Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

International Journal of Infectious Diseases 104 (2021) 415–422

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Clinical applications of detecting IgG, IgM or IgA antibody for the diagnosis of COVID-19: A meta-analysis and systematic review

Mengyu Chen[a,1], Rundong Qin[b,1], Mei Jiang[b], Zhaowei Yang[b], Weiping Wen[a,c,**], Jing Li[b,*]

[a] Department of Otorhinolaryngology, Head and Neck Surgery, First Affiliated Hospital, Sun Yat-Sen University, Guangzhou, Guangdong, PR China
[b] Department of Allergy and Clinical Immunology, State Key Laboratory of Respiratory Disease, First Affiliated Hospital of Guangzhou Medical University, Guangzhou, PR China
[c] Guangzhou Key Laboratory of Otorhinolaryngology, Guangzhou, PR China

ARTICLE INFO

Article history:
Received 19 October 2020
Received in revised form 4 January 2021
Accepted 6 January 2021

Keywords:
COVID-19
SARS-CoV-2
Antibody tests
Specificity
Sensitivity
Diagnostic accuracy

ABSTRACT

Background: The coronavirus disease 2019 (COVID-19) pandemic has had a devastating impact worldwide, and timely detection and quarantine of infected patients are critical to prevent spread of disease. Serological antibody testing is an important diagnostic method used increasingly in clinics, although its clinical application is still under investigation.

Methods: A meta-analysis was conducted to compare the diagnostic performance of severe acute respiratory syndrome coronavirus-2 antibody tests in patients with COVID-19. The test results analysed included: (1) IgM-positive but IgG-negative (IgM^+IgG^-); (2) IgG-positive but IgM-negative (IgG^+IgM^-); (3) both IgM-positive and IgG-positive (IgM^+IgG^+); (4) IgM-positive without IgG information ($IgM^+IgG^{+/-}$); (5) IgG-positive without IgM information ($IgG^+IgM^{+/-}$); (6) either IgM-positive or IgG-positive (IgM^+ or IgG^+); and (7) IgA-positive (IgA^+).

Results: Sixty-eight studies were included. Pooled sensitivities for IgM^+IgG^-, IgG^+IgM^-, IgM^+IgG^+, $IgM^+IgG^{+/-}$, $IgG^+IgM^{+/-}$, and IgM^+ or IgG^+ were 6%, 7%, 53%, 68%, 73% and 79% respectively. Pooled specificities ranged from 98% to 100%. IgA^+ had a pooled sensitivity of 78% but a relatively low specificity of 88%. Tests conducted 2 weeks after symptom onset showed better diagnostic accuracy than tests conducted earlier. Chemiluminescence immunoassay and detection of S protein as the antigen could offer more accurate diagnostic results.

Discussion: These findings support the supplemental role of serological antibody tests in the diagnosis of COVID-19. However, their capacity to diagnose COVID-19 early in the disease course could be limited.
© 2021 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

The coronavirus disease 2019 (COVID-19) pandemic, caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), has affected more than 200 countries, with 15,785,641 confirmed cases and 640,016 deaths worldwide (World Health Organization, 2020). Timely detection and quarantine of infected patients are critical to prevent spread of the disease. Various diagnostic tests for COVID-19 have been reported (Beeching et al., 2020).

Virological testing to detect SARS-CoV-2 is often recommended for the diagnosis of COVID-19 as it provides the strongest evidence for the presence of the virus (Nuccetelli et al., 2020). SARS-CoV-2 RNA in respiratory samples can be detected by reverse transcription polymerase chain reaction (RT-PCR), which is the gold standard diagnostic test recommended by current guidelines (National Institutes of Health, 2020). However, various factors, including inappropriate specimen collection techniques, viral load, time since exposure and specimen source, have been reported to markedly affect the performance of RT-PCR assays, which could contribute to false-negative test results (Kucirka et al., 2020; Lin et al., 2020; Pan et al., 2020; Wang et al., 2020). Therefore, supplementary diagnostic tests are needed urgently.

Serological tests for specific antibodies against SARS-CoV-2, including immunoglobulin M (IgM), IgG and IgA antibodies, have been developed as supplementary diagnostic methods as they can provide information about recent or prior infection (Peeling et al., 2020). Although some studies have reported that serological tests

* Corresponding author at: Department of Allergy and Clinical Immunology, State Key Laboratory of Respiratory Disease, First Affiliated Hospital of Guangzhou Medical University, 151 Yan Jing Road, Guangzhou 510120, PR China.
** Corresponding author at: Department of Otorhinolaryngology, Head and Neck Surgery, First Affiliated Hospital of Sun Yat-sen University, No.58 Zhongshan Er Road, 510080 Guangzhou, PR China.
E-mail addresses: wenwp@mail.sysu.edu.cn (W. Wen), lijing@gird.cn (J. Li).
[1] These authors contributed equally to this paper.

https://doi.org/10.1016/j.ijid.2021.01.016
1201-9712/© 2021 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

had high sensitivity, ranging from 96.0% to 97.8%, and demonstrated improved diagnostic accuracy when combined with PCR (Deeks et al., 2020), high-quality evidence supporting the use of antibody tests in practice for COVID-19 is missing (Lisboa Bastos et al., 2020). Indeed, antibody subtype, antigen used in the serological test kit, detection time and method of measurement varied markedly between studies. Some studies detected both IgM and IgG and reported a positive result if either was positive, while other studies detected IgM or IgG individually. There is no consensus on the interpretation of antibody test results (Cheng et al., 2020). The

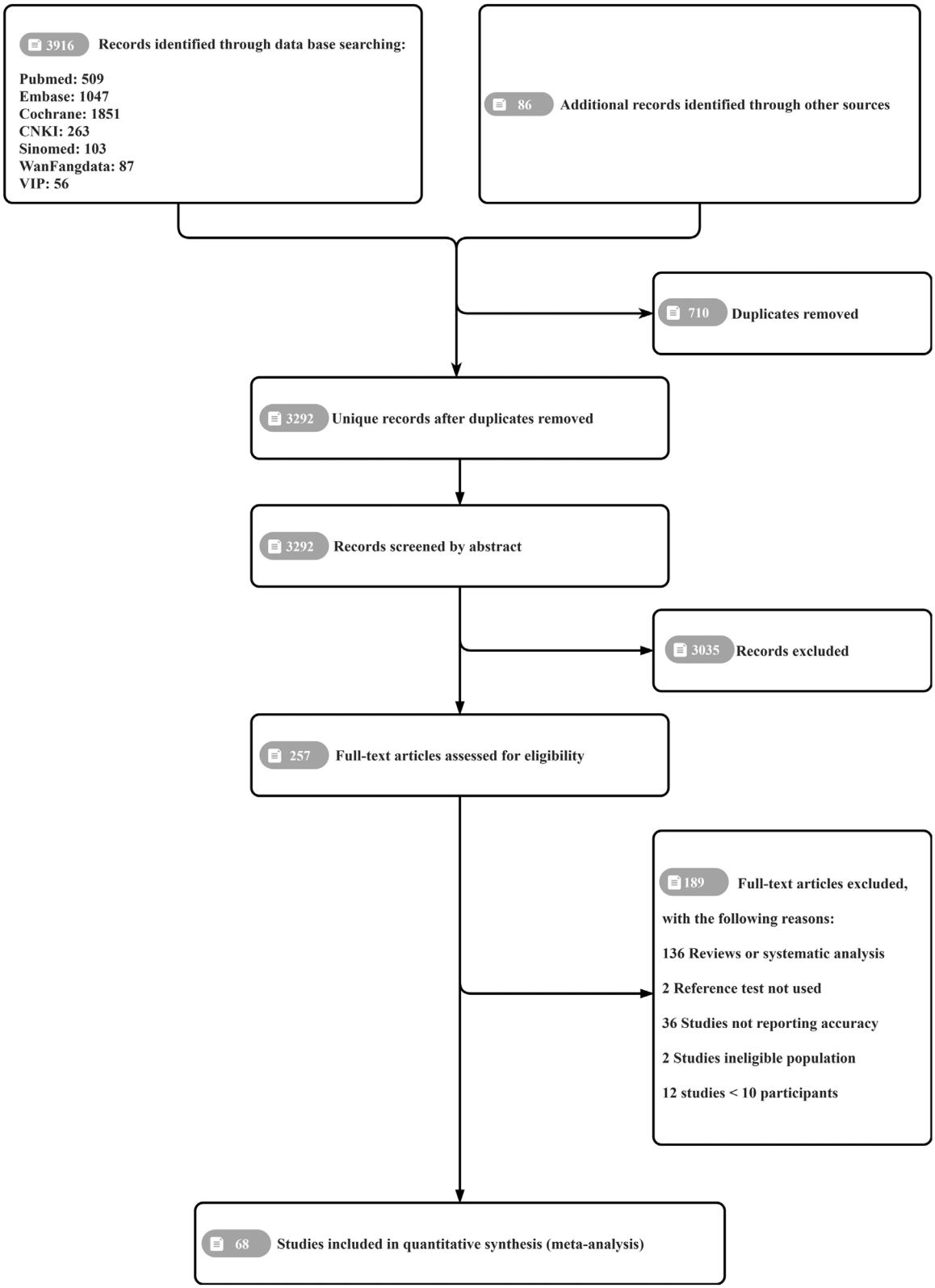

Figure 1. Prisma flow chart.

presence of IgM, IgG and IgA, either alone or in certain combinations, may be related to disease severity and immunization, which could affect diagnostic accuracy.

As such, this meta-analysis aimed to investigate the diagnostic effectiveness of SARS-CoV-2-specific antibodies stratified by different positive results, including: (1) IgM-positive but IgG-negative (IgM^+IgG^-); (2) IgG-positive but IgM-negative (IgG^+IgM^-); (3) both IgM-positive and IgG-positive (IgM^+IgG^+); (4) IgM-positive without IgG information ($IgM^+IgG^{+/-}$); (5) IgG-positive without IgM information ($IgG^+IgM^{+/-}$); (6) either IgM-positive or IgG-positive (IgM^+ or IgG^+); and (7) IgA-positive (IgA^+). For the first three panels, this study provided clear information regarding the presence of antibody types, while previous meta-analyses focused on the diagnostic accuracy of $IgM^+IgG^{+/-}$, $IgG^+IgM^{+/-}$, and IgM^+ or IgG^+ which only offer vague information (Caini et al., 2020; Deeks et al., 2020; Lisboa Bastos et al., 2020; Moura et al., 2020).

Methods

Search strategy

This meta-analysis followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (Moher et al., 2009). Pubmed, Medline, Embase, Cochrane Library, ICTRP, ClinicalTrials.gov, Medrxiv, Biorxiv, CNKI, Sinomed, Wan-Fangdata and Cqvip databases were searched. MESH terms and entry terms for concepts of COVID-19 (or SARS-CoV-2) and serological tests were searched in the titles and abstracts in each database. The simplified search formula was ((COVID-19[Title/Abstract] OR SARS-CoV-2[Title/Abstract]) AND (serological tests [Title/Abstract])). Only articles including human subjects and published between December 2019 and June 2020 were included in this meta-analysis, and no language restrictions were made. The detailed search strategy is shown in Table S1 (see online supplementary material).

The inclusion criteria were as follows: (1) patients with COVID-19 confirmed by RT-PCR, or by a combination of RT-PCR and clinical manifestation; (2) serological diagnostic tests with no method, antibody or antigen restrictions; and (3) prospective or retrospective studies comparing confirmed cases of COVID-19 against non-COVID-19 individuals or suspected cases of COVID-19 that could not be confirmed by microbiological tests.

The exclusion criteria were as follows: (1) case reports with a sample size <10; (2) publications without a 2×2 contingency table; (3) reviews, meta-analyses and systematic analyses; (4) studies focused on ineligible populations, such as pregnant women, elderly people and children; and (5) studies undertaken in a population surveillance setting.

Two independent researchers screened the publications by reading the titles and abstracts. In the case of disagreement, researchers discussed the full-text publications with a third researcher.

Data extraction

Data were extracted independently by two authors. Data extracted from each study included the first author's surname, age and sex of patients with COVID-19, days since symptom onset, test kit manufacturer, study design, reference standard, RT-PCR sample type, blood sample type, methods, antigen, and antibody types of antibody detection. True-positive, false-positive, false-negative and true-negative results were extracted to construct the 2×2 contingency table, and to estimate sensitivity and specificity. This study aimed to investigate the diagnostic accuracy of different antibody combinations, including: (1) IgM^+IgG^-; (2) IgG^+IgM^-; (3) IgM^+IgG^+; (4) $IgM^+IgG^{+/-}$; (5) $IgG^+IgM^{+/-}$; (6) IgM^+ or IgG^+; and (7) IgA^+.

Quality assessment

The methodological quality of included studies was assessed independently by two authors using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool (Whiting et al., 2011). The QUADAS tool consists of four key domains: patient, index test, reference standard, and flow and timing.

Statistical analysis

Estimated sensitivity and specificity based on the 2×2 contingency table were calculated for each individual study. All results are reported with 95% confidence intervals (CI). Data are summarized as paired Forest plots. A bivariate random-effects model was used for the meta-analysis as different studies had different cut-off values.

Summary receiver operating characteristic (ROC) curves were used to estimate the overall diagnostic efficiency of each antibody during different clinical courses. A random-effects logistic regression model was used to compare diagnostic accuracy between different antibodies, different antibody detection methods and different antigens.

Unobserved heterogeneity was quantified using I^2. $I^2 > 50\%$ was considered to indicate significant heterogeneity. Sources of heterogeneity were investigated by performing univariate meta-regression and subgroup analysis. Covariates including country, antigen, antibody detection method, detection time, blood sample type, disease severity, and whether tests were built in-house were considered to contribute to heterogeneity.

All statistical analyses were performed using the meta-analysis modules of STATA Version 16 (Stata Corp., College Station, TX, USA) (metandi and midas) (Dwamena, 2007; Harbord, 2008). Summary ROC curves were plotted using Review Manager 5 Version 5.3 (Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark). $P < 0.05$ was considered to indicate significance.

Study registration

PROSPERO CRD42020195898.

Results

Identification and characteristics of included studies

In total, 4002 records were identified through the database search. After removing 710 duplicates, the titles and abstracts of 3292 unique studies were screened, and 257 full-text articles were assessed for eligibility. Ultimately, 68 eligible publications were included in this review, comprising a total of 20,430 samples (8817 confirmed cases of COVID-19 and 2899 suspected cases). Figure 1 shows the PRISMA flow diagram of search and eligibility results. Of the included studies, 44 were conducted in China. Forty-six studies were published in English and the rest were published in Chinese. Nineteen studies investigated two or more methods for serological testing. Ten studies used tests that had been built in-house. All the included studies are cited in the supplementary references (see online supplementary material).

The included studies varied in terms of sample type (e.g. serum, plasma, venous whole blood, capillary blood), detection method [e.g. chemiluminescence immunoassay (CLIA), enzyme-linked immunosorbent assay (ELISA), lateral flow immunosorbent assay (LFIA), immunoblot, luciferase immunoprecipitation system assay] and antigen type [e.g. nucleocapsid protein (N protein), spike

protein (S protein), membrane protein (M protein)]. Most studies (48/68, 70.59%) used serum-based samples. Antigen information was only available in 36 of the included studies. Regarding the S protein, only 13 studies mentioned using the S1 subunit. All studies conducted RT-PCR as a reference standard. The essential characteristics of the included studies are presented in Tables S2 and S3 (see online supplementary material).

Quality assessment

Figure 2 presents the summary of quality assessment. For the patient selection domain, 54 (79%) studies had a high risk of bias and eight (12%) studies had a low risk of bias. The majority of studies had a case–control design and failed to select patients consecutively or randomly. Moreover, applicability was hindered due to inclusion without outpatients, asymptomatic patients, patients with other respiratory or autoimmune diseases, and healthy individuals. For the index test domain, 60 (88%) studies had high or unclear risk of bias as the blinding method was not clarified and the protocol was not stated when interpreting the results of the index tests and reference standard. For the reference standard domain, unclear risk of bias was seen in 31 (46%) studies, mainly related to unreported sample site and timing of RT-PCR. Thirty-seven (54%) studies were low risk as they stated that RT-PCR was conducted in accordance with national guidance or the guidance of the World Health Organization. For the flow and timing domain, 47 (69%) studies had high or unclear risk of bias due to the inclusion of multiple samples per participant and lack of stratification of the results by days since symptom onset.

Pooled diagnostic accuracy of antibody tests

Detailed results are presented in Table 1. The sensitivities of IgM^+IgG^-, IgG^+IgM^- and IgM^+IgG^+ were 6% (95% CI 4–8%), 7% (95% CI 5–11%) and 53% (95% CI 46–60%), respectively, while the sensitivities of $IgM^+IgG^{+/-}$, $IgG^+IgM^{+/-}$, and IgM^+ or IgG^+ were 68% (95% CI 62–73%), 73% (95% CI 69–77%) and 79% (95% CI 76–83%), respectively. In contrast, the specificity of the serological tests was high, ranging from 98% to 100%. For IgA, sensitivity was comparable to IgM^+ or IgG^+ (78%), but specificity was only 88%. Figure 3 compares the summary ROC curves of IgM^+IgG^-, IgG^+IgM^-, IgM^+IgG^+, and IgM^+ or IgG^+ with stratification by days since symptom onset. Except for IgM^+IgG^-, the area under the ROC curve peaked >14 days after symptom onset regardless of the antibody type. Figure 4 shows the summary ROC curve of IgA.

Analysis of heterogeneity

Meta-regression

Given that the results of pooled diagnostic accuracy presented a high I^2, a meta-regression analysis was performed to investigate the potential source of heterogeneity. The factors, including characteristics of included patients, blood sample type, detection method, antigen types and commercial availability, were included in the meta-regression models for each antibody panel. The results suggest that the marked heterogeneity may have resulted from all factors that were included in the analysis. Detailed results are presented in Table S3 (see online supplementary material).

The detailed subgroup analysis results are presented in Table 2.

Subgroup analysis by detection time

Detection time was classified into three groups: patients who received antibody testing within 1 week, 2 weeks or ≥3 weeks of symptom onset. Due to the lack of adequate studies, analysis was not performed for IgA^+. For patients who received antibody testing within 1 week of symptom onset, sensitivity ranged from 4% to 34%, with the best performance seen for IgM^+ or IgG^+ and the worst performance seen for IgG^+IgM^-.

Subgroup analysis by detection method

Detection methods were classified into three groups: CLIA, ELISA and LFIA. This analysis evaluated studies reporting IgM^+ or IgG^+ as this has the best diagnostic performance. The pooled sensitivities of CLIA, ELISA and LFIA were 86% (95% CI 73–94%), 83%

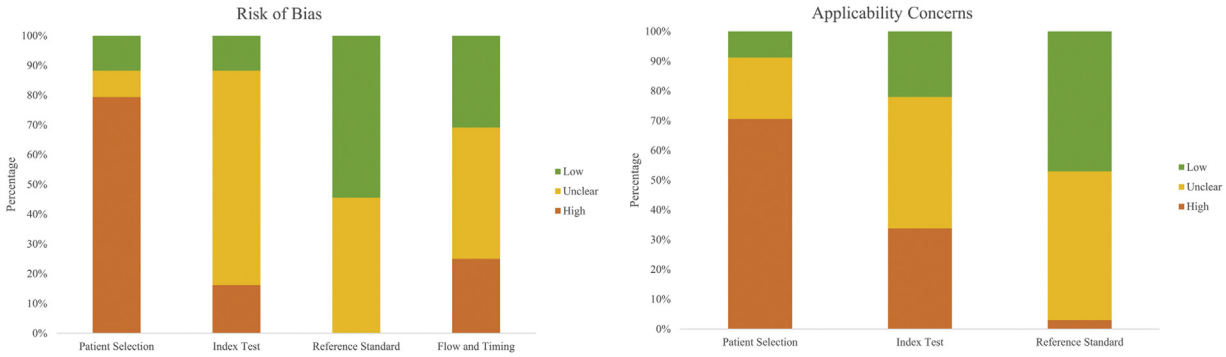

Figure 2. Summary of quality assessment.

Table 1
Pooled sensitivity, pooled specificity and area under the receiver operating characteristic curve (AUROC) of serological tests.

Types of antibody	Number of studies and samples	Sensitivity with 95% CI (%)	Specificity with 95% CI (%)	AUROC with 95% CI (%)
IgM^+IgG^-	28 studies, 6161 samples	6 (4–8)	99 (98–99)	0.64 (1.00–0.00)
IgG^+IgM^-	58 studies, 12,751 samples	7 (5–11)	99 (99–99)	0.91 (1.00–0.00)
IgM^+IgG^+	63 studies, 13,344 samples	53 (46–60)	100 (99–100)	0.94 (1.00–0.00)
$IgM^+IgG^{+/-}$	87 studies, 18,924 samples	68 (62–73)	98 (97–99)	0.96 (0.19–1.00)
$IgG^+IgM^{+/-}$	88 studies, 18,597 samples	73 (69–77)	99 (98–99)	0.97 (0.19–1.00)
IgM^+ or IgG^+	76 studies, 20,065 samples	79 (76–83)	98 (98–99)	0.97 (0.19–1.00)
IgA^+	6 studies, 934 samples	78 (67–85)	88 (82–92)	0.91 (0.88–0.93)

CI, confidence interval.

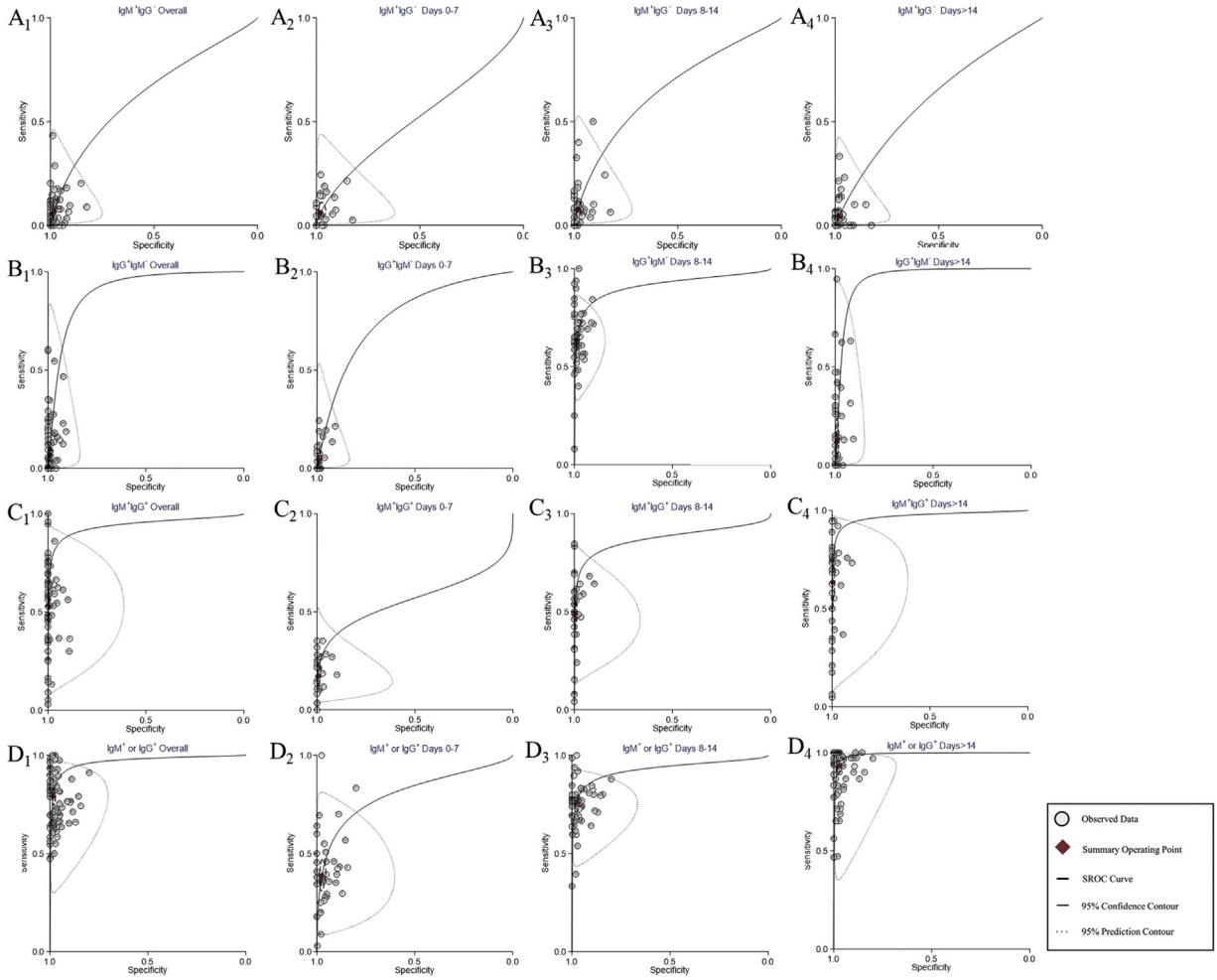

Figure 3. Summary receiver operating characteristic curves of different positive results for antibodies. (A) IgM^+IgG^-. (B) IgG^+IgM^-. (C) IgM^+IgG^+. (D) IgM^+ or IgG^+. A_1–D_1, antibodies were detected without stratification of detection time. A_2–D_2, antibodies were detected 0–7 days after symptom onset. A_3–D_3, antibodies were detected 8–14 days after symptom onset. A_4–D_4, antibodies were detected >14 days after symptom onset.

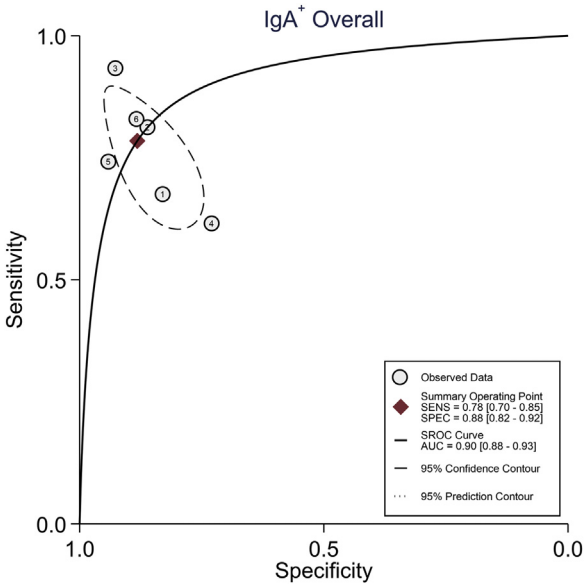

Figure 4. Summary receiver operating characteristic curve of IgA without stratification of detection time.

(95% CI 76–88%) and 75% (95% CI 71–79%), respectively. Due to data limitations, the overall results were pooled without stratification by detection time. The 95% CI overlapped between each test method. However, the sensitivity of LFIA was significantly lower compared with CLIA and ELISA ($P < 0.001$).

Subgroup analysis by antigen protein

Antigen protein was classified into three groups: S protein, N protein, and both S and N proteins. For IgM^+ or IgG^+, the pooled sensitivities of S protein, N protein, and both S and N proteins were 90% (95% CI 84–94%), 79% (95% CI 65–88%) and 88% (95% CI 83–91%), respectively. The pooled sensitivity was higher for S protein ($P < 0.001$).

Pooled diagnostic accuracy of patients with clinical suspicion of COVID-19 but negative RT-PCR results

Patients with clinical suspicion of COVID-19 had typical epidemiological history, signs and symptoms, but did not have a positive RT-PCR result. For these patients, the positive rates of IgM^+IgG^-, IgG^+IgM^-, IgM^+IgG^+, $IgM^+IgG^{+/-}$, $IgG^+IgM^{+/-}$, and IgM^+ or IgG^+ were 5% (95% CI 2–14%), 19% (95% CI 8–37%), 32% (95% CI 11–64%), 47% (95% CI 29–66%), 72% (95% CI 47–89%) and 81% (95% CI 59–92%), respectively. The results are presented in Table 2.

Table 2
Pooled sensitivity, pooled specificity, and area under the receiver operating characteristic curve (AUROC) of serological tests stratified by detection time, detection method and antigen protein.

Category	Antibody	Sensitivity with 95% CI (%)	Specificity with 95% CI (%)	AUROC with 95% CI (%)
Days 0–7	IgM^+IgG^- (25 studies, 3803 samples)	6 (4–10)	98 (96–99)	0.52 (1.00–0.00)
	IgG^+IgM^- (26 studies, 3904 samples)	4 (2–8)	99 (98–100)	0.77 (1.00–0.00)
	IgM^+IgG^+ (25 studies, 3803 samples)	17 (13–22)	100 (99–100)	0.56 (0.07–0.96)
	$IgM^+IgG^{+/-}$ (38 studies, 5710 samples)	34 (27–41)	97 (96–98)	0.81 (1.00–0.00)
	$IgG^+IgM^{+/-}$ (38 studies, 5794 samples)	25 (20–30)	98 (98–99)	0.82 (0.14–0.99)
	IgM^+ or IgG^+ (41 studies, 6708 samples)	39 (33–45)	97 (96–98)	0.80 (0.13–0.99)
Days 8–14	IgM^+IgG^- (36 studies, 5705 samples)	8 (5–11)	98 (97–99)	0.66 (1.00–0.00)
	IgG^+IgM^- (37 studies, 5876 samples)	7 (4–12)	99 (99–100)	0.91 (1.00–0.00)
	IgM^+IgG^+ (35 studies, 5630 samples)	49 (42–55)	100 (99–100)	0.88 (0.07–0.96)
	$IgM^+IgG^{+/-}$ (50 studies, 7759 samples)	66 (60–72)	98 (96–98)	0.94 (1.00–0.00)
	$IgG^+IgM^{+/-}$ (48 studies, 8014 samples)	64 (59–68)	99 (98–99)	0.92 (0.17–1.00)
	IgM^+ or IgG^+ (52 studies, 8966 samples)	76 (72–79)	98 (97–98)	0.93 (0.18–1.00)
>14 days	IgM^+IgG^- (36 studies, 5140 samples)	4 (3–7)	98 (97–98)	0.93 (0.18–1.00)
	IgG^+IgM^- (38 studies, 5387 samples)	12 (7–21)	98 (97–99)	0.61 (1.00–0.00)
	IgM^+IgG^+ (36 studies, 5139 samples)	64 (54–72)	99 (98–100)	0.95 (1.00–0.00)
	$IgM^+IgG^{+/-}$ (53 studies, 7855 samples)	81 (74–87)	100 (99–100)	0.97 (0.19–1.00)
	$IgG^+IgM^{+/-}$ (51 studies, 7634 samples)	89 (85–92)	98 (96–98)	0.98 (1.00–0.00)
	IgM^+ or IgG^+ (54 studies, 8225 samples)	93 (90–95)	99 (98–99)	0.99 (0.20–1.00)
Methods	CLIA (10 studies, 3077 samples)	86 (73–94)	100 (97–100)	0.99 (1.00–0.00)
	ELISA (14 studies, 4217 samples)	83 (76–88)	99 (97–100)	0.97 (1.00–0.00)
	LFIA (50 studies, 12621 samples)	75 (71–79)	97 (96–98)	0.95 (0.19–1.00)
Antigens	S protein (14 studies, 3617 samples)	90 (84–94)	99 (98–100)	0.99 (0.98–1.00)
	N protein (9 studies, 2829 samples)	79 (65–88)	97 (94–99)	0.96 (0.94–0.98)
	S and N protein (9 studies, 5936 samples)	88 (83–91)	98 (96–99)	0.96 (0.94–0.97)
Suspected	IgM^+IgG^- (8 studies, 1447 samples)	5 (2–14)	99 (96–100)	0.74 (0.70–0.77)
	IgG^+IgM^- (8 studies, 1446 samples)	19 (8–37)	100 (94–100)	0.78 (0.75–0.82)
	IgM^+IgG^+ (8 studies, 1426 samples)	32 (11–64)	100 (96–100)	0.98 (0.96–0.99)
	$IgM^+IgG^{+/-}$ (10 studies, 1858 samples)	47 (29–66)	100 (96–100)	0.95 (0.93–0.97)
	$IgG^+IgM^{+/-}$ (10 studies, 1858 samples)	72 (47–89)	99 (96–100)	0.99 (0.97–0.99)
	IgM^+ or IgG^+ (16 studies, 2856 samples)	81 (59–92)	99 (97–99)	0.98 (0.98–1.00)

CI, confidence interval; CLIA, chemiluminescence immunoassay; ELISA, enzyme-linked immunosorbent assay; LFIA, lateral flow immunosorbent assay; Suspected, patients with clinical suspicion of coronavirus disease 2019 had typical epidemiological history, signs and symptoms but did not have a positive reverse transcription polymerase chain reaction result.

Figures S2–S5 (see online supplementary material) present the paired Forest plots of each analysis. Figures S6–S9 (see online supplementary material) show the summary ROC curve for each analysis. Figure S10 (see online supplementary material) summarizes the publication bias of each analysis. Most analyses were performed without risk of bias.

Table 3 summarizes the positive predictive value (PPV) and negative predictive value of serological total antibody tests for COVID-19 with prevalence rates of 5%, 10% or 20%. If 1000 people received total antibody serological tests with a prevalence rate of COVID-19 of 10%, 67 people would be diagnosed correctly and 33 people would be diagnosed incorrectly. Additionally, 879 people would be identified correctly as healthy individuals, and 21 people would be incorrectly diagnosed with COVID-19.

Table 3
Positive and negative predictive values of serological testing for IgM^+ or IgG^+.

Category	PPV5	PPV10	PPV20	NPV5	NPV10	NPV20
Overall	67.52%	81.44%	90.80%	98.88%	97.67%	94.92%
Days 0–7	40.63%	59.09%	76.47%	96.80%	93.47%	86.41%
Days 8–14	66.67%	80.85%	90.48%	98.73%	97.35%	94.23%
>14 days	70.99%	80.85%	90.48%	99.63%	99.21%	98.25%
CLIA	81.90%	90.53%	95.56%	99.26%	98.45%	96.59%
ELISA	81.37%	90.22%	95.40%	99.10%	98.13%	95.88%
LFIA	56.82%	73.53%	86.21%	98.66%	97.22%	93.95%
N antigen	58.09%	74.53%	86.81%	98.87%	97.65%	94.87%
S antigen	82.57%	90.91%	95.74%	99.47%	98.89%	97.54%
N and S antigen	69.84%	83.02%	91.67%	99.36%	98.66%	97.03%

PPV5, positive predictive value at a prevalence of 5%; PPV10, positive predictive value at a prevalence of 10%; PPV20, positive predictive value at a prevalence of 20%; NPV5, negative predictive value at a prevalence of 5%; NPV10, negative predictive value at a prevalence of 10%; NPV20, negative predictive value at a prevalence of 20%; overall, without stratification by detection time; CLIA, chemiluminescence immunoassay; ELISA, enzyme-linked immunosorbent assay; LFIA, lateral flow immunosorbent assay.

Discussion

This meta-analysis attempted to provide comprehensive evidence of the diagnostic accuracy of serum antibodies in the detection of COVID-19. It addressed the question of which positive results could provide optimal diagnostic accuracy. Of the six panels of positive antibody results of IgM and IgG, all had favourable specificity. However, the panels exhibited marked difference in sensitivity. The panel of IgM^+ or IgG^+ showed the highest sensitivity at 79%, followed by $IgG^+IgM^{+/-}$ (73%), $IgM^+IgG^{+/-}$ (68%), IgM^+IgG^+ (53%), IgG^+IgM^- (7%) and IgM^+IgG^- (6%). Additionally, serological tests conducted 2 weeks after symptom onset had higher diagnostic accuracy compared with tests conducted sooner after symptom onset. The subgroup analysis further found that CLIA could offer more accurate diagnosis than either ELISA or LFIA, and use of the S protein or both the S and N proteins as the antigen for antibody preparation could also contribute to improved diagnostic accuracy. Furthermore, IgA seems to be a potential candidate for antibody testing for the diagnosis of COVID-19. This study found that detection of IgM^+ or IgG^+ in patients with COVID-19 had favourable accuracy, especially when testing was conducted 2 weeks after symptom onset. These findings support the supplemental role of serological antibody tests in the diagnosis of COVID-19. However, their capacity for diagnosis of COVID-19 early in the disease course could be limited.

To date, the Emergency Use Authorization has allowed various types of serological test kits to flood the healthcare market (Food and Drug Administration, 2020). In clinical practice, confusion exists when interpreting the results of serological tests containing

two or more types of antibodies. However, previous meta-analysis shed light on the results for $IgM^{+}IgG^{+/-}$, $IgG^{+}IgM^{+/-}$, and IgM^{+} or IgG^{+} and the comparison of detection methods in the possible application of population surveillance, without paying attention to the interpretation of results for individuals (Caini et al., 2020; Deeks et al., 2020; Lisboa Bastos et al., 2020; Moura et al., 2020). In the current study, although the data for the three panels with clear information about the presence of antibodies showed poor sensitivity compared with the three vague panels, all of the antibody tests showed favourable specificity. Based on the high specificity of the serological tests, this study provides evidence for the application of tests with vague antibody information, indicating that a positive result for IgG or IgM should be considered as a positive serological test. The three panels with clear information about the presence of antibodies in each sample showed poor diagnostic performance. This might reflect the diversity of immunological reactions against SARS-CoV-2. Differences in the timing of immunoglobulin isotype switching and differentiation of IgM-positive or IgG-positive B cells to plasma cells could explain the observed patterns of variable IgM/IgG responses. IgG antibodies against SARS-CoV-2 were not confined to a later stage after infection, and could be observed earlier than IgM. Usually, IgM and IgG were detected simultaneously after SARS-CoV-2 infection (Lee et al., 2020; Long et al., 2020; To et al., 2020). Differences in the half-lives of IgM and IgG, distinct detection methods, and competition between IgM and IgG affect the final outcome of IgG and IgM determination in serum. Thus, the presence of IgM failed to identify acute infection, for which diagnosis still relies on virologic tests (Qiu et al., 2020). In addition, the sensitivity of IgG improved at later time points compared with earlier time points. This finding suggests that performing antibody detection at an early stage may reduce diagnostic accuracy. To date, the estimation of prevalence in most areas affected by the pandemic has been <5% (Pollán et al., 2020). Therefore, as shown by the unacceptably low PPV of the IgM^{+} or IgG^{+} (67.52%) panel, serological tests were not capable of identifying nearly one-third of cases. Proper algorithms and the intended detection purpose of serological testing have yet to be determined.

Specificity of serological tests is related to cross-reaction between antigens of SARS-CoV-2 and other coronaviruses. Currently, the S protein and N protein are the main targets for antibody-based detection of SARA-CoV-2. The S protein includes S1 and S2 subunits that aid in host infection, while the N protein plays an important role in the transcription and replication of viral RNA (Dutta et al., 2020; Huang et al., 2020). The N protein is highly homogeneous to SARS-CoV-1 (~90%) and shares sequence similarity with other human coronaviruses (19–45%). However, the S1 subunit of SARS-CoV-2 shares only 64% and 57% sequence homology with SARS-CoV-1 and Middle East respiratory syndrome coronavirus, respectively, and 9–37% sequence homology with other human coronaviruses (Kames et al., 2020). The S2 subunit shows 88% sequence homology with SARS-CoV-1 (Huang et al., 2020). Therefore, the S1 subunit of SARS-CoV-2 potentially demonstrates less cross-reactivity among the endemic coronaviruses. This meta-analysis provides comprehensive evidence showing that use of the S1 protein as the antigen for antibody preparation improved the diagnostic accuracy compared with the N protein; this finding supported recent studies which reported a higher false-positive rate for the N protein compared with the S1 protein (Jiang et al., 2020). In addition, given that neutralizing antibodies primarily target the S1 protein in SARS-CoV-2, the clinical value of serological testing might expand to the evaluation of vaccine efficacy and individual immunity. The data suggest that serological tests against the S1 protein, rather than the N protein, should be used in the diagnosis of COVID-19.

Caution is also warranted because of the discrepancy between different detection methods. This meta-analysis found that CLIA had higher diagnostic accuracy than either ELISA or LFIA for COVID-19. Similar results were reported by a previous meta-analysis that aimed to compare the diagnostic accuracy of different methods (Lisboa Bastos et al., 2020). LFIA have been developed and marketed rapidly to meet the urgent need for a low-cost, rapid and accurate point-of-care test, but their low sensitivity is of particular concern. The factors influencing the analytic detection of LFIA include the size of gold nanoparticles, antibody concentration, conjugation pH and characteristics of membranes (Safenkova et al., 2012). The considerable heterogeneity of included patients and sample types were also potential contributors to the low pooled performance of LFIA. It is worth noting that many sample types were used in serological tests, including serum, plasma and whole blood. However, a recent study demonstrated higher sensitivity when using serum or plasma compared with whole blood, whereas no significant difference was observed between plasma and serum (De Marinis et al., 2020). Therefore, the use of whole blood (or capillary blood) might cause relatively low sensitivity for disease surveillance (Flower et al., 2020). As the COVID-19 pandemic continues to accelerate, it is crucial to determine suitable detection methods depending on the clinical setting.

Targeted at mucosa, SARS-CoV-2 is able to induce strong mucosal immunity, leading to the generation of secretory IgA. The IgA system not only serves as the first-line barrier to prevent pathogens from adhering to the mucosa, but also interacts with the innate and adaptive immune systems for the maintenance of homeostasis (Pabst, 2012; Renegar et al., 2004). Studies have found that IgA and IgG levels were markedly higher in patients with severe disease compared with patients with mild or moderate disease (Yu et al., 2020). Studies on IgA are limited, but there is room for future exploration. In further research, serological tests using a combination of the three antibodies could be a possible way to improve diagnostic accuracy.

This study has several limitations. The majority of eligible studies were retrospective, undertaken with knowledge of whether patients had COVID-19 prior to study enrolment. Lack of blinding could lead to potential bias when interpreting the test results. In addition, the included studies were mainly from China, North America and Europe, and a lack of data from other regions might represent another source of bias. Several case–control studies used healthy individuals as the control group, which concealed possible cross-reactions and led to imprecise estimation of pooled specificity. The use of in-house testing kits, which are not widely commercially available, raises concerns about replicability and generalizability. All tests were performed by experienced laboratory researchers, which may not represent the situation in clinical practice, especially for point-of-care testing. Only two studies performed LFIA using whole blood samples, which raises applicability concerns.

In conclusion, this meta-analysis found that the IgM+ or IgG+ panel–with a diagnosis of COVID-19 if a sample tested positive for either IgM or IgG–demonstrated the highest performance compared with other panels. These findings are particularly important given the rapid development of serological tests, causing massive heterogeneity regarding the types of antibodies used in product making, and variable test results. A consensus on serological tests for the diagnosis of COVID-19 based on the high number of clinical studies is needed urgently.

Ethical approval

As the current study did not involve human subjects, the Ethics Committee at the First Affiliated Hospital of Sun Yat-Sen University exempted this study from the need for ethical approval.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ijid.2021.01.016.

References

Beeching NJ, Fletcher TE, Beadsworth MBJ. Covid-19: testing times. BMJ 2020;369.

Caini S, Bellerba F, Corso F, Díaz-Basabe A, Natoli G, Paget J, et al. Meta-analysis of diagnostic performance of serological tests for SARS-CoV-2 antibodies up to 25 April 2020 and public health implications. Euro Surveillance 2020;25. https://doi.org/10.1002/ghd92c.

Cheng MP, Yansouni CP, Basta NE, Desjardins M, Kanjilal S, Paquette K, et al. Serodiagnostics for severe acute respiratory syndrome-related coronavirus-2: a narrative review. Ann Intern Med 2020;173(6):450–60, doi:http://dx.doi.org/10.7326/M20-2854.

De Marinis Y, Sunnerhagen T, Bompada P, Bläckberg A, Yang R, Svensson J, et al. Serology assessment of antibody response to SARS-CoV-2 in patients with COVID-19 by rapid IgM/IgG antibody test. Infect Ecol Epidemiol 2020;10:1821513.

Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. Cochrane Database Syst Rev 2020;6: CD013652.

Dutta NK, Mazumdar K, Gordy JT. The nucleocapsid protein of SARS-CoV-2: a target for vaccine development. J Virol 2020;94: e00647-20.

Dwamena B. MIDAS: Stata Module for Meta-Analytical Integration of Diagnostic Test Accuracy Studies. Boston: Statistical Software Components; 2007.

Flower B, Brown JC, Simmons B, Moshe M, Frise R, Penn R, et al. Clinical and laboratory evaluation of SARS-CoV-2 lateral flow assays for use in a national COVID-19 seroprevalence survey. Thorax 2020;75:1082–8.

Food and Drug Administration. EUA Authorized Serology Test Performance. Silver Spring, MD: FDA; 2020 Available at: https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/eua-authorized-serology-test-performance [Accessed 20 September 2020].

Harbord R. METANDI: Stata Module to Perform Meta-Analysis of Diagnostic Accuracy. Boston: Statistical Software Components; 2008.

Huang Y, Yang C, Xu X, Xu W, Liu S. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharm Sin 2020;41:1141–9.

Jiang H, Li Y, Zhang H, Wang W, Men D, Yang X, et al. SARS-CoV-2 proteome microarray for global profiling of COVID-19 specific IgG and IgM responses. Nature Communications 2020;11(1)3581, doi:http://dx.doi.org/10.1038/s41467-020-17488-8 2020.03.20.20039495.

Kames J, Holcomb DD, Kimchi O, DiCuccio M, Hamasaki-Katagiri N, Wang T, et al. Sequence analysis of SARS-CoV-2 genome reveals features important for vaccine design. Sci Rep 2020;10:15643.

Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in false-negative rate of reverse transcriptase polymerase chain reaction-based SARS-CoV-2 tests by time since exposure. Ann Intern Med 2020;173(4):262–7, doi:http://dx.doi.org/10.7326/M20-1495.

Lee Y-L, Liao C-H, Liu P-Y, Cheng C-Y, Chung M-Y, Liu C-E, et al. Dynamics of anti-SARS-Cov-2 IgM and IgG antibodies among COVID-19 patients. J Infect 2020;81: e55–8.

Lin C, Xiang J, Yan M, Li H, Huang S, Shen C. Comparison of throat swabs and sputum specimens for viral nucleic acid detection in 52 cases of novel coronavirus (SARS-Cov-2)-infected pneumonia (COVID-19). Clin Chem Lab Med 2020;58, doi:http://dx.doi.org/10.1515/cclm-2020-0187.

Lisboa Bastos M, Tavaziva G, Abidi SK, Campbell JR, Haraoui L-P, Johnston JC, et al. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. BMJ 2020;370:m2516.

Long QX, Liu BZ, Deng HJ, Wu GC, Deng K, Chen YK, et al. Antibody responses to SARS-CoV-2 in patients with COVID-19. Nat Med 2020;26:845–8.

Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.

Moura DTH de, McCarty TR, Ribeiro IB, Funari MP, Oliveira PVAG de, Miranda Neto AA de, et al. Diagnostic characteristics of serological-based COVID-19 testing: a systematic review and meta-analysis. Clinics 2020;75. https://doi.org/10.1002/gg899p.

National Institutes of Health. Coronavirus Disease 2019 (COVID-19) Treatment Guidelines. Bethesda, MD: NIH; 2020 Available at: https://www.covid19treatmentguidelines.nih.gov/ [Accessed 1 September 2020].

Nuccetelli M, Pieri M, Grelli S, Ciotti M, Miano R, Andreoni M, et al. SARS-CoV-2 infection serology: a useful tool to overcome lockdown?. Cell Death Discov 2020;6. https://doi.org/10.1002/gg3bsb.

Pabst O. New concepts in the generation and functions of IgA. Nat Rev Immunol 2012;12:821–32.

Pan Y, Zhang D, Yang P, Poon LLM, Wang Q. Viral load of SARS-CoV-2 in clinical samples. Lancet Infect Dis 2020;20:411.

Peeling RW, Wedderburn CJ, Garcia PJ, Boeras D, Fongwen N, Nkengasong J, et al. Serology testing in the COVID-19 pandemic response. Lancet Infect Dis 2020;20: e245–9.

Pollán M, Pérez-Gómez B, Pastor-Barriuso R, Oteo J, Hernán MA, Pérez-Olmeda M, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet 2020;396:535–44.

Qiu X, Xiang Y, Sun J, Wang X, Chen G, Xu X, et al. Dynamic changes of throat swabs RNA and serum antibodies for SARS-CoV-2 and their diagnostic performances in patients with COVID-19. Emerg Microbes Infect 2020;0:1–33.

Renegar KB, Small PA, Boykins LG, Wright PF. Role of IgA versus IgG in the control of influenza viral infection in the murine respiratory tract. J Immunol 2004;173:1978–86.

Safenkova I, Zherdev A, Dzantiev B. Factors influencing the detection limit of the lateral-flow sandwich immunoassay: a case study with potato virus X. Anal Bioanal Chem 2012;403:1595–605.

To KK-W, Tsang OT-Y, Leung W-S, Tam AR, Wu T-C, 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;20:565–74.

Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA 2020;323:1843–4.

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. Ann Intern Med 2011;155:529–36.

World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard. Geneva: WHO; 2020 Available at: https://covid19.who.int [Accessed 20 September 2020].

Yu H-Q, Sun B-Q, Fang Z-F, Zhao J-C, Liu X-Y, Li Y-M, et al. Distinct features of SARS-CoV-2-specific IgA response in COVID-19 patients. Eur Respir J 2020;56:2001526.