

A Systematic and Meta-Analysis Review on the Diagnostic Accuracy of Antibodies in the Serological Diagnosis of COVID-19.

Arthur Vengesai (✉ arthurvengesai@gmail.com)

University of KwaZulu-Natal College of Health Sciences <https://orcid.org/0000-0002-7629-6843>

Herald Midzi

University of KwaZulu-Natal College of Health Sciences

Maritha Kasambala

University of KwaZulu-Natal School of Life Sciences

Hamlet Mutandadzi

University of Zimbabwe College of Health Sciences

Tariro L. Mduluza-Jokonya

University of KwaZulu-Natal College of Health Sciences

Simbarashe Rusakaniko

University of Zimbabwe College of Health Sciences

Francisca Mutapi

The University of Edinburgh School of Biological Sciences

Thajasvarie Naicker

University of KwaZulu-Natal College of Health Sciences

Takafira Mduluza

University of Zimbabwe Faculty of Science

Research

Keywords: Serology, COVID-19, SARS-CoV2, rRT-PCR, IgG, IgM, specificity, sensitivity

Posted Date: June 12th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-34638/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Systematic Reviews on May 26th, 2021. See the published version at <https://doi.org/10.1186/s13643-021-01689-3>.

Abstract

Background: Serological testing based on different antibody types are an alternative method being used to diagnose SARS-CoV-2 and has the potential of having higher diagnostic accuracy compared to the current gold standard RT-PCR. Therefore, the objective of this review was to evaluate the diagnostic accuracy of IgG and IgM based Point-of-care (POC) lateral flow immunoassays (LFIA), chemiluminescence enzyme immunoassay (CLIA), fluorescence enzyme-linked immunoassay (FIA) and ELISA systems that detect SARS-CoV-2 antigens.

Method: A systematic literature search was carried out in PubMed, Medline complete and MedRxiv. Studies evaluating the diagnostic accuracy of serological assays for SARS-CoV-2 were eligible. Study selection and data-extraction were done by two authors independently. QUADAS-2 checklist tool was used to assess the quality of the studies. The bivariate model and the hierarchical summary receiver operating characteristic curve model were performed to evaluate the diagnostic accuracy of the serological tests. Subgroup meta-analysis analyses was performed to explore the heterogeneity.

Results: The pooled sensitivity for IgG, IgM and IgG-IgM based LFIA tests were 0.5856, 0.4637 and 0.6886 respectively compared to RT-PCR method. The pooled sensitivity for IgG and IgM based CLIA tests were 0.9311 and 0.8516 respectively compared to RT-PCR. The pooled sensitivity the IgG, IgM and IgG-IgM based ELISA tests were 0.8292, 0.8388 and 0.8531 respectively compared to RT-PCR. All tests displayed high specificities ranging from 0.9693 to 0.9991. Among the evaluated tests, IgG based CLIA expressed the highest sensitivity signifying its accurate detection of the largest proportion of infections identified by RT-PCR. ELISA and CLIA tests performed better in terms of sensitivity compared to LFIA. IgG based tests performed better compared to IgM ones expect for the ELISA.

Conclusions: We report that IgG-IgM based ELISA tests have the best overall diagnostic test accuracy. Moreover, irrespective of the method, a combined IgG/IgM test seems to be a better choice in terms of sensitivity than measuring either antibody type independently. Given the poor performances of the current LFIA devices there is need for more research on the development of highly sensitivity and specific POC LFIA that are adequate for most individual patient applications and attractive for large sero-prevalence studies.

Systematic review registration: PROSPERO Registration Number is: CRD42020179112

Introduction

Coronavirus disease 2019 (COVID-19) is a major contagious pandemic of respiratory disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which is also known as novel (new) coronavirus 2019-nCoV^{1 2 3}. The first COVID-19 cases were identified in December 2019 from Wuhan, Hubei Province, China⁴. On May 27 2020, according to the World Health organisation (WHO) COVID-19 Situation Report 127, there were approximately 5 million confirmed cases of COVID-19 and 187 705 deaths worldwide⁵. Although the COVID-19 clinical features are not yet fully known and understood, clinicians have reported clinical manifestations range from asymptomatic cases to patients with mild and severe respiratory illness, with or without pneumonia, fever, cough and shortness of breath. Serious illness including death occurs in 16% of cases. Older people (> 65 years) and people of all ages with severe chronic medical conditions such as lung diseases, heart disease and diabetes seem to have a higher risk of succumbing to severe COVID-19 illness³.

Early and accurate diagnostic testing for COVID-19 is critical for tracking the SARS-CoV-2, understanding the virus epidemiology, informing case management, suppressing transmission and for quarantine purposes^{6 7}. The current diagnostic confirmatory test for COVID-19 is the real-time reverse-transcriptase polymerase chain reaction (rRT-PCR) which identifies the positive nucleic acids of SARS-CoV-2 when present in sufficient quantity in sputum, throat swabs, and secretions of the lower respiratory tract. This molecular test takes between four to six hours which is time consuming. It requires expensive specialist equipment, skilled laboratory personal for sample preparation and testing and PCR reagents, creating diagnostic delays and limiting use in real-life situations when rapid diagnosis is required for fast intervention decisions. Therefore, less expensive and easy implementable tests are needed for detecting SARS-CoV-2. Another limitation of using RT-PCR involves the use of swabs from the upper respiratory tract which can be falsely diagnosed as negative due to the poor quality of the sample or acquiring the sample at an incorrect

timeframe; notably viral load in upper respiratory tract secretions peak in the first week of symptoms but may decline below the limit of detection in patients presenting late with symptoms^{7-9 10}. Missing the time-window of viral replication may also provide false negative results. Moreover, after a variable period of time, one expects the rRT-PCR result to become negative as people stop viral shedding¹¹.

False-negative rRT-PCR results are common during diagnosis of SARS-CoV-2. The Fever Clinic of the Beijing Haidian Hospital collected data from January 2020 which indicated that two out of ten negative cases diagnosed using rRT-PCR test were eventually confirmed to be positive cases for COVID-19. This yielded an approximately 20% false-negative rate of rRT-PCR¹⁰. Zhang *et al* also showed that the current strategy for the detection of viral RNA in oral swabs used for SARS-CoV-2 diagnosis is not 100% accurate. The presence of the virus has been detected in anal swabs or blood samples of patients while their oral swabs diagnosis reports a negative result. This observation implies that a patient cannot be discharged based purely on oral swabs samples being negative¹².

A false negative diagnosis may have grave consequences, especially at this stage of the COVID-19 pandemic by allowing SARS-CoV-19 infected patients to spread the infection thereby hampering the efforts to contain the spread of the virus⁷. Additional screening methods that can detect the presence of infection despite lower viral titres are highly beneficial to ensure timely diagnosis of all COVID-19 patients. Other screening methods that have been used include the use of a chest computed tomography (CT). Chest CT has been used as a confirmatory diagnostic method for patients with high COVID-19 clinical suspicion that are SARS-CoV-2 negative by RT-PCR. The chest CT provides more information, but is inconclusive as not all the patients with COVID-19 develop pneumonia and other pathogens may also cause pneumonia¹³.

Detection of serum specific anti-SARS-CoV-2 antibodies, both immunoglobulin G (IgG) and M (IgM) which are produced rapidly after the infection provide an alternative highly sensitive and accurate solution and compensates for the limitations of RT-PCR. The serological methods could also be a more practical alternative to chest CT^{7,14,15} Immunoglobulin G antibodies permit the use of serological tools to better understand the overall rate of COVID-19 infections including the rate of asymptomatic infections. In addition, the higher speed of antibody testing is useful and may be applied in remote and poor resource areas where rRT-PCR assays cannot be performed⁷.

However the dynamics of blood or serum antibodies in the cases of COVID-19 are not well evaluated. Currently the serological dynamics of COVID-19 patients remain limited. Also, before diagnostic assays are widely deployed, their performance characteristics need to be evaluated. Therefore, the objective of this review was to evaluate the diagnostic accuracy of IgG and IgM (together or separately) based Point-of-care (POC) lateral flow immunoassays (LFIA), chemiluminescence enzyme immunoassay (CLIA), Fluorescence enzyme-linked immunoassay and luciferase immunoprecipitation assay systems (LIPS) to detect antigens against SARS-CoV-2.

Review questions

The primary research questions of this systematic review were:

1. What is the sensitivity and specificity of antibody serology tests for the target condition COVID-19?
2. What is the diagnostic accuracy of antibody serology tests for COVID-19?

Materials And Methods

We conducted a systematic review and meta-analysis in accordance with the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic test accuracy (PRISMA-DTA)¹⁶ (**Additional** Fig. 1). We used the Cochrane recommendations to report systematic reviews and meta-analyses of studies on diagnostic Accuracy¹⁷. We also used protocols from published systematic and meta-analysis reviews to develop our protocol¹⁸⁻²¹. The developed systematic review protocol was registered in the International Prospective Register of Systematic Reviews registration number CRD42020179112.

Eligibility criteria

Cross-sectional studies would be the ideal study type to answer our review questions. However, as we anticipated that serological diagnosis cross-sectional studies for COVID-19 would be very sparse, we decided to include case-control studies. The inclusion criteria comprised studies in which the study population or serum samples from patients diagnosed with COVID-19, were subjected to rRT-PCR testing or genetic sequencing and either one or all of the following serological tests; Point-of-care (POC) lateral flow immunoassays (LFIA), chemiluminescence enzyme immunoassay (CLIA), Fluorescence Immunoassays (FIA) and enzyme-linked immunoassay. The diagnostic accuracy of the tests was defined as the primary outcome. Original studies were included without restriction based on language or geographical location. We included studies between 1 January 2020 and 27 April 2020. We excluded animal studies and *in vitro* based studies. Survey studies investigating seroprevalence were also excluded.

Information sources and search strategies

The following databases were searched for studies: MEDLINE Complete (EBSCO) and PubMed. Grey literature was researched on MedRxiv (a preprint server for health sciences which distributes complete but unpublished manuscripts). We performed the search strategy on studies dated until April 2020. We used the following search strategy for PubMed and adapted it for the other databases: COVID-19 and serologic test (and its synonyms) (**Appendix 1**). (((((COVID-19[Title/Abstract]) OR SARS-CoV-2[Title/Abstract]) OR 2019-nCoV[Title/Abstract]) OR Wuhan Coronavirus[Title/Abstract])) AND (((((((((((Serologic test) OR Serologic method) OR Serological test) OR Serological method) OR Serodiagnosis) OR Immunodiagnosis) OR Immunological test) OR Immunological method) OR Antibody detection) OR Antigen detection) OR IgM) OR IgG) OR Immunochromatography). Additional studies were identified by contacting experts in the field and by searching reference lists from primary studies, review articles and textbook chapters.

Study selection and data extraction

Two authors (AV and HM) assessed the titles identified by the search, excluding those obviously irrelevant to the serological diagnosis of COVID-19. Letters, review articles, and articles clearly irrelevant based on examination of the abstract and other notes and duplicates were excluded next. The eligibility of the remaining potentially relevant articles was be judged on full text publications.

Data extraction was conducted independently by two authors (AV and HM) to avoid bias and discrepancies and were resolved by discussion. Where an agreement could not be reached, a third author was consulted. Where it remained unclear whether a study is eligible for inclusion, it was then excluded. While extracting data, authors also had to decide whether a study was a case-control or a cross-sectional study. The following data were extracted.

- Study authors and publication year
- Study design.
- Case definition (Inclusion/exclusion criteria)
- Participant demographics
- Reference standard (including criteria for positive test).
- Index tests (cut-off values (prespecified or not) and whether the test was a commercial or in-house test).
- Geographical location of data collection
- Index/reference time interval.
- Distribution of severity of disease in those with target condition
- Other diagnoses in those without target condition
- Numbers of TPs, FNs, FPs, and FNs.

Other considerations and exclusion criteria

Studies from which a 2x2 table containing true positives, false positives, false negatives and true negatives could not be drawn, were excluded. Furthermore, studies that were too unspecific in their reporting to ensure that they fulfilled the above criteria, were excluded.

Assessment of methodological quality

The QUADAS-2 tool was used to assess the methodological quality of all studies included in this systematic review²². QUADAS-2 consists of four key domains: patient selection, index test, reference standard, and flow and timing. We assessed all domains for risk of bias (ROB) potential and the first three domains for applicability concerns. Risk of bias was judged as “low,” “high,” or “unclear.” Details are shown in appendix 2. Two review authors (AV and HM) independently completed QUADAS-2. The divergences were resolved by consensus among the researchers.

Statistical analysis and data synthesis

Diagnostic accuracy

For each study we constructed 2 x 2 tables for true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). Where only sensitivity and specificity estimates were reported, we derived the two-by-two table from the reported data. We constructed forest plots displaying sensitivity and specificity of the index tests from contingency tables assuming that the reference method was 100% sensitive and specific. Data were entered into the Review Manager (RevMan) software for Windows v.5.3 (Cochrane Collaboration, Copenhagen, Denmark) and forest plots were created with 95% confidence interval (CI) for sensitivity and specificity for each study.

Studies were submitted to meta-analysis when three conditions were met: sample size was greater than 20; sensitivity and specificity were available for the index and the reference tests; and control group was included in the analysis. We used the two recommended random-effects hierarchical methods: the bivariate model and the hierarchical summary receiver operating characteristic (HSROC) model for performing the meta-analysis. The focus of the bivariate model is estimation of a summary point (summary sensitivity and specificity). HSROC model is on estimating an SROC curve²³. The summary estimates of sensitivity and specificity and 95% CI and the HSROC were calculated using OpenMeta-Analyst for windows 10 (open-source, cross-platform for advanced meta-analysis).

Investigations of heterogeneity

We investigated heterogeneity by adding antigen type as the covariate. The following approach was taken: Firstly, the variation in accuracy between IgG or IgM or IgG-IgM based LFIA, CLIA and ELISA serological testing was analysed (table 2). Then, the effect of the antigen type was investigated using subgroup meta-analysis in OpenMeta-analyst. I^2 values close to 0 % were considered as having no heterogeneity between studies; values close to 25 %, low heterogeneity; values close to 50 %, moderate heterogeneity; and values close to 75 %, high heterogeneity between studies²⁴.

Assessment of Publication bias

In this review we did not assess for reporting bias. The studies included in our meta-analysis showed a lot of heterogeneity; therefore assessments for reporting bias may not yield conclusive results. This was adopted from the approach used by Ochodo *et al.*, (2015)²⁵.

Results

Study inclusion

The results of the search and selection process are presented in Fig. 1. A total of 202 articles were identified. Among these, 40 were MedRxiv preprints and 162 were fully published articles from MEDLINE Complete (EBSCO) and PubMed. 2 articles were identified from other sources for example manual search. After abstract/title exclusion and removing duplicates, 74 articles were submitted to full text screening and 31 of these were included for the systematic review. Most articles were excluded because they did not present sufficient data hence it was not possible to extract data to construct 2 × 2 table and 1 article was excluded because it was only available in Chinese. A total of 29 articles describing the results of 99 independent studies (99 data sets^{19, 23 and 57} investigating IgG or IgM or IgG-IgM based LFIA, CLIA and ELISA serological test respectively) were eligible for the meta-analysis.

Characteristics of the studies

The general characteristics of the included articles are presented in table 1. All the published articles (n=14) included in the review were published in 2020 because COVID-19 is an emerging disease. The 17 unpublished articles were MedRxiv preprints which have been submitted to different journals for publication. Twenty five articles included in the review had a case-control design, comparing a group of well-defined cases with a group of healthy controls or controls with diseases or COVID-19 rRT-PCR negative patients and only six studies were cross sectional studies. One study had no control group and was excluded in the meta-analysis²⁶. Most of the studies (n=22) were conducted in China where the COVID-19 pandemic began and 3 studies were conducted in Italy whilst, USA whilst UK, Denmark, Germany, Spain and Japan each conducted one study.

Study ID	Country	Antibody type	Antigen type	Commercial	Reference standard	Index test	Control group/comparison group
Kai-Wang To ²⁶ P,CS	China	IgG and IgM	S and N	Inhouse	rRT-PCR	ELISA	No control
Cassaniti ³⁰ P,CS	Italy	IgG, IgM and IgG-IgM	S	Commercial	rRT-PCR	POC LFIA	Patients with fever and respiratory syndrome/RT-PCR negative
Duchuan Lin ³¹	China	IgG, IgM and IgG-IgM	N	Inhouse	Epidemiological risk/clinical features/rRT-PCR	CLIA	Healthy individuals and tuberculosis patients
Jie Xiang ³²	China	IgG, IgM and IgG-IgM		Commercial	rRT-PCR	ELISA and POC LFIA	Healthy individuals
Li Guo ¹¹ P	China	IgM	N	Inhouse	Deep sequencing and rRT-PCR	ELISA	Adult patients with acute lower respiratory tract infections (ALRTIs)
Rui Liu ²⁸ CS	China	IgM	N	Inhouse	rRT-PCR		COVID-19 rRT-PCR negative patients
Wanbing Liu ³³ P	China	IgG, IgM and IgG-IgM	S and N	Commercial	rRT-PCR	ELISA	Healthy individuals
Xuefei Cai ³⁴	China	IgG, IgM and IgG-IgM	S	Inhouse	rRT-PCR	Peptide-based Magnetic CLIA	Mixed diseases and Healthy controls
Yu bao Pan ³⁵ CS	China	IgG, IgM and IgG-IgM		Commercial	rRT-PCR	POC LFIA	COVID-19 rRT-PCR negative patients
Yujiao Jin ³⁶ P	China	IgG, IgM and IgG-IgM	S-N	Commercial	rRT-PCR	CLIA	Patients with suspected SARS-CoV-2 infection but with negative rRT-PCR results
Zhao ³⁷ P	China	IgG and IgM	S	Commercial	Chest CT images/Epidemiological history/Clinical diagnosis/rRT-PCR	ELISA	Healthy individuals
Zhengtu Li ³⁸ P	China	IgG, IgM and IgG-IgM	S	Commercial	rRT-PCR	POC LFIA	Healthy individuals
Rongqing Zhao ²⁷	China	IgG-IgM	S	Inhouse	Not clear but all cases were confirmed COVID-19 patients	ELISA	Healthy individuals (samples collected before and during the COVID-19 pandemic)
Pingping Zhang ³⁹	China	IgG-IgM	S	Inhouse	rRT-PCR	POC LFIA	COVID-19 rRT-PCR negative patients

Paradiso ⁴⁰	Italy	IgG-IgM	S	Commercial	rRT-PCR	POC LFIA	Patients with Covid-19 disease orienting-symptoms but rRT-PCR negative
Huan Ma ⁴¹	China	IgA, IgG, IgM, IgG-IgM and Ab	S and N	Inhouse	rRT-PCR	CLIA	Healthy individuals, COVID-19 suspected individuals and Mixed disease group
Qian ¹⁴	China	IgG and IgM	S-N	Commercial	rRT-PCR	CLIA	Healthy individuals and Hospitalised individuals
Ling Zhong ^{42 P}	China	IgG and IgM		Inhouse	rRT-PCR	CLIA and ELISA	Healthy individuals
Jiajia Xie ^{43 P, CS}	China	IgG and IgM	E-N	Commercial	Chest CT images/Epidemiological history/Clinical diagnosis/rRT-PCR	CLIA	Clinically confirmed COVID-19 rRT-PCR negative patients
Infantino ^{44 P}	Italy	IgG and IgM	S-N	Commercial	rRT-PCR	CLIA	Mixed diseases patients and blood donors Pre-COVID-19
Adams ⁴⁵	UK	IgG, IgM and IgG-IgM	S	Inhouse (ELISA) and Commercial (LFIA)	rRT-PCR	ELISA and POC LFIA	Healthy blood and ICU cerebral organ donors before the COVID-19 pandemic
Lassaunière ⁴⁶	Denmark	IgA, IgG and Ab	S	Commercial	rRT-PCR	ELISA and POC LFIA	Healthy individuals and mixed diseases patients (Including Acute respiratory tract infections caused by other corona viruses and non-corona viruses
Qiang Wang ^{47 P}	China	IgG, IgM and IgG-IgM		Commercial	Chest CT images/Epidemiological history/Clinical diagnosis/rRT-PCR	ELISA and POC LFIA	COVID-19 clinical negative mixed diseases patients
Fei Xiang ^{32 P}	China	IgG and IgM	N	Commercial	rRT-PCR	ELISA	Healthy blood donors or from patients with other disease hospitalized
Bin Lou ⁴⁸	China	IgG, IgM and Ab	S and N	Commercial	rRT-PCR	ELISA, CLIA and POC LFIA	Healthy Individuals

Lei Liu ⁴⁹	China	IgG-IgM	N	Commercial	rRT-PCR	ELISA	Randomly-selected ordinary patients and healthy blood donors
Imai ⁵⁰	Japan	IgG, IgM and IgG-IgM		Commercial	rRT-PCR	POC LFIA	Non-COVID-19 patients (from April to October 2019)
Pérez-García ⁵¹	Spain	IgG, IgM and IgG-IgM		Commercial	rRT-PCR	POC LFIA	Healthy individuals (samples collected before the COVID-19 pandemic)
Zhenhua Chen ^{13 P}	China	IgG	N	Inhouse	rRT-PCR	POC LFIA	Clinically suspicious for the presence of anti-SARS-CoV-2
Dohla ^{52 P, CS}	German	IgG, IgM and IgG-IgM		Commercial	RT-qPCR	POC LFIA	COVID-19 RT-qPCR negative patients
Burbelo ²⁹	USA	Ab	S and N	Inhouse	RT-PCR	LIPS	Subjects with COVID-19-like symptoms or household contacts of persons with COVID-19 (not tested by PCR), and blood donors who donated samples before 2018.
Key							
<ul style="list-style-type: none"> • Studies with P superscripts were published articles and without P superscripts were MedRxiv preprints. • Studies with CS superscripts are cross sectional studies and without CS superscripts are case control studies. • IgG-IgM means that either one of them or both were detected in serum. • Ab means total antibodies. 							

Table 1
The general characteristics of the studies included in the review.

Most articles (n=26) included in the review clearly stated that the gold standard nucleic acid tests (rRT-PCR or deep sequencing) were used as the reference standard. However, five articles used a combination of epidemiological risk, clinical features, chest CT images and rRT-PCR. In one article the reference standard used was not stated but all the patients in the study were COVID-19 patients ²⁷.

Point-of-care (POC) lateral flow immunoassays (LFIA) were used in 14 articles, CLIA were used in 9 articles and ELISA were used in 13 articles. We did not identify articles using FIA that met our inclusion criteria. One study did not specify the serological assay used and it was excluded from the review ²⁸. One study used a LIPS which is performed in solution, thus maintaining the native antigen conformation ²⁹. Most of the serological assay test kits were commercial (n=21) and 12 were in-house. Three SARS-CoV-2 antigens, Spike protein (S), nucleocapsid protein (N) and envelope protein (E) were used together or separately in studies included in the review. The spike protein and nucleocapsid were used as the antigen in 9 articles and 6 articles respectively. Five articles

used both S and N as the antigens separately. In three articles S and N antigens (S-N) were used together as the antigen. In one article N and E antigens (N-E) were used together as the antigen. In seven articles the name of antigen used was not given.

Methodological quality of included studies

The methodological quality of the included studies for the IgG or IgM or IgG-IgM based LFIA, CLIA and ELISA summarised across all studies are shown in **Figures 2b, 3b and 4b**. **Figures 2a, 3a and 4a** show for the risk of bias and applicability concerns summary results for the LFIA, CLIA and ELISA individual studies respectively. None of the studies included in this review had low risk of bias in all four QUADAS-2 domains. Generally case control studies were of high risk of bias and high concern in the patients and timing and flow domains and cross sectional studies were of low risk of bias and low concern in all domains.

Patient selection domain

Generally most studies included were at risk of bias and had high concerns regarding applicability. Studies were mostly case control studies and they did not include a consecutive or random series of participants implying that the patients that were included are not representative for clinical use. All thirteen ELISA studies were at high risk of bias and had high concerns regarding applicability. For CLIA all the nine studies included had high risk of bias and only one cross sectional study had low applicability concerns. Generally LFIA had more studies (n=4) with low risk of bias and applicability concerns in the patient selection domain because there were 4 LFIA cross sectional studies.

Index test domain

In the index test domain, studies generally had a low risk of bias (13/14, 5/9 and 9/13 for LFIA, CLIA and ELISA respectively). This was because most studies had a pre-specified threshold (cut-off value to decide whether a test is positive or negative). The studies that had high risk of bias did not have a pre-specified threshold and in two studies the risk of bias could not be determined as it was not clear whether the threshold was pre-specified or not. Likewise, studies generally had low applicability concerns in the index test domain (12/14, 5/9 and 8/13) for LFIA, CLIA and ELISA respectively because they used commercial index tests.

Reference standard domain

Like the index test domain, studies generally had a low risk of bias (10/14, 8/9 and 10/13 for LFIA, CLIA and ELISA respectively) in the reference standard domain. Generally the studies were of low applicability concern, 10/14, 8/9 and 11/13 for LFIA, CLIA and ELISA respectively.

Flow and timing domain

All the CLIA studies (n=9) and ELISA studies (n=13) were at high risk of bias in the flow and timing domain. These studies were all case control studies. Most of the LFIA studies were also at a high risk of bias however four cross sectional studies LFIA studies were at low risk of bias.

Quantitative synthesis and meta-analysis

Firstly, we considered performance of the LFIA devices using RT-PCR-confirmed cases as the reference standard. The forest plots in **Figure 5** show the sensitivity, specificity range, and heterogeneity for the three IgG or IgM or IgG-IgM based LFIA detecting COVID-19 across the included studies. Overall, the sensitivity varied widely across studies in contrast to the specificity which did not vary much except for 2 studies, Yunbao Pan, 2020 and Qiang Wang, 2020, which had the lowest and second lowest specificities respectively. Amongst the IgG based LFIA tests (n=17) the sensitivity estimates ranged from 0.14 (95% CI 0.09-0.21) (Imai, 2020) to 1.00 (95% CI 0.77-1.00) (Qiang Wang, 2020) and specificity estimates ranged from 0.41 (95% CI 0.21-0.64) (Yunbao Pan, 2020) to

1.00 (95% CI 0.97-1.00) (Bin Lou, 2020) (**Figure 5a**). For the IgM based LFIA tests (n=16) the sensitivity estimates ranged from 0.05 (95% CI 0.01-0.18) Adams assay 4 to 1.00 (95% CI 0.77-1.00) (Qiang Wang, 2020) and specificity estimates ranged from 0.64 (95% CI 0.41-0.83) (Yunbao Pan, 2020) to 1.00 (95% CI 0.94-1.00) (Adams assays 4 and 5) (**Figure 5b**). For the IgG-IgM based LFIA tests (n=24) the sensitivity estimates ranged from 0.18 (95% CI 0.08-0.34) (Cassaniti, 2020) to 1.00 (95% CI 0.77-1.00) (Qiang Wang, 2020), with most of the studies having sensitivities over 0.55 and specificity estimates ranged from 0.36 (95% CI 0.17-0.59) (Yunbao Pan, 2020) to 1.00 (95% CI 0.94-1.00) (Adams assays 2 and 3) (**Figure 5c**)

We then considered performance of the different IgG or IgM or IgG-IgM based CLIA test using RT-PCR-confirmed cases as the reference standard (**Figures 6a, 6b and 6c**). Considering any positive result (IgM positive, IgG positive or both), CLIA serological tests achieved sensitivity ranging from 0.48 (95% CI 0.29-0.68%) (Yujiao Jin, 2020) to 1.00 (95% CI 0.79-1.00) with most studies being between 0.80 and 1. The specificity was over 0.80 in most tests except for two tests, one IgG based test and one IgM based test which had the lowest 0.00 (95% CI 0.00-0.009) and second lowest 0.15 (95% CI 0.06-0.30) specificities respectively.

Lastly, we evaluated the performance of the different IgG or IgM or IgG-IgM based ELISA tests using RT-PCR-confirmed cases as the reference standard (**Figures 7a, 7b and 7c**). The sensitivities and specificities were generally high, ranging from 0.80 to 1.00 and 0.95 to 1.00 in most studies. For all the IgG based ELISA tests (n=10) the sensitivity estimates ranged from 0.65 (95% CI 0.57-0.72) (Zhao, 2020) to 1.00 (95% CI 0.79-1.00) (Kai-Wang To, 2020) and specificity estimates from 0.86 (95% CI 0.51-0.89) to 1.00 (95% CI 0.98-1.00) (Ling Zhong, 2020) (**Figure 7a**). In the IgM based tests (n=11), the sensitivity and specificity in the individual studies ranged from 0.44 (95% CI 0.32-0.58) (Jie Xiang, 2020) to 1.00 (95% CI 0.77-1.00) (Qiang Wang, 2020) and 0.69 (95% CI 0.57-0.80) (Qiang Wang, 2020) to 1.00 (95% CI 0.99-1.00) (Ling Zhong, 2020), respectively (**Figure 7b**). The sensitivity across the 5 studies included in the IgG-IgM based ELISA tests ranged from 0.80 (95% CI 0.74-0.85) (Wanbing Liu, 2020) to 0.87 (95% CI 0.77-0.94) (Rongqing Zhao, 2020). On the other hand, specificity across the 5 studies ranged from 0.97 (95% CI 0.92-0.99) (Lei Liu, 2020) to 1.00 (95% CI 0.98-1.00) (Rongqing Zhao, 2020) (**Figure 7c**).

We also constructed the SROC curves for all the three antibody based serological tests, figure 8, however we did not calculate the area under the ROC (AUROC). From the SROC we visually assessed heterogeneity between the different tests. Diagonal line indicated useless tests and the best tests were clustered further up to the top left hand corner.

The bivariate model and the hierarchical summary receiver operating characteristic curve (HSROC) model were performed to evaluate the diagnostic accuracy of the serological tests. The outputs of the meta-analysis (bivariate and HSROC parameter estimates, as well as the summary values of sensitivity and specificity) are presented in **Table 2 and Figure 9**. The pooled sensitivity for the IgG, IgM and IgG-IgM based LFIA tests were 0.5856, 0.4637 and 0.6886 respectively compared to RT-PCR. The pooled sensitivity for the IgG and IgM based CLIA tests were 0.9311 and 0.8516 respectively compared to RT-PCR. The pooled sensitivity for the IgG, IgM and IgG-IgM based ELISA tests were 0.8292, 0.8388 and 0.8531 respectively compared to RT-PCR. All the tests had high specificities ranging from 0.9693 to 0.9991 compared to RT-PCR. The estimated SROC curves for bivariate models are not presented.

HSROCs were also to visually access the overall performance of the diagnostic tests, access the overall diagnostic accuracy of the tests and compare the diagnostic accuracy of the different tests used for diagnosing COVID-19 in the review (**Figure 9**). The overall diagnostic test accuracy was measured by the closeness of the curve to the top left corner which represents high sensitivity and specificity. The closer the curve was to the upper left hand corner, the better the diagnostic accuracy⁵³. From figure 9 it can be observed that ELISA and CLIA have better diagnostic accuracy compared to LFIA and IgG-IgM based ELISA tests have the best overall diagnostic test accuracy.

Test type	Antibody type	Number of studies/tests	Sensitivity (95 %-CI)	Specificity (95 %-CI)	Correlation
LFIA	IgG	17	0.5856 (0.4397-0.7179)	0.9896 (0.9561-0.9976)	-0.4454
CLIA	IgG	9	0.9311 (0.9309-0.9312)	0.9757 (0.9757-0.9758)	-0.511
ELISA	IgG	10	0.8292 (0.7416-0.8915)	0.9948 (0.9675-0.9992)	-0.1709
LFIA	IgM	16	0.4637 (0.3016-0.6339)	0.9734 (0.9275-0.9905)	-0.7925
CLIA	IgM	10	0.8516 (0.7356-0.9221)	0.9693 (0.855-0.9941)	-0.7074
ELISA	IgM	11	0.8388 (0.7307-0.909)	0.9991 (0.9778-1)	-0.7247
LFIA	IgG-IgM	24	0.6886 (0.5878-0.7742)	0.9757 (0.9466-0.9892)	0.1011
CLIA	IgG-IgM	3	-	-	-
ELISA	IgG-IgM	5	0.8531 (0.7851-0.9023)	0.9901 (0.9287-0.9987)	-0.6771

Table 2

Summary estimates of test accuracy.

We identified one study (Burbelo, 2020) reporting total antibody (Ab) based luciferase immunoprecipitation assay system (LIPS) using N and S antigens with sensitivities and specificities of 0.91 (95 % CI 0.77-0.99) and 1.00 (0.80-1.00) and 1.00 (0.92-1.00) and 1.00 (0.92-1.00) respectively. We also identified studies reporting other Ab based serological assays and IgA based serological assays but results are not reported in this review.

Heterogeneity investigations

Generally high overall I^2 values above 85 %, which indicate high heterogeneities, were observed for both the sensitivities and specificities when we performed antigen subgroup meta-analysis with the exception of IgG-IgM based ELISA. IgG-IgM based ELISA had an overall sensitivity I^2 value of 52.12 % which is considered moderate heterogeneity and overall specificity I^2 value of 0 % which is considered to be low heterogeneity. However it should be noted that only 5 studies were included for this subgroup meta-analysis. Overall I^2 values for sensitivities and specificities heterogeneities for the antigen type subgroup meta-analysis are shown in **table 3**. We did not investigate heterogeneity for LFIA because most studies included in the review did not specify the type of antigen they used in their serological tests.

Detailed results of heterogeneity for the different antigen type sensitivities and specificities for each test type and antibody type combination are presented in **Additional file 2**.

Test type	Antibody type	Heterogeneity (I^2)	
		Sensitivity	Specificity
LFIA	IgG	-	-
	IgM	-	-
	IgG-IgM	-	-
CLIA	IgG	93.56 %	86.5 %
	IgM	93.42 %	95.17 %
	IgG-IgM	-	-
ELISA	IgG	78.07 %	84.97 %
	IgM	85.47 %	90.08 %
	IgG-IgM	52.12 %	0 %

Table 3

Overall antigen type subgroup meta-analysis heterogeneity

Discussion

COVID-19 is a major healthcare challenge globally. One key aspect of limiting SARS-CoV-2 spread is to ensure early and accurate diagnosis of the viral infection. In this study, we performed a meta-analysis to evaluate the diagnostic accuracy of IgG and IgM based serological assays offered to detect antigens against RT-PCR positive SARS-CoV-2 patients. The meta-analysis showed that all serological assays yielded high specificities ranging from 0.9693 (95% CI 0.855–0.9941) to 0.9991 (95% CI 0.9778–1) in comparison to RT-PCR. The sensitivities of all serological assays varied widely across studies. COVID-19 serological assays rely on antibodies binding to SARS-CoV-2 nucleocapsid protein, spike protein, or spike protein fragments (i.e. receptor binding domain). Some tests or test devices even use a combination of the N and S proteins and protein fragments. This may have resulted in the inconsistencies observed in the different serological assays⁵⁴. Overall IgG-IgM based ELISA had superior diagnostic accuracy with sensitivity and specificity of 0.8531 (95% CI 0.7851–0.9023) and 0.9991 (95% CI 0.9778–1). IgG based CLIA had the highest sensitivity 0.9311 (95% CI 0.9309–0.9312). The POC LFIA had the lowest sensitivities compared to both CLIA and LFIA across the different serological assays. The IgM based serological assays had the lowest sensitivities compared IgG based serological tests in each respective test method. Low antibody concentrations and especially timing of the IgM based serological assays could potentially explain the low ability of the IgM based tests to identify people infected with SARS-CoV-2. Immediately after a person is infected, antibodies may not have been developed yet or too late IgM antibodies may have decreased or disappeared⁵⁵.

Generally combined IgG-IgM based serological assays seem to be a better choice in terms of sensitivity than measuring either antibody type alone, though in this review we did not estimate the pooled sensitivity of the IgG-IgM based CLIA due to the limited number of studies. Subgroup meta-analysis using Open Meta-Analyst showed that IgM based serological assays that use the S antigen are more sensitive than IgM based serological assays that used the N antigen-based tests probably due to higher sensitivity and earlier immune response to the S antigen³³. However subgroup meta-analysis showed that IgG based serological assays that use N antigen are more sensitive than IgG based serological assays that use S antigen.

The sensitivity and specificity estimates from this review were used to indicate the consequences of testing when the test is being used in clinical practice. Prevalence estimates were calculated using metaDTA⁵⁶ to predict how many patients in practice you would expect to have true positive, false positive, true negative and false negative results for a given prevalence based on the meta-analysis results giving the results some clinical context. Using IgG-IgM based ELISA test which had the best overall diagnostic accuracy as an example, one can see that with 1000 patients and a COVID-19 prevalence of 50% we would expect 432 (95% CI 394–486) patients to test positive for SARS-CoV-2, of which 427 (95% CI 393–451) will be true positives (are diseased and test positive) and 5 (95% CI 1–35) will be false positives (are not diseased but test positive). We also noted 568 (95% CI 514–606) patients test negative for COVID-19, of which 495 (95% CI 465–499) were true negatives (are not diseased and test negative) and 73 (95% CI 49–107) were false negatives (are diseased but test negative).

A prevalence of 1% and 1000 patients would mean that 19 (95% CI 9–79) patients test positive for SARS-CoV-2, 9 (95% CI 8–9) of these will be true positives and 10 (95% CI 1–70) false positives and 981 (95% CI 921–991) patients will test negative for SARS-CoV-2, 980 (95% CI 920–989) of these will be true negatives and 1 (95% CI 1–2) patient will be false negative. This means that the false positive will be treated for COVID-19 while they have another disease, thus delaying their final diagnosis and subsequent treatment. In the false-negatives, COVID-19 diagnosis will be missed or delayed and the patients will not be quarantined and they will thus spread the SARS-CoV-2 to other patients in the hospital/clinic. Nonetheless, despite calculations like these providing insight in the consequences of testing, they should be taken with caution.

Using the inferences for LIFA using the IgG-IgM based test [sensitivity 0.6886 (0.5878–0.7742) and specificity 0.9757 (0.9466–0.9892)]. We see that for 1000 patients and a COVID-19 prevalence of 50% we would expect 356 (95% CI 299–414) patients to test positive for SARS-CoV-2, 324 (95% CI 294–387) of which will be true positives and 12 (95% CI 5–27) will be false positives. We would also expect 644 (95% CI 586–701) patients to test negative for the COVID-19, 488 (95% CI 473–495) of which will be true negatives and 156 (95% CI 113–206) are false negatives. A prevalence of 1% and 1000 patients would mean 31 (95% CI 16–62) patients test positive for SARS-CoV-2, 7 (95% CI 6–8) of these will be true positives and 24 (95% CI 10–54) false positives and 969

(95% CI 938–984) patients to test negative for COVID, 966 (95% CI 936–980) of these will be true negatives and 3 (95% CI 2–4) are false negatives. Compared to IgG-IgM based ELISA, the IgG-IgM based LFIA has higher rates of false positives and false negatives.

Limitation

Serological tests for SARS-CoV-2 have accuracy issues that deserve attention. They measure specific antibody responses which may take some weeks to develop after disease onset reducing the sensitivities of the assays. If blood samples were collected during the early stage of the infection, they may produce false negative results. They do not directly detect the presence of the virus. Further, antibodies may be present when SARS-CoV-2 is no longer present giving false positive case diagnosis. Moreover, since the identity of the N protein of SARS-CoV-2 and SARS-CoV reached up to 91.2%, there is probability of a cross reaction between the N protein of SARS-CoV-2 and antibodies against other human coronaviruses. Other molecules including interferon, rheumatoid factor and non-specific IgM may cause false positive results ⁴².

Most studies included in the meta-analysis were case-control studies. These may be easy to perform in a laboratory setting than cross-sectional designs, but their results are less representative for clinical practice. The performance of diagnostic tests very much depends on the population in which the test is being used. Future studies should therefore be prospective cross-sectional studies including a consecutive sample of presenting patients.

Conclusion

Given the poor performances of the current LFIA devices we recommend more research to develop highly sensitivity and specific POC LFIA that are adequate for most individual patient applications and attractive for large seroprevalence studies. The use of CLIA and ELISA for diagnosis have high sensitivity and are comparable to using RT-PCR. They may be calibrated to be specific for detecting and quantifying SARSCoV-2 IgM and IgG. More serological data should be collected to elucidate the clinical and epidemiological utility of IgG and IgM serological measurements to detect symptomatic and asymptomatic cases of COVID-19.

Abbreviations

COVID-19	Coronavirus disease 2019
SARS	Severe acute respiratory syndrome
SARS-CoV-2	Severe acute respiratory syndrome Coronavirus 2
STARD	Standards for the reporting of diagnostic accuracy studies
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies-2
rRT-PCR	Real-time reverse-transcriptase polymerase chain reaction
PRISMA-DTA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic test accuracy
DTA	Diagnostic test accuracy
RT-qPCR	Quantitative real-time polymerase chain reaction
POC	Point-of-care
LFIA	Lateral flow immunoassays
CLIA	Chemiluminescence enzyme immunoassay
FIA	Fluorescence
EISA	Enzyme-linked immunoassay
LIPS	Luciferase immunoprecipitation assay systems
S	SARS-CoV-2 spike protein
N	SARS-CoV-2 nucleocapsid protein
E	SARS-CoV-2 envelope protein
HSROC	Hierarchical summary receiver operating characteristic
TP	True positive
FN	False negative
FP	False positive
TN	True negative

Appendix 1

Search strategy in PubMed

Search	Query	Items found
#20	Search (((((COVID-19[Title/Abstract] OR SARS-CoV-2[Title/Abstract] OR 2019-nCoV[Title/Abstract] OR Wuhan Coronavirus[Title/Abstract])) AND (((((((((((Serologic test) OR Serologic method) OR Serological test) OR Serological method) OR Serodiagnosis) OR Immunodiagnosis) OR Immunological test) OR Immunological method) OR Antibody detection) OR Antigen detection) OR IgM) OR IgG) OR Immunochromatography)	78
#19	Search (((((((((((Serologic test) OR Serologic method) OR Serological test) OR Serological method) OR Serodiagnosis) OR Immunodiagnosis) OR Immunological test) OR Immunological method) OR Antibody detection) OR Antigen detection) OR IgM) OR IgG) OR Immunochromatography	822733
#18	Search Immunochromatography	69624
#17	Search IgG	143941
#16	Search IgM	66041
#15	Search Antigen detection	82456
#14	Search Antibody detection	99123
#13	Search Immunological method	39637
#12	Search Immunological test	456120
#11	Search Immunodiagnosis	450622
#10	Search Serodiagnosis	184574
#9	Search Serological method	24322
#8	Search Serological test	191599
#7	Search Serologic method	16939
#6	Search Serologic test	184758
#5	Search (((COVID-19[Title/Abstract] OR SARS-CoV-2[Title/Abstract] OR 2019-nCoV[Title/Abstract] OR Wuhan Coronavirus[Title/Abstract]	6734
#4	Search Wuhan Coronavirus[Title/Abstract]	14
#3	Search 2019-nCoV[Title/Abstract]	547
#2	Search SARS-CoV-2[Title/Abstract]	1898
#1	Search COVID-19[Title/Abstract]	5985

Appendix 2. Rating of QUADAS-2 items

1. Patient selection

1a. Risk of bias, signalling questions

- Was a case-control design avoided?
 - Case-control designs, especially if they include healthy controls, carry a high risk of bias. Therefore, all case-control studies are automatically rated to be of high risk of bias in the overall judgement.
- Was a consecutive or random sample of patients enrolled?
- Did the study avoid inappropriate exclusions?
- Overall judgement:
 - Case-control studies were always rated as having a high risk of bias.

- Cross-sectional studies: only low risk of bias if the other two signalling questions are answered with 'yes'. If one of the questions was answered 'no', then high risk of bias. Otherwise 'unclear'.

1b. Concerns regarding applicability: this concerns the extent to which the patients (both cases and controls) that were included in a study are representative for the patients which will receive these serology tests.

- Is there concern that the included patients do not match the review question?
 - All case-control studies are by default rated 'high concern'. All cross-sectional studies are by default 'low concern', except when the used case definition was not very clear.

2. Index test

2a. Risk of bias, signalling questions

- Were the index test results interpreted without knowledge of the results of the reference standard?
- If a threshold was used, was it pre-specified? By selecting the cut-off value with the highest sensitivity and/or specificity, researchers artificially optimise the accuracy of their tests, which may lead to an overestimation of sensitivity and specificity.
- Overall judgement:
 - If the second question is answered with 'yes', overall judgement is automatically 'low'. This is because the first question is usually not reported or answered with 'yes'.
 - If the latter is rated as 'unclear', then overall is also 'unclear'; if the latter receives a 'no', the overall risk is considered 'high'.

2b. Concerns regarding applicability: this concerns the extent to which the index test evaluated is representative of the tests that will be used in practice.

- Are there concerns that the index test, its application or interpretation deviate from the review question?
 - All in-house tests are automatically rated as 'high concern'.

Risk of bias and concerns regarding applicability should be assessed for each test separately.

3. Reference standard

3a. Risk of bias, signalling questions

- Is the reference standard likely to correctly classify the target condition?
 - Assumption: This will likely be the case for case-control studies that use the 'correct' case definitions
 - This is also likely for cross-sectional studies which use the 'correct' case definitions.
- Were the reference standard results interpreted without knowledge of the results of the index test?
 - Assumption: This will likely be the case for most case-control studies, but only if serology was not part of the case definition.
 - For cross-sectional studies, this should be explicitly stated.
- Overall judgement risk of bias:
 - Case-control studies with clear case definitions were rated as having a 'low' risk of bias.
 - case-control studies with unclear/wrong case definitions rated as 'unclear' Or 'high risk' of bias respectively
 - Cross-sectional studies with a clear case definition and the second question answered with 'yes': low risk of bias.
 - Otherwise 'unclear'.

3b. Concerns regarding applicability: Are there concerns that the target condition as defined by the reference standard does not match the review question?

- If serology is included in the case definition, there is an incorporation bias and thus a high risk of bias.
- If a case-control study uses clear criteria and does not include serology in these criteria: 'low' concern of bias.

4. Risk of bias regarding flow and timing, signalling questions

- Was there an appropriate interval between index test(s) and reference standard?
 - We expected that studies with a cross-sectional design conducted most tests on a date sufficiently close to the final diagnosis. If we had reason to suspect that the patient status changed between the time of testing and the time of diagnosis, we rated this as 'no'.
 - For case-control studies, this was always rated as 'no', because serology was always determined after the case definitions were defined, sometimes with a long delay.
- Did patients receive the same reference standard?
- Were all patients included in the analysis?
 - This was rated 'no' for all case-control studies.
- Overall judgement:
 - Case-control studies were always rated as having a high risk of bias.
 - For cross-sectional studies, we perceived a low risk of bias if all three questions were answered with 'yes'; a high risk of bias was perceived if at least one of them was answered 'no'. All other possibilities were rated as 'unclear'.

(European Centre for Disease Prevention and Control. A systematic literature review on the diagnostic accuracy of serological tests for Lyme borreliosis. Stockholm: ECDC; 2016.)

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

The raw data (data extraction results) will be provided for sharing after reasonable request.

Conflict of interest

The authors declare that they have no competing interests.

Funding

This research was commissioned by the National Institute for Health Research (NIHR) Global Health Research programme (16/136/33) using UK aid from the UK Government. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Authors' contributions

AV and TM conceived and designed the study protocol. AV, HM, HM, TLMJ and MK conducted a comprehensive literature search, screened the literature and extracted the data. AV did the statistical analysis. TN, FM, SR and TM provide guidance. All authors read and approved the final version of the manuscript.

Acknowledgements

Not applicable

References

1. Ludvigsson JF. Systematic review of COVID-19 in children shows milder cases and a better prognosis than adults. *Acta Paediatr* [Internet]. 2020 Jun 14 [cited 2020 May 27];109(6):1088–95. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/apa.15270>
2. Ren L-L, Wang Y-M, Wu Z-Q, Xiang Z-C, Guo L, Xu T, et al. Identification of a novel coronavirus causing severe pneumonia in human. *Chin Med J (Engl)* [Internet]. 2020 May 5 [cited 2020 May 28];133(9):1015–24. Available from: <http://journals.lww.com/10.1097/CM9.0000000000000722>
3. Situation Summary | CDC [Internet]. [cited 2020 May 27]. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/summary.html>
4. NIH clinical trial of investigational vaccine for COVID-19 begins | National Institutes of Health (NIH) [Internet]. [cited 2020 May 27]. Available from: <https://www.nih.gov/news-events/news-releases/nih-clinical-trial-investigational-vaccine-covid-19-begins>
5. Coronavirus disease (COVID-19) Situation Report-127.
6. Coronavirus disease 2019 (COVID-19) Situation Report-63 HIGHLIGHTS.
7. Guo L, Ren L, Yang S, Xiao M, Chang D, Yang F, et al. *Clinical Infectious Diseases* Clinical Infectious Diseases ® 2020;XX(X):1-8. [cited 2020 May 28]; Available from: <https://academic.oup.com/cid/advance-article-abstract/doi/10.1093/cid/ciaa310/5810754>
8. To KKW, Tsang OTY, Leung WS, Tam AR, Wu TC, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis*. 2020 May 1;20(5):565–74.
9. Wikramaratna P, Paton RS, Ghafari M, Lourenco J. Estimating false-negative detection rate of SARS-CoV-2 by RT-PCR. *medRxiv* [Internet]. 2020 Apr 14 [cited 2020 May 28];2020.04.05.20053355. Available from: <https://doi.org/10.1101/2020.04.05.20053355>
10. Li D, Wang D, Dong J, Wang N, Huang H, Xu H, et al. False-negative results of real-time reverse-transcriptase polymerase chain reaction for severe acute respiratory syndrome coronavirus 2: Role of deep-learning-based ct diagnosis and insights from two cases. *Korean J Radiol*. 2020 Apr 1;21(4):505–8.
11. Guo L, Ren L, Yang S, Xiao M, Chang D, Yang F, et al. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19). *Clin Infect Dis* [Internet]. 2020 Mar 21 [cited 2020 Apr 2]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/32198501>
12. Zhang W, Du RH, Li B, Zheng XS, Yang X Lou, Hu B, et al. Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. *Emerg Microbes Infect* [Internet]. 2020 Jan 1 [cited 2020 May 28];9(1):386–9. Available from: <https://www.tandfonline.com/doi/full/10.1080/22221751.2020.1729071>
13. Chen Z, Zhang Z, Zhai X, Li Y, Lin L, Zhao H, et al. Rapid and Sensitive Detection of anti-SARS-CoV-2 IgG, Using Lanthanide-Doped Nanoparticles-Based Lateral Flow Immunoassay. *Anal Chem*. 2020 May 5;
14. Qian C, Zhou M, Cheng F, Lin X, Gong Y, Xie X, et al. Development and Multicenter Performance Evaluation of The First Fully Automated SARS-CoV-2 IgM and IgG Immunoassays. *medRxiv*. 2020 Apr 21;2020.04.16.20067231.
15. The critical role of laboratory medicine during coronavirus disease 2019 (COVID-19) and other viral outbreaks in: *Clinical Chemistry and Laboratory Medicine (CCLM)* - Ahead of print [Internet]. [cited 2020 May 28]. Available from:

<https://www.degruyter.com/view/journals/cclm/ahead-of-print/article-10.1515-cclm-2020-0240/article-10.1515-cclm-2020-0240.xml>

16. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies The PRISMA-DTA Statement. *JAMA - J Am Med Assoc.* 2018 Jan 23;319(4):388–96.
17. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Chapter 10 Analysing and Presenting Results [Internet].* [cited 2020 May 31]. Available from: <http://srdta.cochrane.org/>.
18. Zeller H, Van Bortel W. A systematic literature review on the diagnosis accuracy of serological tests for Lyme borreliosis. 2016.
19. Steingart KR, Henry M, Laal S, Hopewell PC, Ramsay A, Menzies D, et al. A systematic review of commercial serological antibody detection tests for the diagnosis of extrapulmonary tuberculosis. Vol. 83, *Postgraduate Medical Journal.* BMJ Publishing Group Ltd; 2007. p. 705–12.
20. Volpe Chaves CE, do Valle Leone de Oliveira SM, Venturini J, Grande AJ, Sylvestre TF, Poncio Mendes R, et al. Accuracy of serological tests for diagnosis of chronic pulmonary aspergillosis: A systematic review and meta-analysis. Agarwal R, editor. *PLoS One [Internet].* 2020 Mar 17 [cited 2020 May 31];15(3):e0222738. Available from: <https://dx.plos.org/10.1371/journal.pone.0222738>
21. Abba K, Deeks JJ, Olliaro PL, Naing C-M, Jackson SM, Takwoingi Y, et al. Rapid diagnostic tests for diagnosing uncomplicated *P. falciparum* malaria in endemic countries . *Cochrane Database Syst Rev.* 2011 Jul 6;(7).
22. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. Quadas-2: A revised tool for the quality assessment of diagnostic accuracy studies. Vol. 155, *Annals of Internal Medicine.* American College of Physicians; 2011. p. 529–36.
23. Takwoingi Y, Riley RD, Deeks JJ. Meta-analysis of diagnostic accuracy studies in mental health. *Evid Based Ment Health.* 2015 Nov 1;18(4):103–9.
24. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Vol. 327, *British Medical Journal.* British Medical Journal Publishing Group; 2003. p. 557–60.
25. Ochodo EA, Gopalakrishna G, Spek B, Reitsma JB, van Lieshout L, Polman K, et al. Circulating antigen tests and urine reagent strips for diagnosis of active schistosomiasis in endemic areas. Vol. 2015, *Cochrane Database of Systematic Reviews.* John Wiley and Sons Ltd; 2015.
26. Kai-Wang To K, Tak-Yin Tsang O, Leung W-S, Raymond Tam A, Wu T-C, Christopher Lung D, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. 2020 [cited 2020 May 31]; Available from: www.thelancet.com/infection
27. Zhao R, Li M, Song H, Chen J, Ren W, Feng Y, et al. Serological diagnostic kit of SARS-CoV-2 antibodies using CHO-expressed full-length SARS-CoV-2 S1 proteins. *medRxiv [Internet].* 2020 Mar 27 [cited 2020 May 31];2020.03.26.20042184. Available from: <https://www.medrxiv.org/content/10.1101/2020.03.26.20042184v1>
28. Liu R, Liu X, Han H, Shereen MA, Niu Z, Li D, et al. The comparative superiority of IgM-IgG antibody test to real-time reverse transcriptase PCR detection for SARS-CoV-2 infection diagnosis. *medRxiv.* 2020 Mar 30;2020.03.28.20045765.
29. Burbelo PD, Riedo FX, Morishima C, Rawlings S, Smith D, Das S, et al. Detection of Nucleocapsid Antibody to SARS-CoV-2 is More Sensitive than Antibody to Spike Protein in COVID-19 Patients. *medRxiv.* 2020 Apr 24;2020.04.20.20071423.
30. Cassaniti I, Novazzi F, Giardina F, Salinaro F, Sachs M, Perlini S, et al. Performance of VivaDiag COVID-19 IgM/IgG Rapid Test is inadequate for diagnosis of COVID-19 in acute patients referring to emergency room department. *J Med Virol [Internet].* 2020 Apr 8 [cited 2020 May 28];jmv.25800. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmv.25800>
31. Lin D, Liu L, Zhang M, Hu Y, Yang Q, Guo J, et al. Evaluations of serological test in the diagnosis of 2019 novel coronavirus (SARS-CoV-2) infections during the COVID-19 outbreak. *medRxiv [Internet].* 2020 Mar 30 [cited 2020 May 31];2020.03.27.20045153. Available from: <https://www.medrxiv.org/content/10.1101/2020.03.27.20045153v1>
32. Xiang J, Yan M, Li H, Liu T, Lin C, Huang S, et al. Evaluation of Enzyme-Linked Immunoassay and Colloidal Gold-Immunochemical Assay Kit for Detection of Novel Coronavirus (SARS-Cov-2) Causing an Outbreak of Pneumonia

- (COVID-19). medRxiv [Internet]. 2020 Mar 1 [cited 2020 May 31];2020.02.27.20028787. Available from: <https://doi.org/10.1101/2020.02.27.20028787>
33. Liu W, Liu L, Kou G, Zheng Y, Ding Y, Ni W, et al. Evaluation of Nucleocapsid and Spike Protein-based ELISAs for detecting antibodies against SARS-CoV-2. *J Clin Microbiol*. 2020 Mar 30;58(6).
 34. Chen J, Hu J, Long Q, Deng H, Fan K, Liao P, et al. A Peptide-based Magnetic Chemiluminescence Enzyme Immunoassay for 1 Xue-fei. medrxiv.org [Internet]. 2020 Feb 25 [cited 2020 May 31];2020.02.22.20026617. Available from: <https://doi.org/10.1101/2020.02.22.20026617>
 35. Pan Y, Li X, Yang G, Fan J, Tang Y, Zhao J, et al. Serological immunochromatographic approach in diagnosis with SARS-CoV-2 infected COVID-19 patients. medRxiv. 2020 Mar 17;2020.03.13.20035428.
 36. Jin Y, Wang M, Zuo Z, Fan C, Ye F, Cai Z, et al. Diagnostic value and dynamic variance of serum antibody in coronavirus disease 2019. *Int J Infect Dis* [Internet]. 2020 [cited 2020 May 31];94:49–52. Available from: <https://doi.org/10.1016/j.ijid.2020.03.065>
 37. Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. *Clin Infect Dis* [Internet]. 2020 Mar 28 [cited 2020 Apr 2]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/32221519>
 38. Li Z, Yi Y, Luo X, Xiong N, Liu Y, Li S, et al. Development and Clinical Application of A Rapid IgM-IgG Combined Antibody Test for SARS-CoV-2 Infection Diagnosis. *J Med Virol* [Internet]. 2020 Feb 27 [cited 2020 Apr 2]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/32104917>
 39. Zhang P, Gao Q, Wang T, Ke Y, Mo F, Jia R, et al. Evaluation of recombinant nucleocapsid and spike proteins for serological diagnosis of novel coronavirus disease 2019 (COVID-19). medRxiv [Internet]. 2020 Mar 20 [cited 2020 May 31];2020.03.17.20036954. Available from: <https://www.medrxiv.org/content/10.1101/2020.03.17.20036954v1>
 40. Paradiso AV, Summa S De, Loconsole D, Procacci V, Sallustio A, Centrone F, et al. Clinical meanings of rapid serological assay in patients tested for SARS-Co2 RT-PCR. medRxiv. 2020 Apr 6;2020.04.03.20052183.
 41. Ma H, Zeng W, He H, Zhao D, Yang Y, Jiang D, et al. COVID-19 diagnosis and study of serum SARS-CoV-2 specific IgA, IgM and IgG by a quantitative and sensitive immunoassay. medRxiv. 2020 Apr 30;2020.04.17.20064907.
 42. Zhong L, Chuan J, Gong B, Shuai P, Zhou Y, Zhang Y, et al. Detection of serum IgM and IgG for COVID-19 diagnosis. Vol. 63, *Science China Life Sciences*. Science in China Press; 2020. p. 777–80.
 43. Xie J, Ding C, Li J, Wang Y, Guo H, Lu Z, et al. Characteristics of patients with coronavirus disease (COVID-19) confirmed using an IgM-IgG antibody test. *J Med Virol* [Internet]. 2020 May 7 [cited 2020 May 31];jmv.25930. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmv.25930>
 44. Infantino M, Grossi V, Lari B, Bambi R, Perri A, Manneschi M, et al. Diagnostic accuracy of an automated chemiluminescent immunoassay for anti-SARS-CoV-2 IgM and IgG antibodies: an Italian experience. *J Med Virol* [Internet]. 2020 May 10 [cited 2020 May 31];jmv.25932. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmv.25932>
 45. Crook D. Evaluation of antibody testing for SARS-CoV-2 using ELISA and lateral flow immunoassays. [cited 2020 May 31]; Available from: <https://doi.org/10.1101/2020.04.15.20066407>
 46. Lassaunière R, Frische A, Harboe ZB, Nielsen AC, Fomsgaard A, Krogfelt KA, et al. Evaluation of nine commercial SARS-CoV-2 immunoassays. medRxiv. 2020 Apr 10;2020.04.09.20056325.
 47. Wang Q, Du Q, Guo B, Mu D, Lu X, Ma Q, et al. A method to prevent SARS-CoV-2 IgM false positives in gold immunochromatography and enzyme-linked immunosorbent assays. *J Clin Microbiol*. 2020 Apr 10;58(6).
 48. Lou B, Li T, Zheng S, Su Y, Li Z, Liu W, et al. Serology characteristics of SARS-CoV-2 infection since the exposure and post symptoms onset. medRxiv. 2020 Mar 27;2020.03.23.20041707.
 49. Liu L, Liu W, Wang S, Zheng S. A preliminary study on serological assay for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in 238 admitted hospital patients. medRxiv. 2020 Mar 8;2020.03.06.20031856.
 50. Imai K, Tabata S, Ikeda M, Noguchi S, Kitagawa Y, Matuoka M, et al. Clinical evaluation of an immunochromatographic IgM/IgG antibody assay and chest computed tomography for the diagnosis of COVID-19. medRxiv. 2020 Apr 27;2020.04.22.20075564.

51. Garcia FP, Tanoira RP, Cabrera JPR, Serrano TA, Herruz PG, Gonzalez JC. Rapid diagnosis of SARS-CoV-2 infection by detecting IgG and IgM antibodies with an immunochromatographic device: a prospective single-center study. medRxiv. 2020 Apr 24;2020.04.11.20062158.
52. Döhla M, Boesecke C, Schulte B, Diegmann C, Sib E, Richter E, et al. Rapid point-of-care testing for SARS-CoV-2 in a community screening setting shows low sensitivity. Public Health. 2020 May 1;182:170–2.
53. Jones CM, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. Ann Thorac Surg. 2005 Jan 1;79(1):16–20.
54. Zhang X, Wu X, Wang D, Lu M, Hou X, Wang H, et al. Proteome-wide analysis of differentially-expressed SARS-CoV-2 antibodies in early COVID-19 infection. medRxiv [Internet]. 2020 May 2 [cited 2020 May 31];2020.04.14.20064535. Available from: <https://www.medrxiv.org/content/10.1101/2020.04.14.20064535v2?%253fcollection=>
55. Kontou PI, Braliou GG, Dimou NL, Nikolopoulos G, Bagos PG. Antibody tests in detecting SARS-CoV-2 infection: a meta-analysis. medRxiv [Internet]. 2020 Apr 25 [cited 2020 May 31];2020.04.22.20074914. Available from: <http://medrxiv.org/content/early/2020/04/25/2020.04.22.20074914.abstract>
56. Freeman SC, Kerby CR, Patel A, Cooper NJ, Quinn T, Sutton AJ. Development of an interactive web-based tool to conduct and interrogate meta-analysis of diagnostic test accuracy studies: MetaDTA. BMC Med Res Methodol. 2019 Apr 18;19(1).

Figures

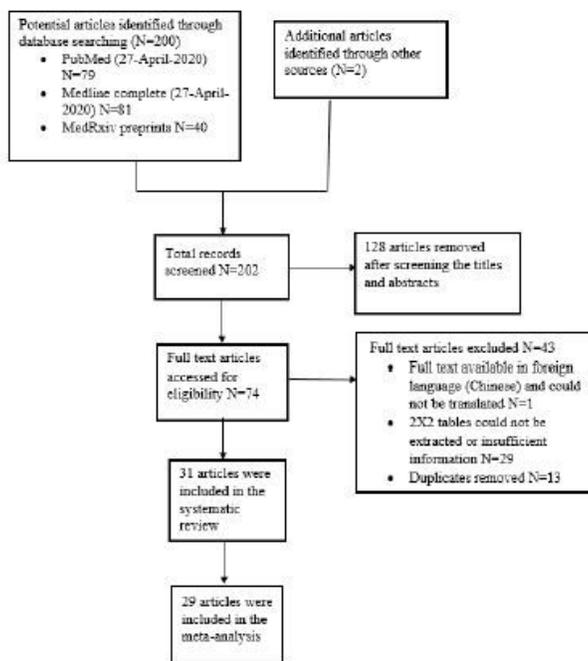


Figure 1

PRISMA flow diagram for selection of articles for meta-analysis.

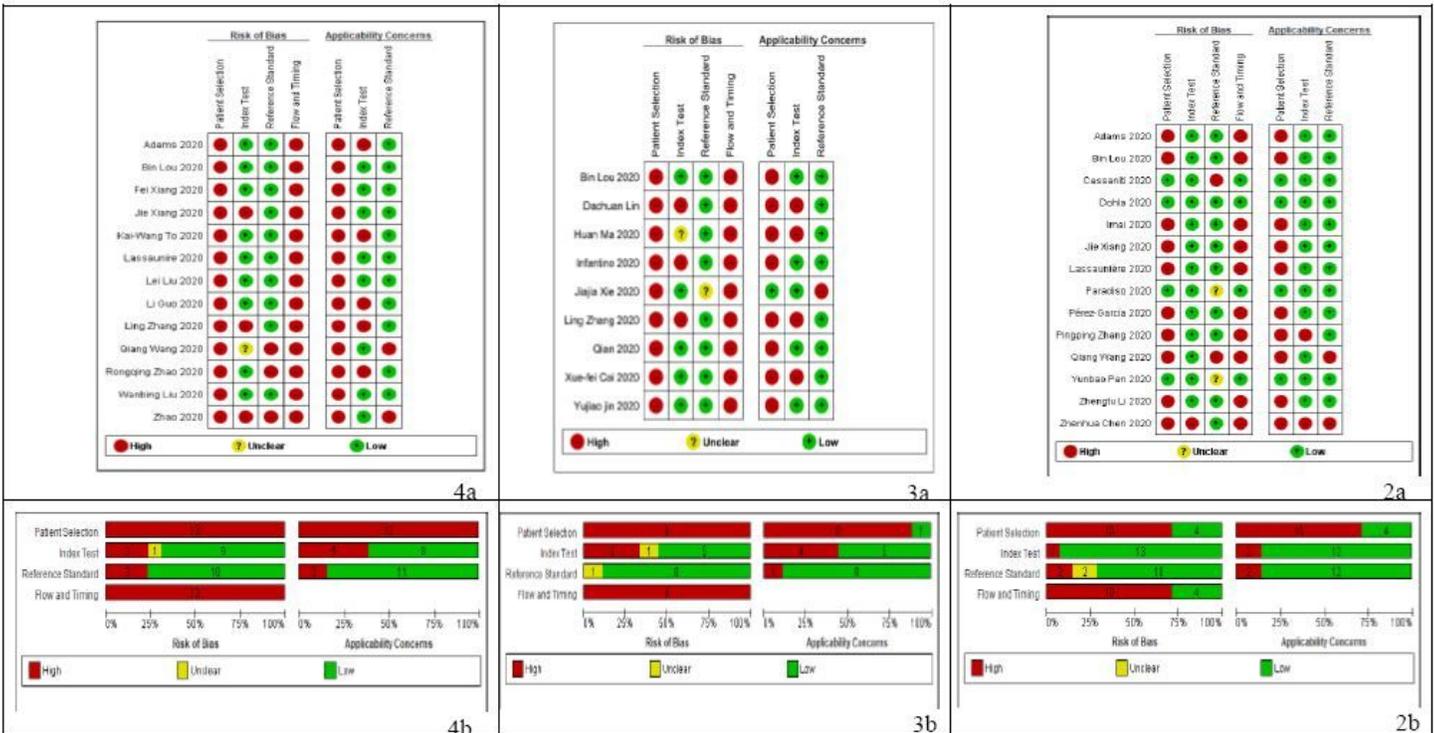


Figure 2. LFIA methodological quality summary table and graph.
 Figure 3. CLIA methodological quality summary table and graph.
 Figure 4. ELISA methodological quality summary table and graph.

- (a) Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.
- (b) QUADAS-2 bias assessment and QUADAS-2 applicability assessment item presented as percentages across all included studies. On the left Risk of bias graph and on the right applicability concerns graph.
- (c) Risk of bias and applicability concerns summary: review authors' Low: low risk of bias; high: high risk of bias; unclear: bias is unclear

Figure 2

Figure 2. Figure 3. Figure 4.

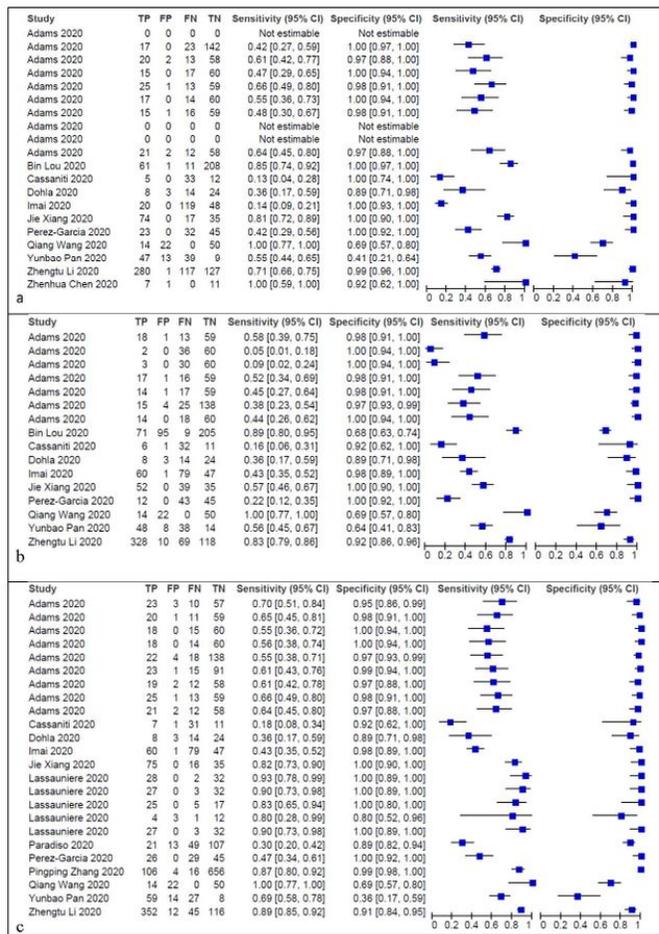


Figure 3

Forest plot of sensitivity, specificity and heterogeneity of serological LFIA diagnosis of COVID-19. 5a Forest plot for the IgG LFIA. 5b Forest plot for the IgM based LFIA. 5c Forest plot for the IgG-IgM based LFIA.

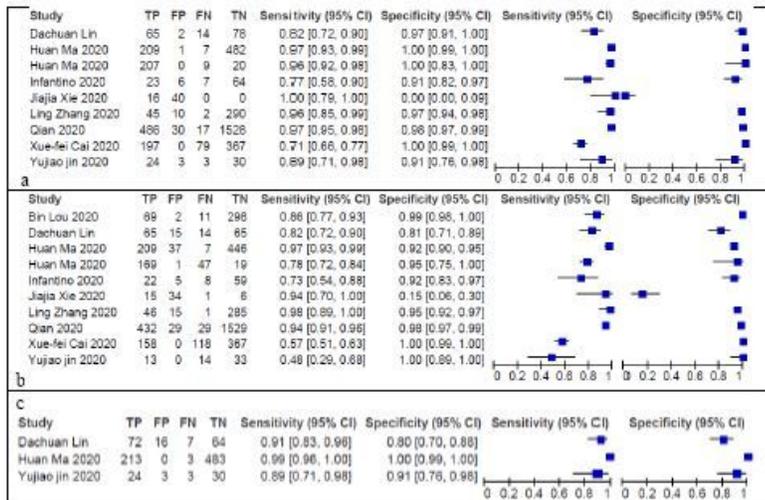


Figure 4

Forest plot of sensitivity, specificity and heterogeneity of serological CLIA diagnosis of COVID-19. 6a Forest plot for the IgG CLIA. 6b Forest plot for the IgM based CLIA. 6c Forest plot for the IgG-IgM based CLIA.

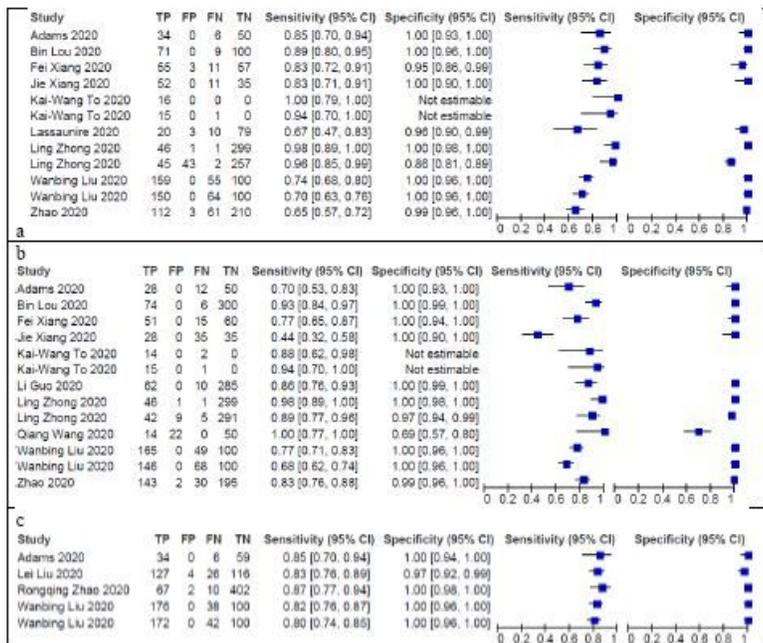


Figure 5

Forest plot of sensitivity, specificity and heterogeneity of serological ELISA diagnosis of COVID-19. 7a Forest plot for the IgG ELISA. 7b Forest plot for the IgM based ELISA. 7c Forest plot for the IgG-IgM based ELISA.

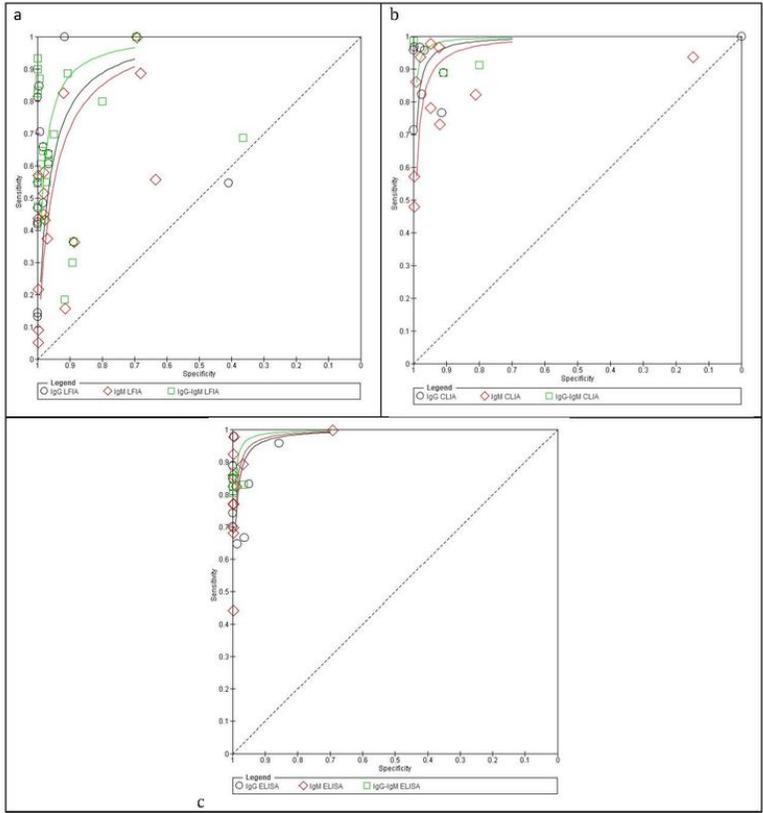


Figure 6

Summary ROC curves for the three antibody serological test groups. Every symbol reflects a 2×2 table, one for each test. One study may have contributed more than one 2×2 table. The curves are shown for the different test-types.

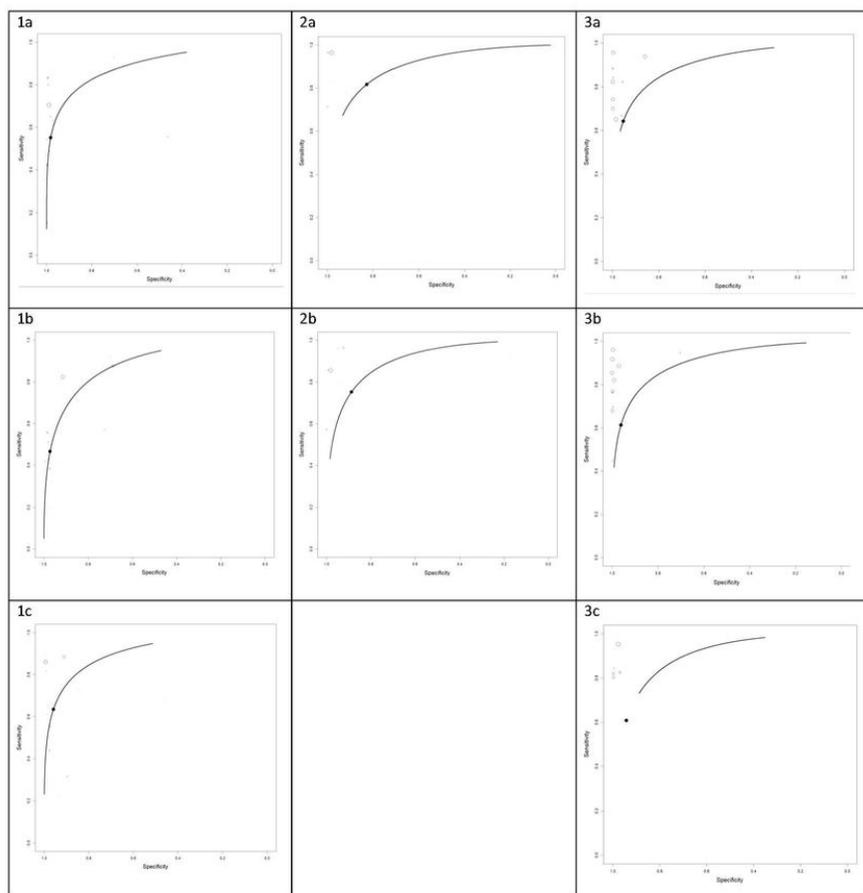


Figure 7

Hierarchical summary receiver operating characteristic (HSROC) curve obtained using OpenMeta-Analyst. Every circle represents the sensitivity and specificity estimates of individual studies in the meta-analysis and the size of the circle reflects the sample size. The black dots indicates summary points of sensitivity and specificity; HSROC curve is the line passing through summary point. The curve is the regression line that summarises the overall diagnostic accuracy. a HSROC for IgG serological tests; b HSROC for IgM serological tests and c HSROC IgG-IgM serological tests. 1 LFIA HSROC, 2 CLIA HSROC and 3 ELISA HSROC.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement12.pdf](#)
- [supplement13.pdf](#)