studies - part 1

Study	Is the population clearly described?	Are competing alternatives clearly described?	Is a well-defined research question posed in an answerable form?	Is the chosen time horizon appropriate to include relevant consequences?	Are all important and relevant outcomes for each alternative identified?	Are all outcomes measured appropriately?	Are outcomes valued appropriately?
Clifford 2020	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gostic 2020	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quilty 2020	Yes	Yes	Yes	No	No	Yes	Yes
Wells 2020	No	No	No	Yes	No	No	No
Zhang 2020	Yes	Yes	Yes	Yes	Unclear	Yes	Yes

Appendix 12. Assessment of methodological quality for modelling studies - part 2

Study	Is an analysis of outcomes of alternatives performed?	Are all important variables whose values are uncertain, appropriately subjected to sensitivity analysis?	Do the conclusions follow from the data reported?	Does the study discuss the generalizability of the results to other settings and patient/client groups?	Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	Are ethical and distributional issues discussed appropriately?	Overall assessment
Clifford 2020	Yes	Yes	Yes	Yes	Yes	Yes	Minor or no concerns
Gostic 2020	Yes	Yes	Yes	Yes	Yes	Yes	Minor or no concerns
Quilty 2020	Yes	Yes	Yes	Yes	Yes	Yes	Moderate concerns
Wells 2020	No	No	Yes	Yes	No	Yes	Major concerns
Wells 2020	Yes	No	Yes	Yes	No	Unclear	Major concerns

Appendix 13. Assessment of methodological quality for studies of diagnostic test accuracy - part 1

Study	Participant selection - was a consecutive or random sample of participants enrolled?	Participant selection - was case-control design avoided?	Participant selection - did the study avoid inappropriate exclusions?	Domain rating: participant selection - could the selection of participants have introduced bias?	Domain rating: applicability - is there concern that the included participants do not match the review question?
Arons 2020	No	Yes	Yes	Low	High
Baggett 2020	Yes	Yes	Yes	Low	Low
Barrett 2020	Yes	No	Yes	High	High
Graham 2020	Yes	Yes	Yes	Low	High
Gudbjartsson 2020b	Yes	Yes	Yes	Low	Low
Guery 2020	Yes	Yes	Yes	Low	High
Hoehl 2020	Yes	Yes	Yes	Low	Low
Kimball 2020	Yes	Yes	Yes	Low	High
Lavezzo 2020	Unclear	Yes	Unclear	Low	Unclear
Lytras 2020	Yes	Yes	Yes	Low	Low
Nishiura 2020	Yes	Yes	Yes	Low	Low
Olalla 2020	Yes	Yes	Yes	Low	Low
Osterdahl 2020	Yes	Yes	Yes	Low	High
Roxby 2020	Yes	Yes	Yes	Low	High
Samuels 2020	Yes	Yes	Yes	Low	Low
Sriwijitalai 2020	Unclear	Yes	Unclear	Unclear	Low
Sutton 2020	Yes	Yes	Yes	Low	Low

Appendix 14. Assessment of methodological quality for studies of diagnostic test accuracy - part 2

Study	Index test - were the index test results interpreted without knowledge of the results of the reference standard?	Index test - if a threshold was used, was it pre-specified?	Domain rating: index test - could the conduct or interpretation of the index test have introduced bias?	Domain applicability: index test - is there concern that the index test, its conduct, or interpretation differ from the review question?	Reference standard - is the reference standard likely to correctly classify the target condition?	Reference standard - were the reference standard results interpreted without knowledge of the results of the index test?	Domain rating: reference standard - could the reference standard, its conduct, or interpretation have introduced bias?	Domain applicability: reference standard - is there concern that the target condition as defined by the reference standard does not match the review question?
Arons 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Baggett 2020	Yes	Unclear	Unclear	Low	Unclear	Unclear	Unclear	Unclear
Barrett 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Graham 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Gudbjartsson 2020b	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Guery 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Hoehl 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Kimball 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Lavezzo 2020	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Lytras 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Nishiura 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Olalla 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Osterdahl 2020	Yes	Unclear	Unclear	Unclear	Unclear	No	Unclear	Unclear
Roxby 2020	Yes	Unclear	Unclear	Unclear	Yes	Unclear	Unclear	Unclear
Samuels 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear

(Continued)

Sriwijitalai 2020	Yes	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Sutton 2020	Yes	Unclear	Unclear	Unclear	Yes	Unclear	Unclear	Unclear

Appendix 15. Assessment of methodological quality for studies of diagnostic test accuracy - part 3

Study	Flow and timing - was there an appropriate interval between index test(s) and reference standard?	Flow and timing - did all participants receive a reference standard?	Flow and timing - did all participants receive the same reference standard?	Flow and timing - Were all participants included in the analysis?	Domain rating Flow and timing - could the participant flow have introduced bias?
Arons 2020	Yes	Yes	Yes	Unclear	Low
Baggett 2020	Yes	Yes	Yes	Unclear	Low
Barrett 2020	Yes	Yes	Yes	Unclear	Low
Graham 2020	Yes	Yes	Yes	Unclear	Low
Gudbjartsson 2020b	Yes	Yes	Yes	Unclear	Low
Guery 2020	Yes	Yes	Yes	Unclear	Low
Hoehl 2020	Yes	Yes	Yes	Unclear	Low
Kimball 2020	Yes	Yes	Yes	Unclear	Low
Lavezzo 2020	Yes	Yes	Yes	Unclear	Low
Lytras 2020	Yes	Yes	Yes	Unclear	Low
Nishiura 2020	Yes	Yes	Yes	Unclear	Low
Olalla 2020	Yes	Yes	Yes	Unclear	Low
Osterdahl 2020	Yes	Yes	Yes	No	Low
Roxby 2020	Yes	Yes	Yes	Unclear	Low
Samuels 2020	Yes	Yes	Yes	Unclear	Low
Sriwijitalai 2020	Unclear	Unclear	Unclear	Unclear	Unclear
Sutton 2020	Yes	Yes	Yes	Unclear	Low

Appendix 16. Assessment of methodological quality for studies of diagnostic test accuracy - part 4

Study	What is your overall rating of risk of bias?	Comments on rating
Arons 2020	Unclear	Blinding unclear, accuracy of reference standard unclear

(Continued)

Baggett 2020	Unclear	Threshold for symptom assessment unclear, accuracy of reference standard unclear, blinding unclear
Barrett 2020	High	Drew from two separate populations with likely different prevalences. Threshold for symptom assessment unclear, accuracy of reference standard unclear, blinding unclear
Graham 2020	Unclear	Threshold for symptom assessment unclear, accuracy of reference standard unclear, blinding unclear
Gudbjartsson 2020b	Unclear	Although marked as low, the population may reflect higher baseline prevalence because not all people were tested. Although the population screening was open to all, those with symptoms may have selectively sought registration
Guery 2020	Unclear	Threshold for symptom assessment unclear, accuracy of reference standard unclear, blinding unclear
Hoehl 2020	Unclear	One person refused the reference test, but in other respects, there are no obvious sources of bias, other than the lack of information on some signalling questions.
Kimball 2020	Unclear	Two considerations that could potentially pose bias considerations (1) the accuracy of the reference standard is unclear (2) not all residents could be included, and it's not clear why some residents were excluded
Lavezzo 2020	Unclear	Limited information on participant selection, index test, and reference standard
Lytras 2020	Unclear	Limited information on index test and reference standard
Nishiura 2020	Unclear	No information on index threshold, no information on blinding when interpreting results
Olalla 2020	Unclear	Limited information on index text and reference standard
Osterdahl 2020	Unclear	No threshold values specified, blinding on tests not specified. Tests were not conducted on the same day for all participants but likely the impact on bias is low. 6 of 15 tested positive on RT-PCR on day 1/3 and got the RT-LAMP on day 4. 15 of 21 who tested positive
Roxby 2020	Unclear	Potential risk of recall bias, because people were asked about symptoms in the past 14 days. Several items not clearly described
Samuels 2020	Unclear	Threshold for symptom assessment unclear, accuracy of reference standard unclear, blinding unclear
Sriwijitalai 2020	Unclear	Very little information provided on most aspects of design. No information on participants' negative test results
Sutton 2020	Unclear	No clarity on threshold for index test, no information on blinding in interpreting results

RT: LAMP: reverse transcriptase loop-mediated isothermal amplification; **RT-PCR:** reverse transcription polymerase chain reaction

HISTORY

Review first published: Issue 9, 2020

CONTRIBUTIONS OF AUTHORS

Study design: MV, GG, LCK

Literature search: IK

Study selection, data collection: MV, GG, LCK, KG, AD, CH, EP, BT, JM

Quality assessment of studies: MV, GG, LCK, BJ

GRADEing the certainty of evidence: MV, GG, LCK, BJ

Narrative synthesis and data interpretation: MV, GG, LCK, BJ

Figures and tables: MV, LCK, KG, EP

Writing the manuscript: MV, GG, LCK, BJ, IK, KG

Supervision: MV, GG

Critical revision and feedback to the manuscript: MV, GG, LCK, BJ, KG, IK, AD, CH, EP, BT, JM

DECLARATIONS OF INTEREST

Meera Viswanathan: none declared

Leila Kahwati: none declared

Beate Jahn: none declared

Kayla Giger: none declared

Andreea Iulia Dobrescu: none declared

Christine Hill: none declared

Irma Klerings: none declared

Jana Meixner: none declared

Emma Persad: none declared

Birgit Teufer: none declared

Gerald Gartlehner: none declared

SOURCES OF SUPPORT

Internal sources

- RTI International, USA
- Danube University Krems, Austria
- University for Health Sciences, Medical Informatics and Technology, Austria

Investigators were supported by internal funds from their respective institutions, listed above.

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We revised the title from 'Screening asymptomatic persons for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)' to 'Universal screening for SARS-CoV-2'. This change was intended to clarify the scope of the review and the specific outcomes being considered under the review. We also switched the order of the questions to be reviewed, to focus on effectiveness first, and then accuracy, and clarified eligible study designs for each question. We removed antibody tests from the list of eligible screening approaches because these are unlikely to be used for mass screening to identify active infections. Finally, we added false alarm and false reassurance to the effectiveness question to address potential harms from screening (Viswanathan 2020).