

Title page:

Diagnostic Accuracy of FDA Authorized Serology Tests to Detect SARS-CoV-2 Antibodies: A Systematic Review and Meta-analysis

Rajesh Pandey, ^a M.D., Anand Gourishankar, ^b M.B.B.S., MRCP, MAS

Affiliations:

^a Division of Neonatology, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX, USA.

^b Division of Pediatric Hospital Medicine, Children's National Hospital, The George Washington School of Medicine and Health Sciences, Washington DC, USA.

Address correspondence to

Rajesh Pandey, MD

Assistant Professor,

Department of Pediatrics, Division of Neonatology

McGovern Medical School

University of Texas Health Science Center at Houston

6431 Fannin Street, MSB: 3.230

Houston, TX, 77030

Email: Rajesh.Pandey@uth.tmc.edu

Funding Source: No external funding for this manuscript.

Financial Disclosure: All authors have indicated they have no financial relationships relevant to this article to disclose.

Potential conflict of Interest: The authors have no conflicts of interest relevant to this article to disclose

Clinical Trial Registration: Not applicable

Abbreviations: Coronavirus disease (COVID-19); reverse transcription-polymerase chain reaction (RT-PCR), World Health Organization (WHO), Centers for Disease Control and Prevention (CDC), Food and Drug Administration (FDA), Immunoglobulin G (IgG), Immunoglobulin M (IgM), severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), Area under curve (AUC), Summary receiver operating characteristic (SROC)

Words: 2372

Tables: 2

Figures: 3

Supplement material: 4 files

Key– points:

Questions: What is the pooled diagnostic accuracy of FDA authorized serology tests to detect SARS-CoV-2 antibodies?

Findings: In this systematic review and meta-analysis of seven reports from FDA authorized serology tests to detect antibodies against SARS-CoV2 antibodies (3336 patients/ samples) pooled sensitivity was 87%, and pooled specificity was almost 100%. There was a wide variation in test performance based on the duration between the onset of symptoms and the tests.

Meaning: FDA authorized tests are highly accurate to detect antibodies against SARS-CoV-2 antibodies if tests are performed under a similar condition, as presented in the original report. There is a wide variation in the test performance based on the time interval between the onset of symptoms to the tests.

69

70 **Abstract:**

71 Importance: Serology tests are diagnostic and complementary to molecular tests during the
72 COVID-19 pandemic.

73 Objective: To evaluate the diagnostic accuracy of FDA authorized serology tests for the
74 detection of SARS-CoV-2 infection.

75 Data sources: A search of MEDLINE, SCOPUS, CINAHL Plus, and EMBASE up to April 4,
76 2020, was performed to identify studies using the "COVID 19 testing" and "meta-analysis." FDA
77 website was accessed for the list of tests for emergency use authorization (EUA).

78 Study Selection: Manufacturer reported serology tests published in the FDA website were
79 selected. Two reviewers independently assessed the eligibility of the selected reports.

80 Data extraction and synthesis: The meta-analysis was performed in accordance with the
81 PRISMA guidelines. A bivariate analysis using the "random-effects model" was applied for
82 pooled summary estimates of sensitivity, specificity, and the summary receiver operating
83 characteristic curves.

84 Main outcomes and measures: The primary outcome was the diagnostic accuracy of the serology
85 test for detecting SARS-CoV-2 infection. Subgroup analysis of the diagnostic accuracy with lag
86 time between symptom onset and testing were studied.

87 Results: Seven manufacturer listed reports were included. The pooled sensitivity was 87% (95%
88 CI, 78% - 93%), the pooled specificity was 100% (95% CI, 97% - 100%), and the area under the
89 hierarchical summary receiver operating characteristic curve was 0.97. At ≤ 7 days, sensitivity
90 was 44% (95% CI, 21% - 70%), and for 8-14 days, sensitivity was 84% (95% CI, 67% - 94%).

For blood draws ≥ 15 days after the onset of symptoms, sensitivity was 96% (95% CI, 93% - 98%). Heterogeneity was substantial, and the risk of bias was low in this analysis.

Conclusions and relevance: FDA authorized serology tests demonstrate high diagnostic accuracy for SARS-CoV-2 infection (certainty of evidence: moderate). There is a wide variation in the test accuracy based on the duration between the onset of symptoms and the tests (certainty of evidence: low).

Background:

Coronavirus disease (COVID-19) is a current global threat, and diagnostic tests are critical. COVID-19 was discovered in Hubei Province, China, in December 2019, caused by a pathogen of unknown origin,^{1,2} which was subsequently named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). COVID-19 has currently caused a pandemic.^{3,4} The genome sequence of SARS-CoV-2 was added to the GenBank sequence repository on January 10, 2020,^{2,5} which enabled laboratories across the world to develop and use reverse transcription-polymerase chain reaction (RT-PCR) test for the diagnosis. Both the World Health Organization (WHO) and the Centers for Disease Control and Prevention (CDC) recommend an RT-PCR for the diagnosis of SARS-CoV-2 infection.^{6,7}

SARS- CoV-2 is a new pathogen in humans, so there are no antibodies found in a person not exposed to this virus.⁷ The rise in antibody titer between acute and convalescent-phase sera tested in parallel is definite evidence of the viral infection.⁷ Seroconversion after SARS-CoV-2 infection appears as early as the fourth day after the onset of symptoms,⁸ but can be delayed beyond 15 days.⁹ Previous publications have supported the use of serological tests for the detection of SARS-CoV-2 infection.^{9,10} WHO has affirmed the use of seroconversion or an increase in antibody titer for the diagnosis of SARS- CoV-2.⁷ However, the CDC does not recommend serological tests for diagnosis of SARS-CoV-2. Nonetheless, the Food and Drug Administration (FDA) has authorized emergency use authorizations (EUA) for diagnostic serological tests for COVID-19^{6, 11} as a valuable tool for the management of this pandemic.

Serological tests offer several advantages including (1) contact tracing; (2) serologic surveillance; (3) identification of those who have already had the virus infection and thus may (if there is protective immunity) be immune;¹² (4) to identify individuals as a potential source for convalescent serum treatments (currently an experimental treatment or prophylaxis);^{13, 14} (5) to do epidemiological researches; (6) theoretically to diagnose a person in a late stage of disease when RT-PCR may be false negative;¹⁵ and (7) evaluation and development of a vaccine.^{15,16}

With the novel Coronavirus surge, testing brings challenges with variability in performance, including sensitivity and specificity.¹⁵ Therefore, the purpose of this review was to evaluate the diagnostic accuracy of the FDA authorized diagnostic serological tests for SARS-CoV-2 infection.

Methods

Literature search

We used the FDA website to search for all FDA authorized, under emergency use authorization (EUA), in-vitro serology diagnostic tests for SARS-CoV-2. We performed a search for systematic reviews up to April 4, 2020, using the term "COVID 19 testing" and "meta-analysis" in the following databases: MEDLINE, CINAHL Plus, EMBASE, and SCOPUS. This search was limited to publication in the "English" language.

Inclusion criteria

We included all FDA authorized diagnostic serological tests (before April 26, 2020) available to use in the United States of America. We collected information from the authorization reports, either the manufactures' instructions/package insert.¹¹ Inclusions also had the following criteria: index test must include qualitative detection of anti-SARS-CoV-2

antibodies [Immunoglobulin G (IgG) or Immunoglobulin M (IgM), or both], and target condition must be confirmed by RT-PCR (the reference standard).

Exclusion criteria

(1) Serology tests not authorized by the FDA; (2) serology tests not authorized for diagnostic use; (3) molecular tests authorized by the FDA.

Data extraction and quality assessment

Two authors (RP and AG) reviewed all eligible reports and extracted the data independently. Disagreements between the researchers were resolved by consensus. We extracted the following data: (1) manufacturer of the test; (2) settings including sample size, disease status at the time of specimen collection, other clinical contexts available in the report; (3) information regarding test-assay methods, specimens used; and (4) data on a true positive, false positive, false negative and true negative. There was a complete agreement on the final data used in the analysis.

We assessed the methodological quality of each study by following the guidelines from the Cochrane Screening and Diagnostic Test Methods Group, a tool adapted from the QUADAS-2. The four domains assessed for risk of bias were patient selection, index test, reference standard, and flow and timing. We assessed the applicability concerns in the first three domains. In each domain, we answered the signaling questions with 'Yes,' 'No,' or 'Unclear.' For each domain, the risk of bias was judged as 'Low,' 'High,' or 'Unclear' risk. One review author assessed study quality, which a second review-author verified. We summarized the overall quality of evidence using the GRADE methodology recommended for diagnostic tests. The assessment of publication bias was not applicable based on our research design.

Data synthesis and analysis

We constructed 2x 2 tables for all reports based on dichotomous data from the reference standard (positive or negative). We performed data analysis using methods described in the Cochrane Handbook of Diagnostic Test Accuracy (DTA) Reviews.¹⁷ We created forest plots with 95% confidence intervals (CI) for sensitivity and specificity for each study using Review Manager 5.¹⁸ Summary statistics of four sets of primary data, namely true positive (TP), false positive (FP), false negative (FN), and true negative (TN), were analyzed. Seroconversion, which determines the accuracy of tests, after SARS-CoV-2 infection depends on the time interval between the onset of disease and testing.^{9, 19, 20} To address the test accuracy and variation in sensitivity of tests with the lag time, we performed subgroup analysis for tests done at (1) at or before 7 days; (2) between 8-14 days and (3) at or after 15 days after the onset of symptoms. A bivariate random-effects model was used for summary effect size, and Area under curve (AUC) using R software version 3.2.^{21, 22} This bivariate model takes account into binomial distribution for directly modeling sensitivity and specificity for within-study variations, and assumes bivariate normal distribution for between-study variation.²³ We used inverse variance when calculating the weights of individual studies. The DerSimonian-Laird estimator was used when calculating between-study variations for statistical assumptions and conservative estimates. The confidence interval was estimated by the Clopper-Pearson method. Heterogeneity was categorized as follows: not important (0% - 40%); moderate (30% - 60%), substantial (50% - 90%); and considerable (75% - 100%).²⁴ In addition to ROC, the Spearman's correlation coefficient (Rho) between the sensitivity and false-positive rates was calculated for the presence of "cut-off" point (≥ 0.6).²⁵ GRADEpro software was used to create a summary of findings

table.²⁶ Publication bias deferred in this analysis as we used manufacturer quoted reports published by the FDA.

Results

The systematic literature search yielded no meta-analysis on serological tests for SARS-CoV-2 infection as of April 4, 2020. A total of 48 EUA summaries were identified from the FDA website accessed on April 25, 2020, of which seven reports which met eligibility criteria were included (eFigure 1 in the supplement).

Characteristics of the included reports

All seven studies are exclusively obtained from the FDA website. Information was collected from the information for use (IFU) documents. The characteristics of the included reports are summarized in Table 1.

The sample sizes in each report ranged from 70 to 1166 (total n = 3336). Five reports used a qualitative detection of IgM, IgG, or both antibodies against the SARS-CoV-2, while the remaining two used qualitative detection for the detection of IgG only. Demographics such as age, ethnicity, gender were not available. Comorbidities of the patients who tested positive for SARS-CoV-2 infection were unknown. As the cases are rising rapidly, there was no information available regarding the prevalence of the disease.

There was an incomplete and variable reporting of the time interval between the onset of symptoms and the test. Autobio[®] Anti-SARS-CoV-2 Rapid Test had information on the timing of the test relative to the onset of symptoms while Cellex[®] qSARS-CoV-2 IgG/IgM Rapid Test did not have such information. Chembio DPP[®] COVID-19 IgM/IgG System test had information on days of blood samples collected from symptom onset in COVID patients. LIAISON[®] SARS-CoV-2 S1/S2 IgG had information on the lag time between the onset of symptoms and the tests.

They performed serial tests, up to three times, for negative results. We used only the first of the serial measurements in this review. Mount Sinai Laboratory COVID-19 ELISA IgG Antibody Test had information on the timing of sample collection (7 to 14 days after onset of symptoms and retested at 21 days if the initial test was negative). VITROS[®] Anti-SARS-CoV-2G had information on all patients. But we did not include in our sub-group analysis as their reported time did not fit into our proposed analysis. VITROS[®] Anti-SARS-CoV2T had information on 15 out of 30 subjects. The characteristics of included reports are summarized in table 1. We believe the seven studies provided overall good methodological quality and low risk of bias (Figure 1). See supplemental material for details of the risk of bias assessment (eFigure 2 in the supplement).

Diagnostic performance of serology tests

The sensitivity of the individual tests ranged from 54% to 94% and specificity ranged from 90% to 100%. We created a coupled forest plot with 95 % confidence interval for individual reports with sensitivity and specificity (Figure 2). The total effect size sensitivity (pooled sensitivity) of all seven studies by the random-effects model was 87% (95% CI, 78% to 93%) (eFigure 3 in the supplement). The Higgins' I² was 86%, and the Cochrane Q statistic was 60; $p < 0.001$, which suggests substantial heterogeneity. The total effect size specificity (pooled specificity) of all seven studies by the random-effects model was 100% (95% CI, 97% to 100%) (eFigure 4 in the supplement). The Higgins' I² was 93%, and the Cochrane Q statistic was 26; $p < 0.002$, which suggests considerable heterogeneity. AUC for the SROC was significant (0.97) (Figure 3). The Spearman's correlation coefficient between the sensitivity and specificity was negative 0.6, also indicating the presence of an implicit cut-off points. Overall, the diagnostic

accuracy of the test was 96% (95% CI, 95% to 97%). The number needed to diagnose was 1.2 (95% CI, 1.1 to 1.2). Around 12 persons need to be tested to return 10 positive tests.

Subgroup analysis

Subgroup analysis detected a notable variation on test sensitivity depending on the timing of blood drawn after the onset of symptoms (lag time). At ≤ 7 days (51 samples, 4 reports), sensitivity was 44% (95% CI, 21% to 70%), for 8-14 days (152 samples, 5 reports), sensitivity was 84% (95% CI, 67% to 94%). For blood draws ≥ 15 days after the onset of symptoms (329 samples, 3 reports), sensitivity was 96% (95% CI, 93% to 98%) (Table 2).

Discussion

Pooled sensitivity of 87% and specificity of 100 % support that the FDA authorized serology tests are useful in ruling out SARS-CoV-2 infection if the tests are done under similar conditions as conducted in the reports. We noticed variations in testing time from the onset of symptoms as well as serial measurements on some reports. Although heterogeneity of specificity was higher (93%) compared to those of sensitivity (86%), SROC curve with less scattering of test values and moderate Spearman's correlation coefficient of negative 0.59 validates these studies. The diagnostic accuracy and the number needed to diagnose reflect the strength of the serology tests.

After the onset of infection with SARS-CoV-2, seroconversion occurs variably, ranging from days to weeks.²⁷ Zhao J et al.⁹ showed empirical evidence for the role of antibody titers with a lag time similar to our study findings. In that study, among 173 patients, the median duration for IgM and IgG antibodies seroconversion was 12 days and 14 days, respectively. Moreover, less than 40% had antibody detected within the first week of illness. In another study of 285 patients, median seroconversion was 13 days for IgG, with 100% positive within 19 days

of symptoms.¹⁹ IgM antibodies are detectable earlier than IgG antibody (208 plasma sample: IgM median 5 (IQR, 3-6 days) versus IgG median 14(IQR, 10-18 days).²⁰ We found an increase in test sensitivity with an increase in time between the illness and the test [sensitivity 96% for tests done ≥ 15 days from symptom onset (Table 2)]. This high sensitivity finding from only three reports warrants validation from more extensive studies. Negative tests, on appropriate clinical settings, should be interpreted with caution, mostly if the tests are done early in the courses of the disease.

Cross-reactivity of SARS-CoV-2 antibodies with other proteins in the blood is under investigation. Reports of false-positive test results are published.^{28 29} Cross-reactivity between SARS-CoV-2 antibodies with other known antibodies were tested in all reports. Chembio DPP® COVID-19 IgM/IgG System reported cross-reactivity against human coronavirus HKU1 and human coronavirus 229E. LIAISON® SARS-CoV-2 S1/S2 IgG reported cross-reactivity against the Hepatitis B Virus, Anti-Influenza A, and rheumatoid factors. These cross-reactivity limitations should be considered while interpreting positive tests.

Limitations

Information on the study population is not available to critically appraise individual reports. We could not verify the spectrum bias. Per protocol, we limited our search to FDA authorized serology tests for detecting antibodies against SARS-CoV-2. None of the serology tests have been independently verified for accuracy by the FDA. Five reports used detection of IgG or IgM or both, and the remaining two used detection of IgG only for the diagnosis of SARS-CoV-2. We accepted the threshold detection of antibodies as defined by individual reports.

Heterogeneity could not be explained by threshold as the serology tests were qualitative. Heterogeneity was substantial. We could not perform meta-regression or sensitivity analysis as limited information was available. Higher heterogeneity in this review can arise from chance, patient selection, and the type of diagnostic test.

We found a significant area under the curve (AUC), which represents the global summary performance of the test. The symmetric shoulder of SROC represents the variability in the studies and the trade-off between sensitivity and specificity.

Conclusions

WHO and CDC recommend RT-PCR for the diagnosis of SARS-CoV-2 infection. Both CDC and WHO agree that serology tests play a pivotal role in the COVID-19 pandemic considering cost, diagnostic accuracy, and relatively quick results. The overall sensitivity and specificity of the FDA authorized serological tests for detecting SARS-CoV-2 infection are high (certainty of evidence: moderate). There is a wide variation in the test accuracy based on the duration between the onset of symptoms and the tests (certainty of evidence: low).

Funding: None

Acknowledgment: Drs. Lisa M. Scheid and Eric W. Reynolds for their review and comments on the draft, and Dr. Audrey Wanger for useful discussions.

References:

1. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579(7798):270-273. doi: 10.1038/s41586-020-2012-7 [doi].
2. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in china. *Nature*. 2020;579(7798):265-269. doi: 10.1038/s41586-020-2008-3 [doi].
3. World Health Organization. Coronavirus disease (COVID-19) pandemic. <https://www.who.int>. Updated 2020. Accessed April 25, 2020.
4. World Health Organization. Coronavirus disease (COVID 2019) situation reports <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>. Updated 2020. Accessed April 25, 2020.
5. Lu R, Zhao X, Li J, et al. Genomic characterization and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565-574. doi: S0140-6736(20)30251-8 [pii].

- 337 6. Centers for Disease Control and Prevention. Testing for COVID-19
338 <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/testing.html>. Updated 2020.
339 Accessed May 2, 2020.
- 340 7. World Health Organization. Use of laboratory methods for SARS diagnosis.
341 <https://www.who.int/csr/sars/labmethods/en/#lab>. Updated 2020. Accessed April 25, 2020.
- 342 8. Xiang F, Wang X, He X, et al. Antibody detection and dynamic characteristics in patients with
343 COVID-19. Clin Infect Dis. 2020. doi: ciaa461 [pii].
- 344 9. Zhao J, Yuan Q, Wang H, et al. Antibody responses to SARS-CoV-2 in patients of novel
345 coronavirus disease 2019. Clin Infect Dis. 2020. doi: ciaa344 [pii].
- 346 10. Jin Y, Wang M, Zuo Z, et al. Diagnostic value and dynamic variance of serum antibody in
347 coronavirus disease 2019. Int J Infect Dis. 2020; 94:49-52. doi: S1201-9712(20)30198-3 [pii].
- 348 11. Food and Drug Administrations. Emergency use authorizations.
349 [https://www.fda.gov/medical-devices/emergency-situations-medical-devices/emergency-use-](https://www.fda.gov/medical-devices/emergency-situations-medical-devices/emergency-use-authorizations#covid19ivd)
350 [authorizations#covid19ivd](https://www.fda.gov/medical-devices/emergency-situations-medical-devices/emergency-use-authorizations#covid19ivd). Updated 2020. Accessed April 26, 2020.
- 351 12. World Health Organization. "Immunity passports" in the context of COVID-19.
352 [https://www.who.int/news-room/commentaries/detail/immunity-passports-in-the-context-of-](https://www.who.int/news-room/commentaries/detail/immunity-passports-in-the-context-of-covid-19)
353 [covid-19](https://www.who.int/news-room/commentaries/detail/immunity-passports-in-the-context-of-covid-19). Updated 2020. Accessed April 25, 2020.
- 354 13. Teixeira da Silva, J A. Convalescent plasma: A possible treatment of COVID-19 in India.
355 Med J Armed Forces India. 2020. doi: 10.1016/j.mjafi.2020.04.006 [doi].

- 356 14. FDA. Recommendations for investigational COVID-19 convalescent plasma.
357 [https://www.fda.gov/vaccines-blood-biologics/investigational-new-drug-ind-or-device-](https://www.fda.gov/vaccines-blood-biologics/investigational-new-drug-ind-or-device-exemption-ide-process-cber/recommendations-investigational-covid-19-convalescent-plasma)
358 [exemption-ide-process-cber/recommendations-investigational-covid-19-convalescent-plasma](https://www.fda.gov/vaccines-blood-biologics/investigational-new-drug-ind-or-device-exemption-ide-process-cber/recommendations-investigational-covid-19-convalescent-plasma).
359 Updated 2020. Accessed April 29, 2020.
- 360 15. Patel R, Babady E, Theel ES, et al. Report from the American society for microbiology
361 COVID-19 international summit, 23 March 2020: Value of diagnostic testing for SARS-CoV-
362 2/COVID-19. mBio. 2020;11(2): e00722-20. doi: 10.1128/mBio.00722. doi:
363 10.1128/mBio.00722-20 [doi].
- 364 16. Loeffelholz MJ, Tang YW. Laboratory diagnosis of emerging human coronavirus infections -
365 the state of the art. Emerg Microbes Infect. 2020;9(1):747-756. doi:
366 10.1080/22221751.2020.1745095 [doi].
- 367 17. Beynon R, Leeftang MM, McDonald S, et al. Search strategies to identify diagnostic
368 accuracy studies in MEDLINE and EMBASE. Cochrane Database Syst Rev.
369 2013;(9):MR000022. Doi (9):MR000022. doi: 10.1002/14651858.MR000022.pub3 [doi].
- 370 18. The Nordic Cochrane Centre, The Cochrane Collaboration. Review manager (RevMan)
371 [computer program]. version 5.3. copenhagen:2014.
- 372 19. Long QX, Liu BZ, Deng HJ, et al. Antibody responses to SARS-CoV-2 in patients with
373 COVID-19. Nat Med. 2020. doi: 10.1038/s41591-020-0897-1 [doi].
- 374 20. Guo L, Ren L, Yang S, et al. Profiling early humoral response to diagnose novel coronavirus
375 disease (COVID-19). Clin Infect Dis. 2020. doi: 10.1093/cid/ciaa310 [doi].

21. R Core Team (2014). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. <https://www.r-project.org>. Updated 2020. Accessed April 25, 2020.
22. Shim SR, Kim S, Lee J, Shim SR. Diagnostic test accuracy: Application and practice using R software. Epidemiology and health. 2019;41: e2019007. doi: 10.4178/epih.e2019007.
23. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982-990. doi: S0895-4356(05)00162-9 [pii].
24. Higgins J GS. Cochrane handbook for systematic reviews of interventions. version 6. the cochrane collaboration. <https://training.cochrane.org/handbook/current>. Updated 2020. Accessed May 7, 2020.
25. Deville WL, Buntinx F, Bouter LM, et al. Conducting systematic reviews of diagnostic studies: Didactic guidelines. BMC Med Res Methodol. 2002; 2:9-2288. doi: 1471-2288-2-9 [pii].
26. McMaster University, 2015 (developed by Evidence Prime. GRADEpro GDT: GRADEpro Guideline development tool [software] <https://gradepro.org>. Updated 2020. Accessed April 26, 2020.
27. Mallapaty S. Will antibody tests for the coronavirus really change everything? Nature. 2020. doi: 10.1038/d41586-020-01115-z [doi].
28. Yan G, Lee CK, Lam LTM, et al. Covert COVID-19 and false-positive dengue serology in Singapore. Lancet Infect Dis. 2020. doi: S1473-3099(20)30158-4 [pii].

29. Zhao J, Yuan Q, Wang H, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. Clin Infect Dis. 2020. doi: ctaa344 [pii].

Table 1: Characteristics of the reports included in the systematic review

Tests ^a	Sample source	Antibody detected	TYPE ^b	Target protein	TP ^c	FN ^c	FP ^c	TN ^c	Total
Autobio ^d	Serum/plasma	IgM and IgG	RDT	Spike protein	357	48	3	309	717
Cellex ^e	Plasma/serum /whole blood	IgM & IgG	RDT	Nucleocapsid	120	8	11	239	378
Chembio ^f	Plasma/serum/ whole blood	IgM & IgG	RDT	Nucleocapsid	29	2	4	35	70
LIAISON ^g	Serum/plasma	IgG	RDT	Spike protein	41	35	8	1082	1166
MSL ^h	Serum/plasma	IgG	ELISA	Spike protein	37	3	0	74	114
VITROS – CoV2G ⁱ	Serum/plasma	IgG	Modified ELISA	Spike protein	42	6	0	407	455
VITROS-COV2T ^j	Serum/plasma	IgM and IgG	Modified ELISA	Spike protein	30	6	0	400	436

- a. See text for description
- b. RDT – Rapid Diagnostic test, ELISA – Enzyme Linked Immunosorbent Assay
- c. TP – True Positive, FN – False Negative, FP - False Positive, TN – True Negative
- d. Autobio® Anti-SARS-CoV-2 Rapid Test: Included 717 patients (405 PCR positive and 312 PCR negative for SARS-CoV-2 infection), done January to Mid- March 2020. They reported information on the time of testing after onset of symptoms.
- e. Cellex® qSARS-CoV-2 IgG/IgM Rapid Test: included 378 (98 samples collected from individuals, who exhibited mild or no clinical symptoms, 30 samples were collected from hospitalized, with severe symptoms, 250 negative samples collected prior to September 2019. The day of sample collection relative to the onset of illness was unknown.
- f. Chembio DPP® COVID-19 IgM/IgG System: They had 31 (RT PCR confirmed positive) symptomatic subjects, in New York or East Asia during COVID -19 pandemic, 41 symptomatic subjects who were confirmed negative by PCR. They reported information on the time of testing after onset of symptoms. Reported cross-reactivity to Human Coronavirus HKU1 and Human Coronavirus 229E
- g. LIAISON® SARS-CoV-2 S1/S2 IgG: 76 European patients (44 patients hospitalized with moderate symptoms and 32 severe symptoms), 1090 presumed SARS-CoV-2 negative samples from a European laboratory routine (n=90) and European blood donors (n=1000) were collected prior to COVID-19. They performed a serial antibody measurement based on testing time after days of diagnosis, Cross reactive to: Hepatitis B Virus, Anti-Influenza A and rheumatoid factors
- h. Mount Sinai Laboratory COVID-19 ELISA IgG Antibody Test: Total 114 samples reported on table (58 samples (collected between 7 and 14 days after onset of symptoms. Fifty-five presumed true negative samples collected prior to the onset of the emergence of COVID-19 pandemic (that adds to 113)
- i. VITROS® Anti-SARS-CoV-2G: 48 samples collected from patients confirmed to be SARS-CoV-2 positive by PCR were tested. Of the 48 PCR positive samples 42 were reactive in the VITROS Anti-SARS-CoV-2 IgG assay and 6 were non-reactive 407 PCR negative control. They reported information on the time of testing after onset of symptoms in test positive subjects. They were not included in our subgroup analysis because the time frame did not fit in to our proposed protocol.
- j. VITROS® Anti-SARS-CoV2T: 36 samples collected from patients confirmed to be SARS-CoV-2 positive by PCR were tested, 400 presumed SARS-CoV-2 antibodies negative healthy blood donor. They reported information on the time of testing after onset of symptoms in 15/30 test positive subjects.

Table 2: Summary of Findings: Primary outcomes and subgroup analysis

Diagnostic Accuracy of FDA Authorized Serology Tests to Detect SARS-CoV-2 Antibodies

Number of studies (subjects) : 7 (3336)

True positive: 710

True Negative: 2546

New test : Anti SARS- CoV-2 antibody (IgM and/ or IgG)

Reference test : Reverse Transcription Polymerase Chain Reaction

Pooled sensitivity : 0.87 (95% CI: 0.78 to 0.93) | Pooled specificity : 0.99 (95% CI: 0.97 to 1.00)

Risk of Bias: Serious

Indirectness: Not serious

Inconsistency: Not serious

Imprecision: Not serious

Publication bias: Not applicable

Test result	Number of results per 100,000 patients tested (95% CI)			Number of participants (studies)	Certainty of the Evidence (GRADE)
	Prevalence 10% Typically seen in	Prevalence 20% Typically seen in	Prevalence 60% Typically seen in		
True positives	8720 (7750 to 9310)	17440 (15500 to 18620)	52320 (46500 to 55860)	3336 (7)	⊕⊕⊕○ MODERATE
False negatives	1280 (690 to 2250)	2560 (1380 to 4500)	7680 (4140 to 13500)		
True negatives	89550 (86940 to 89910)	79600 (77280 to 79920)	39800 (38640 to 39960)	3336 (7)	⊕⊕⊕○ MODERATE
False positives	450 (90 to 3060)	400 (80 to 2720)	200 (40 to 1360)		

CI: Confidence interval

Diagnostic Accuracy of FDA Authorized Serology Tests to Detect SARS-CoV-2 Antibodies based on the lag period between on onset of symptoms and testing.

Accuracy of test result (blood draw ≤ 7 days after onset of symptoms)

Risk of Bias: Serious

Indirectness: Not serious

Inconsistency: Serious

Imprecision: Not serious

Publication bias: Not applicable

Pooled sensitivity : 0.44 (95% CI: 0.21 to 0.70)

Test result	Number of results per 100,000 patients tested (95% CI)	Number of participants (studies)	Certainty of the Evidence (GRADE)
	Prevalence 100% Typically seen in		
True positives	44000 (21000 to 70000)	51 (4)	⊕⊕○○ LOW
False negatives	56000 (30000 to 79000)		

Accuracy of test result (blood draw between 8-14 days of onset of symptoms)

Risk of Bias: Serious

Indirectness: Not serious

Inconsistency: Serious

Imprecision: Not serious

Publication bias: Not applicable

Pooled sensitivity : 0.84 (95% CI: 0.67 to 0.94)

Test result	Number of results per 100,000 patients tested (95% CI)	Number of participants (studies)	Certainty of the Evidence (GRADE)
	Prevalence 100% Typically seen in		
True positives	84000 (67000 to 94000)	152 (5)	⊕⊕○○ LOW
False negatives	16000 (6000 to 33000)		

Accuracy of test result (blood draw ≥15 days of onset of symptoms)

Risk of Bias: Serious

Indirectness: Not serious

Inconsistency: Serious

Imprecision: Not serious

Publication bias: Not applicable

Pooled sensitivity : 0.96 (95% CI: 0.93 to 0.98)

Test result	Number of results per 100,000 patients tested (95% CI)	Number of participants (studies)	Certainty of the Evidence (GRADE)
	Prevalence 100% Typically seen in		
True positives	96000 (93000 to 98000)	329 (3)	⊕⊕○○ LOW
False negatives	4000 (2000 to 7000)		

CI: Confidence interval

534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556

557

558




559

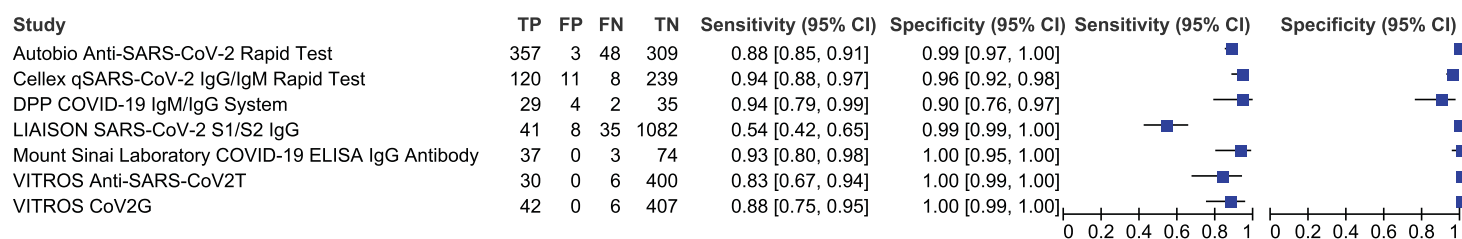
Figure 1. Forest plot: Sensitivity and specificity of serological antibody tests for diagnosing SARS-CoV-2 infection.

Figure 2. Summary receiver operating characteristic (SROC) plot of serological antibody tests for SARS-CoV-2 infection. Study estimates of sensitivity and specificity are shown with the SROC curve.

Figure 3. Risk of bias and applicability concerns summary: as judged by the review authors.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Autobio® Anti-SARS-CoV-2 Rapid Test	?	+	+	+	+	?	+
Cellex® qSARS-CoV-2 IgG/IgM Rapid Test	+	-	+	+	+	?	+
Chembio DPP® COVID-19 IgM/IgG System	+	-	+	+	+	+	+
LIAISON® SARS-CoV-2 S1/S2 IgG	?	-	+	+	+	+	+
Mount Sinai Laboratory COVID-19 ELISA IgG Antibody	+	+	+	+	+	+	+
VITROS® Anti-SARS-CoV-2G	+	-	+	+	+	+	+
VITROS® Anti-SARS-CoV2T	+	-	+	+	+	?	+

 High
 Unclear
 Low



SROC (bivariate model) for Diagnostic Test Accuracy

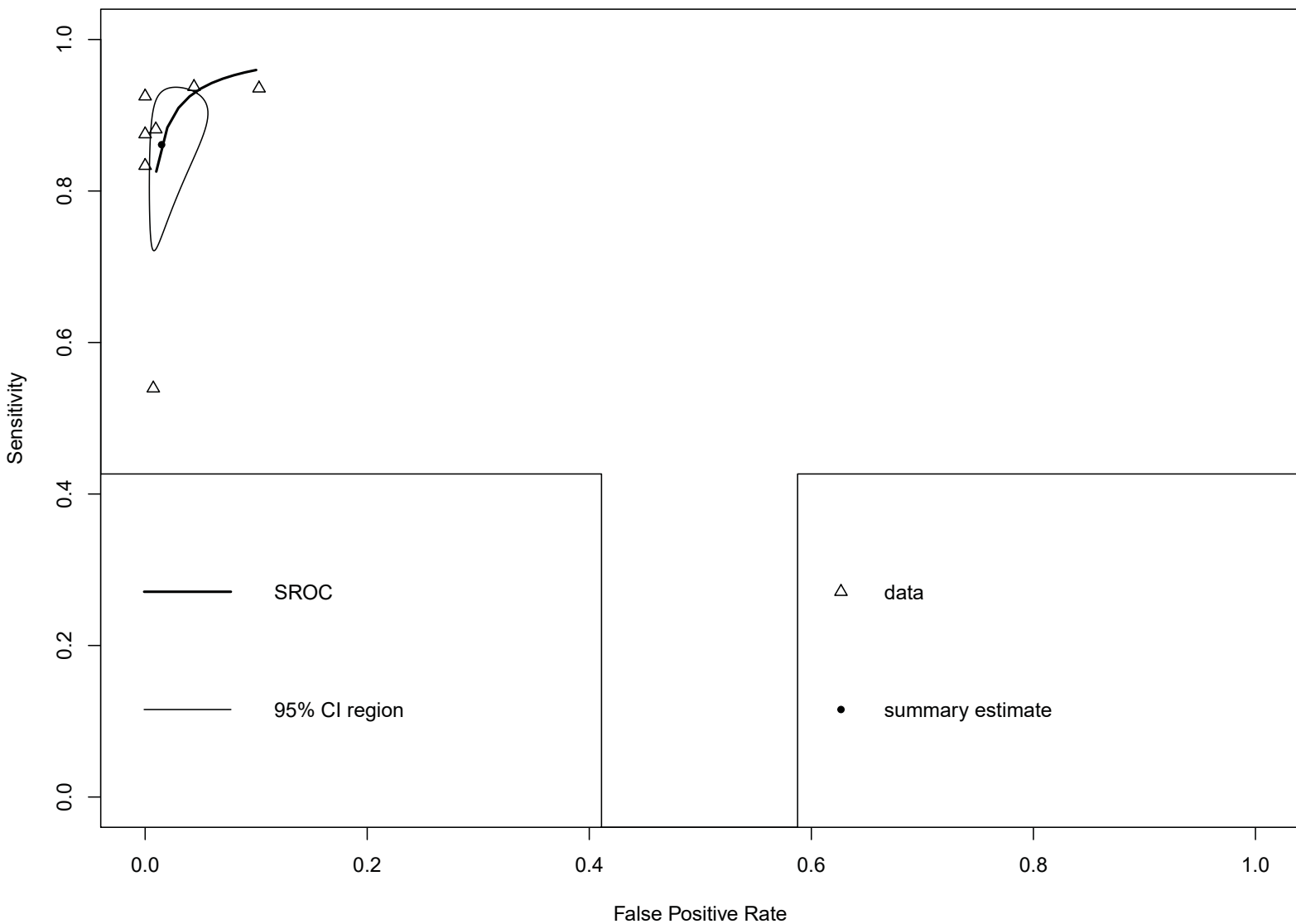


Table 1: Characteristics of the reports included in the systematic review

Tests ^a	COD ^b	Sample source	Antibody detected	TYPE ^c	Target protein	Time to result	TP ^d	FN ^d	FP ^d	TN ^d	Total	Cross reactivity
Autobio ^e	China	Serum/plasma	IgM and IgG	RDT	Spike protein	10-30 minutes	357	48	3	309	717	None
Cellex ^f	USA /China	Plasma/serum /whole blood	IgM & IgG	RDT	Nucleocapsid	10-30 minutes	120	8	11	239	378	None
Chembio ^g	USA	Plasma/serum/ whole blood	IgM & IgG	RDT	Nucleocapsid	10-30 minutes	29	2	4	35	70	2 ^l
LIAISON ^h	USA	Serum/plasma	IgG	RDT	Spike protein	10-30 minutes	41	35	8	1082	1166	3/168 positive ^m
MSL ⁱ	USA	Serum/plasma	IgG	ELISA	Spike protein	1-5 hours	37	3	0	74	114	None
VITROS – CoV2G ^j	USA	Serum/plasma	IgG	Modified ELISA	Spike protein	48 minutes	42	6	0	407	455	None
VITROS-COV2T ^k	USA	Serum/plasma	IgM and IgG	Modified ELISA	Spike protein	48 minutes	30	6	0	400	436	None

a. See text for description

b. COD – Country of development

c. RDT – Rapid Diagnostic test, ELISA – Enzyme Linked Immunosorbent Assay

d. TP – True Positive, FN – False Negative, FP - False Positive, TN – True Negative

e. Autobio® Anti-SARS-CoV-2 Rapid Test: Included 717 patients (405 PCR positive and 312 PCR negative for SARS-CoV-2 infection), done January to Mid- March 2020. They reported information on the time of testing after onset of symptoms.

f. Cellex® qSARS-CoV-2 IgG/IgM Rapid Test: included 378 (98 samples collected from individuals, who exhibited mild or no clinical symptoms, 30 samples were collected from hospitalized, with severe symptoms, 250 negative samples collected prior to September 2019. The day of sample collection relative to the onset of illness was unknown.

g. Chembio DPP® COVID-19 IgM/IgG System: They had 31 (RT PCR confirmed positive) symptomatic subjects, in New York or East Asia during COVID -19 pandemic, 41 symptomatic subjects who were confirmed negative by PCR. They reported information on the time of testing after onset of symptoms.

h. LIAISON® SARS-CoV-2 S1/S2 IgG: 76 European patients (44 patients hospitalized with moderate symptoms and 32 severe symptoms), 1090 presumed SARS-CoV-2 negative samples from a European laboratory routine (n=90) and European blood donors (n=1000) were collected prior to COVID-19. They performed a serial antibody measurement based on testing time after days of diagnosis

i. Mount Sinai Laboratory COVID-19 ELISA IgG Antibody Test: Total 114 samples reported on table (58 samples (collected between 7 and 14 days after onset of symptoms. Fifty-five presumed true negative samples collected prior to the onset of the emergence of COVID-19 pandemic (that adds to 113)

j. VITROS® Anti-SARS-CoV-2G: 48 samples collected from patients confirmed to be SARS-CoV-2 positive by PCR were tested. Of the 48 PCR positive samples 42 were reactive in the VITROS Anti-SARS-CoV-2 IgG assay and 6 were non-reactive 407 PCR negative control. They reported information on the time of testing after onset of symptoms in test positive subjects. They were not included in our subgroup analysis because the time frame did not fit in to our proposed protocol.

- k. VITROS ®Anti-SARS-CoV2T: 36 samples collected from patients confirmed to be SARS-CoV-2 positive by PCR were tested, 400 presumed SARS-CoV-2 antibodies negative healthy blood donor. They reported information on the time of testing after onset of symptoms in 15/30 test positive subjects.
- l. Cross reactive to: Human coronavirus HKU1 and Human coronavirus 229E
- m. Cross reactive to: Hepatitis B Virus, Anti-Influenza A and rheumatoid factors

Table 2: Summary of Findings
Diagnostic Accuracy of FDA Authorized Serology Tests to Detect SARS-CoV-2 Antibodies

Number of studies (subjects) : 7 (3336)

True positive: 710

True Negative: 2546

New test : Anti SARS- CoV-2 antibody (IgM and/ or IgG)

Reference test : Reverse Transcription Polymerase Chain Reaction

Pooled sensitivity : 0.87 (95% CI: 0.78 to 0.93) | Pooled specificity : 0.99 (95% CI: 0.97 to 1.00)

Risk of Bias: Serious

Indirectness: Not serious

Inconsistency: Not serious

Imprecision: Not serious

Publication bias: Not applicable

Test result	Number of results per 100,000 patients tested (95% CI)			Number of participants (studies)	Certainty of the Evidence (GRADE)
	Prevalence 10% Typically seen in	Prevalence 20% Typically seen in	Prevalence 60% Typically seen in		
True positives	8720 (7750 to 9310)	17440 (15500 to 18620)	52320 (46500 to 55860)	3336 (7)	⊕⊕⊕○ MODERATE
False negatives	1280 (690 to 2250)	2560 (1380 to 4500)	7680 (4140 to 13500)		
True negatives	89550 (86940 to 89910)	79600 (77280 to 79920)	39800 (38640 to 39960)	3336 (7)	⊕⊕⊕○ MODERATE
False positives	450 (90 to 3060)	400 (80 to 2720)	200 (40 to 1360)		

CI: Confidence interval

Diagnostic Accuracy of FDA Authorized Serology Tests to Detect SARS-CoV-2 Antibodies based on the lag period between on onset of symptoms and testing.

Accuracy of test result (blood draw ≤ 7 days after onset of symptoms)

Risk of Bias: Serious

Indirectness: Not serious

Inconsistency: Serious

Imprecision: Not serious

Publication bias: Not applicable

Pooled sensitivity : 0.44 (95% CI: 0.21 to 0.70)

Test result	Number of results per 100,000 patients tested (95% CI)	Number of participants (studies)	Certainty of the Evidence (GRADE)
	Prevalence 100% Typically seen in		
True positives	44000 (21000 to 70000)	51 (4)	⊕⊕○○ LOW
False negatives	56000 (30000 to 79000)		

Accuracy of test result (blood draw between 8-14 days of onset of symptoms)

Risk of Bias: Serious

Indirectness: Not serious

Inconsistency: Serious

Imprecision: Not serious

Publication bias: Not applicable

Pooled sensitivity : 0.84 (95% CI: 0.67 to 0.94)

Test result	Number of results per 100,000 patients tested (95% CI)	Number of participants (studies)	Certainty of the Evidence (GRADE)
	Prevalence 100% Typically seen in		
True positives	84000 (67000 to 94000)	152 (5)	⊕⊕○○ LOW
False negatives	16000 (6000 to 33000)		

Accuracy of test result (blood draw ≥15 days of onset of symptoms)

Risk of Bias: Serious

Indirectness: Not serious

Inconsistency: Serious

Imprecision: Not serious

Publication bias: Not applicable

Pooled sensitivity : 0.96 (95% CI: 0.93 to 0.98)

Test result	Number of results per 100,000 patients tested (95% CI)	Number of participants (studies)	Certainty of the Evidence (GRADE)
	Prevalence 100% Typically seen in		
True positives	96000 (93000 to 98000)	329 (3)	⊕⊕○○ LOW
False negatives	4000 (2000 to 7000)		

CI: Confidence interval