

Meta-analysis of the clinical performance of commercial SARS-CoV-2 nucleic acid, antigen and antibody tests up to 22 August 2020

Van Walle I^{1*}, Leitmeyer K¹ and Broberg E K¹, on behalf of the European COVID-19 microbiological laboratories group

Abstract

We reviewed the clinical performance of SARS-CoV-2 nucleic acid, viral antigen and antibody tests based on 94739 test results from 157 published studies and 20205 new test results from 12 EU/EEA Member States. Pooling the results and considering only results with 95% confidence interval width $\leq 5\%$, we found 4 nucleic acid tests, among which 1 point of care test, and 3 antibody tests with a clinical sensitivity $\geq 95\%$ for at least one target population (hospitalised, mild or asymptomatic, or unknown). Analogously, 9 nucleic acid tests and 25 antibody tests, among which 12 point of care tests, had a clinical specificity of $\geq 98\%$. Three antibody tests achieved both thresholds. Evidence for nucleic acid and antigen point of care tests remains scarce at present, and sensitivity varied substantially. Study heterogeneity was low for 8/14 (57.1%) sensitivity and 68/84 (81.0%) specificity results with confidence interval width $\leq 5\%$, and lower for nucleic acid tests than antibody tests. Manufacturer reported clinical performance was significantly higher than independently assessed in 11/32 (34.4%) and 4/34 (11.8%) cases for sensitivity and specificity respectively, indicating a need for improvement in this area. Continuous monitoring of clinical performance within more clearly defined target populations is needed.

¹European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden

*Correspondence: ivo_van_walle@yahoo.com

Introduction

Testing is one of the central pillars of public health actions in epidemic and pandemic situations to allow timely identification, contact tracing, and isolation of infectious cases to reduce the spread of infectious diseases. In addition, it allows e.g. estimating disease incidence, disease prevalence, and prevalence and duration of humoral immunity. Reliable testing for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and timely reporting of the data to public health authorities is therefore key for management of the coronavirus disease 2019 (COVID-19) pandemic. This requires appropriate and accurate diagnostic tests, with sufficient availability, for both identifying individuals that are currently infected with SARS-CoV-2 as well as those that have been infected in the past. Timely access to testing, sufficient supply of testing materials, availability of tests and related reagents and consumables as well as high-throughput testing are pivotal in this context.

A large number of commercial tests for SARS-CoV-2 RNA (nucleic acid tests or NATs) or viral antigen detection are currently available, as well as serological tests for SARS-CoV-2 specific antibodies. The various types of tests can be used for different purposes and many of these tests have the CE-IVD certificate that indicates compliance with the European in vitro diagnostics directive (98/79/EC) and can thus be marketed within the EU/EEA countries. In addition, the US Food and Drug Administration has granted emergency use authorisations for use of many commercial tests in the US and the World Health Organisation maintains an emergency use listing of commercial tests.[1,2] It is however important to note that the CE-marking is based on a self-declaration of the test manufacturer, including the claims on performance of the test. Independent information on the clinical performance of these tests in terms of sensitivity and specificity is still limited, and yet this is critical for proper interpretation of results.

For this reason, the European Centre for Disease Prevention and Control (ECDC) launched a continuous call to EU/EEA Member States and the UK on 1 April 2020 to provide any such clinical performance data for sharing with other Member States. These data, provided by 12 Member States, are presented in this paper. In addition, publicly available data were added. Finally, minimal performance criteria for different intended uses were gathered from public sources and aided by a survey conducted among EU/EEA member states and the UK from 20 May to 1 June 2020.

Methods

Search strategy and selection criteria

Studies containing potentially usable data on clinical performance of SARS-CoV-2 nucleic acid, antigen and antibody tests were first extracted from systematic reviews on this topic. These reviews were identified through an initial Pubmed (Medline) search for systematic reviews and meta-analysis for 'COVID-19' and 'SARS-CoV-2', followed by snowballing using the 'find similar articles' feature. The selection was then extended with the studies listed in the Foundation for Innovative Diagnostics (FIND) database and the European Commission (EC) COVID-19 In Vitro Diagnostic Devices and Test Methods Database. Both databases attempt to exhaustively identify peer-reviewed as well as grey literature on clinical performance of COVID-19 tests and are continuously updated.[3,4] Results from the latter were further filtered on those with a description indicating that they contain clinical performance results. Results produced by United States Food and Drug Administration were also included.[5] Finally, Pubmed was searched according to the query in supplementary Figure S1.

The resulting studies were subsequently assessed for eligibility. At present there are virtually no clinical performance studies that can be judged as being at low risk of bias and low applicability concerns, and systematic reviews up to this point have not used risk of bias or applicability concerns as exclusion criteria.[6-9] This was not done in this work either. Instead, studies were excluded if they did not contain data on commercial tests, or if one or more of the authors were employed by the developer or manufacturer of the index test, to avoid possible conflicts of interest. Studies with an ineligible design, such as blinded tests, analytical validation only, use of another threshold for positivity than in the instructions for use, comparisons between different specimen types or use of an antibody rather than nucleic acid test as reference test for any type of index test were subsequently excluded as well.

Further exclusions were done at sample level based on the reference test employed. Samples classified as actual negatives, i.e. used for determining specificity, had to either (i) be taken before the COVID-19 outbreak, in practice before 2020, (ii) be taken from an individual without COVID-19 compatible symptoms, or (iii) be taken from an individual with COVID-19 compatible symptoms but who was confirmed with another respiratory illness. Samples classified as actual negatives that were taken during the outbreak and were negative according to a nucleic acid test were therefore excluded. This was done to maximally reduce misclassification as actual negatives due to known issues with sensitivity of nucleic acid tests. Such misclassified samples would artificially lower index test specificity in particular when the index test is more sensitive than the reference test.[10-16] For the same reason the reported sensitivity of nucleic acid or antigen index tests, based on a nucleic acid reference test, was considered to be a positive agreement instead, calculated as part of a head-to-head comparison between the two tests. For antibody index tests on the other hand, a nucleic acid test was considered to be a valid reference test to determine actual positive samples and sensitivity, in accordance with WHO interim guidelines.[17]

Manufacturer reported clinical sensitivity and specificity data were extracted from instructions for use where available, or otherwise from the manufacturer's website. Sensitivity results derived from contrived samples spiked with purified viral RNA were excluded.

Original clinical performance data

Primary clinical performance data generated by the COVID-19 microbiological laboratories group's co-authors were assessed by ECDC according to the same criteria as those of the literature review.

Statistical analysis

Meta-analysis of the included clinical sensitivity and specificity results was performed per test and per target, i.e. the genomic region for nucleic acid tests and the antibody isotype for antibody tests. Antigen targets were not distinguished further. Antibody test sensitivity results below the threshold days after onset were excluded. Sensitivity and positive agreement results were further stratified by case population: hospitalised cases, mild or asymptomatic cases or unknown. Pooled sensitivity and specificity values were calculated using fixed effects analysis, i.e. separately summing and dividing the number of correct predictions by the total number of samples in the group. Wilson 95% confidence intervals were calculated for pooled results. Study heterogeneity was assessed through the I^2 statistic, calculated through random effects analysis using R version 4.0.2 and the metafor package.[18] I^2 values <50.0% were considered low heterogeneity, 50.0-74.9% moderate and $\geq 75\%$ high heterogeneity.

Results

Minimum performance criteria

As of 1 June 2020, minimum performance criteria for tests were publicly available from Belgium, France, the Netherlands and the United Kingdom (supplementary Table S1). All were applicable solely to antibody (Ab) tests. The intended uses included diagnosis of COVID-19, determination of exposure to SARS-CoV-2 and determination of the immune status against SARS-CoV-2. Minimum clinical sensitivity for all of the specified intended uses ranged from 85 to 98%, with a median of 95%. These thresholds applied to samples collected at least >14 days post onset of symptoms (dpo), taking into account the time to seroconversion. Minimum clinical specificity for all of the specified intended uses was 98% in six countries and 98.5% in one. For nucleic acid and antigen confirmatory tests, the draft WHO Target Product Profiles for priority diagnostics to support response to the COVID-19 pandemic state >95% - >98% sensitivity (acceptable/desired) and >99% specificity.[19] General thresholds of >95% sensitivity and >98% specificity were used for further analysis, together with a maximum 95% confidence interval (CI) width of ≤5%. For IgM only results, an upper limit of ≤28 dpo, or the highest dpo category having a lower limit ≤28 dpo, was added as well since IgM antibodies decrease fairly rapidly and such tests are not intended to be used long after exposure.[20]

Primary clinical performance data

Eight systematic reviews were identified, including one by health technology assessment bodies not listed as a peer-reviewed study, and the primary studies included in the analysis extracted.[6-9,21-24] The full list of studies in the FIND and EC databases was retrieved on 22 August 2020. Pubmed was searched on the same date. From the EC database, 268 out of 385 studies were screened out since their description did not indicate that they would contain clinical performance data on commercial tests. Analogously, 1520 out of 1738 studies were screened out from the Pubmed results. From the combined list of 364 studies, 99 had no clinical performance data on commercial tests, 34 were excluded due to a potential conflict of interest and 74 were excluded due to an ineligible design, leaving a total 157 included studies. A complete overview of the study selection is given in Figure 1. After exclusion of antibody test sensitivity results ≤14 dpo and ineligible specificity results, a total of 38202 and 56537 index test results remained for calculation of sensitivity and specificity, respectively. After addition of original, previously unpublished results provided by the authors of this study, this increased to 48310 and 66634 index test results, respectively, for 201 tests (Supplementary Tables S2-S4).

Meta-analysis

Pooled estimates for clinical sensitivity and specificity per test, target and, for sensitivity, case population were made as described in methods. For antibody tests, results were restricted to those estimates that had a 95% CI width of ≤5% and were derived from at least two studies, to be able to assess study heterogeneity. Based on the minimum performance criteria analysis, results above ≥95% sensitivity and/or ≥98% specificity for a particular population are highlighted (Table 1). Among these results, there were two CLIAs, one ELISA and no LFIAs/POCs, that had ≥95% sensitivity. Analogously, there were nine CLIAs, four ELISAs and twelve LFIAs/POCs, that had ≥98% specificity, among which the three with ≥95% sensitivity. Study heterogeneity was low for 4/10 (40.0%) sensitivity and 53/69 (76.8%) specificity results with CI width of ≤5%. There were few sensitivity results for IgG for mild or asymptomatic cases, IgA and total antibody, none of which had a CI width of ≤5% (Table 1). Compared to the same test used for hospitalised cases, drops in sensitivity were

observed of 7.4%, 11.0%, 13.1% and 19.2% for IgG, 28.8% for IgA and -6.0% for total antibody. The latter negative value is likely due to low number of samples for both populations.

For nucleic acid tests that were not point of care (POC) tests, results were restricted as for antibody tests (Table 2). Three tests had $\geq 95\%$ positive agreement with a CI width of $\leq 5\%$, and nine had $\geq 98\%$ specificity. Study heterogeneity was low for all 4 sensitivity and all 15 specificity results with CI width of $\leq 5\%$. Limited data were available for five POC antigen tests and five POC nucleic acid tests (Table 3). Large variability in positive agreement was observed. The best performing nucleic acid POC achieved a positive agreement of 98.9% (97.3-99.6%), i.e. with a CI width of $\leq 5\%$. Virtually no data were available on specificity, but no false positives were observed among them for either antigen or nucleic acid POCs.

The correlation between independently assessed clinical performance results and manufacturer reported results is shown in Figure 2. Only independently assessed results with CI width of $\leq 5\%$ are included. A total of 11/32 (34.4%) of sensitivity and 4/33 (12.1%) specificity results were significantly different ($p < 0.05$).

Discussion

This review represents, to our knowledge, the most complete independent overview so far of clinical performance of commercially available COVID-19 tests. A substantial amount of previously unpublished data from European countries are included as well. To date, there are numerous commercial tests for which sufficient performance data are available to allow calculation of clinical sensitivity or positive agreement, and specificity with narrow confidence interval ranges. Reassuringly, the clinical performances of several nucleic acid and antibody tests exceeded the minimum performance criteria. As time progresses, the list of tests with sufficient available performance data is expected to increase.

At the same time, the available evidence for point of care nucleic acid and antigen tests remains scarce, even though these tests can have substantial practical advantages for e.g. screening. We therefore recommend more emphasis on the validation of these tests, including as part of a testing algorithm, whereby the sensitivity and specificity of taking two tests with a number of days in between is assessed, and which can e.g. be useful to reduce the duration of a quarantine period.

The comparison between the independently assessed clinical performance data and manufacturer reported clinical performance revealed that in particular sensitivity is frequently (40.7% of the cases in this study) significantly overestimated by the manufacturer. At a minimum, this emphasises that such independent assessments are clearly necessary. In the longer term, an explicit and proactive regulatory mechanism in Europe to compare available independently generated evidence on these tests with manufacturer reported values, coupled with appropriate regulatory action, could be useful. This could also be rewarding towards those manufacturers that do provide robust estimates of their product's performance.

Limitations of this paper include that most of the included studies have a substantial risk of bias in the sample selection, especially for the sensitivity panel, as established also in the assessments performed in the systematic reviews that were used as a source. Results were mainly based on hospitalised cases or poorly defined populations, whereas the population of interest consists often of symptomatic cases in general or even asymptomatic cases, and differences in performance may exist depending on disease severity. While this review addresses a pressing need for actionable

clinical performance data, ideally, the clinical performance should be assessed through prospective studies or clinical trials with a guaranteed unbiased sample selection for a clearly defined target population and intended use of the test. Given the difficulty of assessing and extracting the data from individual studies in a coherent way, we recommend that the Standard for Reporting of Diagnostic Accuracy Studies (STARD) should also be followed when publishing the results.[25]

In this context, the selection of the reference test is particularly important with respect to reference negative samples. As described in some of the assessed studies, it should be avoided that index test results are considered as false positives while the samples are from actual cases and for this reason we excluded nucleic acid negative samples from suspected COVID-19 patients altogether. We expect therefore little bias in the specificity results, except potentially from under or overrepresentation of confounders. This is especially relevant for seroprevalence studies, where, in the current low-prevalence situation, in particular the specificity of the test needs to be well-defined and high. On the other hand, sensitivity results using a nucleic acid test as reference should be interpreted with caution, since the positive samples may exclude some actual cases.

Possibilities to improve the reference test can include testing - potentially only the false positives - with a second reference nucleic acid test preferably targeting different genes, testing more than one sample from the same patient including for antibodies at a later time point, testing samples from both upper and lower respiratory tracts, and sequencing the sample. The handling of intermediate index test results is an issue that needs to be described in studies, and in general these should be considered as positive results rather than either negatives or being excluded from the validation, since they would normally require further follow-up to confirm the positivity of the sample.

For the above reasons, the authors and organisations contributing to this study in no way recommend the use of the listed commercial tests over other not listed commercial or in-house tests. Finally, it should be kept in mind that this study is a snapshot in time and that the clinical performance of tests may change over time as the virus population evolves. We therefore recommend a continuous monitoring of clinical performance both in Europe and globally, which is key for reliable monitoring of the pandemic and which will also support vaccine and antiviral development. These results should be shared in a timely manner publically

Acknowledgements

We would like to acknowledge the technicians in the European microbiology laboratories who work hard to support the control of COVID-19 and supported the validations described in this manuscript.

We would like to acknowledge the companies that made some of the kits available for evaluation to some of the laboratories.

European COVID-19 microbiological laboratories group

Marjan Van Esbroeck (Institute of Tropical Medicine, Antwerpen, Belgium), Pieter Vermeersch (Clinical Department of Laboratory Medicine and National Reference Center for Respiratory Pathogens, University Hospitals Leuven, Leuven, Belgium), Kurt Beuselinck (Clinical Department of Laboratory Medicine and National Reference Center for Respiratory Pathogens, University Hospitals Leuven, Leuven, Belgium), Christos Karagiannis (Nicosia General Hospital, Cyprus), Merit Melin (Department of Health Protection, Expert Microbiology Unit, Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Nina Ekström (Department of Health Protection, Expert Microbiology Unit, Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Iris Erlund (Department of Government Services, Forensic Toxicology Unit, Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Terhi Vihervaara (Department of Government Services, Forensic

Toxicology Unit, Finnish Institute for Health and Welfare (THL, Helsinki, Finland), Vanessa Escuret (Laboratoire de Virologie des HCL, Institut des Agents Infectieux, CNR des virus à transmission respiratoire (dont la grippe), Groupement Hospitalier Nord, Lyon, France), Emilie Frobert (Laboratoire de Virologie des HCL, Institut des Agents Infectieux, CNR des virus à transmission respiratoire (dont la grippe), Groupement Hospitalier Nord, Lyon, France), Alexandre Gaymard (Laboratoire de Virologie des HCL, Institut des Agents Infectieux, CNR des virus à transmission respiratoire (dont la grippe), Groupement Hospitalier Nord, Lyon, France), Andreas Mentis (Hellenic Pasteur Institute, Athens, Greece), Stavroula Lampropoulou (Hellenic Pasteur Institute, Athens, Greece), Ivan-Christian Kurolt (Research unit, University Hospital for Infectious Diseases "Dr. Fran Mihaljević", Zagreb, Croatia), Tamir Abdelrahman (Department of Microbiology, Laboratoire national de santé, Luxembourg), Trung Nguyen (Department of Microbiology, Laboratoire national de santé, Luxembourg), Guillaume Fournier (Department of Microbiology, Laboratoire national de santé, Luxembourg), Chantal B.E.M. Reusken (Centre for Infectious Disease Control, National Institute for Public Health and the Environment, The Netherlands), Maaikje J.C. van den Beld (Centre for Infectious Disease Control, National Institute for Public Health and the Environment, The Netherlands), Janette Rahamat-Langendoen (Department of Medical Microbiology, Radboud University Medical Center, Nijmegen), Marjolijn C.A. Wegdam-Blans (Department of Medical Microbiology, PAMM, Veldhoven, The Netherlands), Jeroen H. T. Tjhie (Department of Medical Microbiology, PAMM, Veldhoven, The Netherlands), Peter Croughs (Department of Medical Microbiology and Infectious Diseases, Erasmus Medical Center, Rotterdam, The Netherlands), Corine H. GeurtsvanKessel (Department of Virology, Erasmus Medical Center, Rotterdam, The Netherlands), Johan Reimerink (Centre for Infectious Diseases Research, Diagnostics and Laboratory Surveillance, Centre for Infectious Disease Control, National Institute for Public Health and the Environment, The Netherlands), David S.Y. Ong ([a] Department of Medical Microbiology and Infection Control, Franciscus Gasthuis & Vlietland, Rotterdam, The Netherlands, [b] Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands), Hans G.M. Koeleman (Department of Medical Microbiology and Infection Control, Franciscus Gasthuis & Vlietland, Rotterdam, The Netherlands), Hannke Berkhout (Canisius-Wilhelmina hospital, Nijmegen, The Netherlands), Christel F.M. van der Donk (Canisius-Wilhelmina hospital, Nijmegen, The Netherlands), Menno D. de Jong (Department of Medical Microbiology & Infection prevention, Amsterdam University Medical Centers, The Netherlands), Rens Zonneveld (Department of Medical Microbiology, Amsterdam University Medical Center, Amsterdam, The Netherlands), Suzanne Jurriaans (Department of Medical Microbiology, Amsterdam University Medical Center, Amsterdam, The Netherlands), Nathalie Van Burgel (Hagaziekenhuis, The Hague, The Netherlands), Bas B. Wintermans (Department of Medical Microbiology and Immunology, Admiraal de Ruyter Hospital, Vlissingen, The Netherlands), Ger T. Rijkers ([a] Department of Medical Microbiology and Immunology, Admiraal de Ruyter Hospital, Goes, The Netherlands, [b] Elisabeth-Tweesteden Hospital, Tilburg, The Netherlands), Jean-Luc Murk (Elisabeth-Tweesteden Hospital, Tilburg, The Netherlands), Khoa T.D. Thai ([a] Unit of Medical Microbiology, Star-shl Medical Diagnostic Center, Rotterdam, The Netherlands, [b] Department of Medical Microbiology and Infectious Diseases, Erasmus Medical Center, Rotterdam, The Netherlands), Melanie J de Graaf ([a] Department of Medical Microbiology, University Medical Centre, Utrecht, The Netherlands, [b] Saltro Diagnostic Centre, Utrecht, The Netherlands), Annemarie van 't Veen ([a] Department of Medical Microbiology, University Medical Centre, Utrecht, The Netherlands, [b] Saltro Diagnostic Centre, Utrecht, The Netherlands), Cornelis P. Timmerman (Central Bacteriology and Serology Laboratory, Tergooi Hospital, Hilversum, The Netherlands), Annette van Corteveen-Splinter (Central Bacteriology and Serology Laboratory, Tergooi Hospital, Hilversum, The Netherlands), Felix Geeraedts (Laboratory for Medical Microbiology and Public Health, Hengelo, The Netherlands), Adrian Klak (Laboratory for Medical Microbiology and Public Health, Hengelo, The Netherlands), Maria M. Konstantinovski (Reinier Haga Medical Diagnostic Centre, Delft, The Netherlands), Manou R. Batstra (Reinier Haga Medical Diagnostic Centre, Delft, The Netherlands), K. A. Heemstra (Alrijne Zorggroep, Leiderdorp, The Netherlands), Jos J. Kerremans (Alrijne Zorggroep, Leiderdorp, The Netherlands), Inge H. M. van Loo ([a] Department of Medical Microbiology, Maastricht University Medical Center, The Netherlands, [b] Care and Public Health Research Institute, Maastricht University), Paul H. M. Savelkoul ([a] Department of Medical Microbiology, Maastricht University Medical Center, The Netherlands, [b] Care and Public Health Research Institute, Maastricht University), Johan Kissing (Department of Medical Microbiology and Infection prevention, Gelre Hospitals, Apeldoorn, The Netherlands), Paul Martijn den Reijer (Department of Medical Microbiology and Infection prevention, Gelre Hospitals, Apeldoorn, The Netherlands), Anne Russcher (Department of Medical Microbiology, Medical Meander Center, Amersfoort, The Netherlands), Moniek Heusinkveld (Department of Medical Microbiology, Hospital Gelderse Vallei, Ede, The Netherlands), Ellen van Lochem (Department of Medical Microbiology and Immunology, Hospital Rijnstate, The Netherlands), Steven F. T. Thijsen (Medical Microbiology and Immunology, Diaconessen Hospital, Utrecht, The Netherlands), Michiel Heron (Medical Microbiology and Immunology, Diaconessen Hospital, Utrecht, The Netherlands), Susanne P. Stoof (Department of Medical Microbiology, Comicro, Hoorn, The Netherlands), Sim van Gysegem (Department of Medical Microbiology, Comicro, Hoorn, The Netherlands), Sylvia B. Debast (Laboratory of Clinical Microbiology and Infectious Diseases, Isala Hospital, Zwolle, The Netherlands), Claudy Oliveira dos Santos (Laboratory of Clinical Microbiology and Infectious Diseases, Isala Hospital, Zwolle, The Netherlands), Bjorn L. Herpers (Regional Public Health Laboratory Kennemerland, The Netherlands), Theo Mank (Regional Public Health Laboratory Kennemerland, The Netherlands), Kin Ki Jim ([a] Department of Medical Microbiology and Infection Control, Jeroen Bosch Hospital, 's-

Hertogenbosch, The Netherlands, [b] Department of Medical Microbiology and Infection Prevention, Amsterdam University Medical Centers, Amsterdam institute for Infection and Immunity, Amsterdam, The Netherlands), Peter C. Wever (Department of Medical Microbiology and Infection Control, Jeroen Bosch Hospital, 's-Hertogenbosch, The Netherlands), Jutte J.C. de Vries (Department of Medical Microbiology, Leiden University Medical Center, Leiden, The Netherlands), Martine Hoogewerf (Department of Medical Microbiology, Northwest Hospital Group, Alkmaar, The Netherlands), Deborah J. Kaarsenhout (Atalmedial Medical Microbiology Laboratory, Amsterdam, the Netherlands), Annette M. Stemerding (Deventer Ziekenhuis, Deventer, the Netherlands), Babette C. van Hees (Deventer Ziekenhuis, Deventer, the Netherlands), Vishal Hira (Department of Medical Microbiology and Infection Prevention, Groene Hart Ziekenhuis, Gouda, the Netherlands), Anne E. Bos (Department of Medical Microbiology and Infection Prevention, Groene Hart Ziekenhuis, Gouda, the Netherlands), Leontine Mulder (Clinical Laboratory, Medlon B.V., Enschede, The Netherlands), Michiel van Rijn (Medical Laboratory, Ikazia Hospital, Rotterdam, The Netherlands), Aleksander Michalski (Epidemiological Response Centre of The Polish Armed Forces, Warsaw, Poland), Marta Pakieła (Voivodeship Sanitary Epidemiological Station, Warsaw, Poland), Anna Siewierska-Puchlerska (Voivodeship Sanitary Epidemiological Station, Warsaw, Poland), Jarosław Paciorek (Voivodeship Sanitary Epidemiological Station, Warsaw, Poland), Ewa Gajda (Epidemiological Response Centre of The Polish Armed Forces, Warsaw, Poland), Katarzyna Pancer (Department of Virology, BSL3 Laboratory, COVID-19 NIPH-NIH team, National Institute of Public Health-National Institute of Hygiene, Warsaw, Poland), Agnieszka Kotakowska-Kulesza (Department of Virology, COVID-19 NIPH-NIH team, National Institute of Public Health-National Institute of Hygiene, Warsaw, Poland), Magdalena Rzczkowska (Department of Bacteriology and Biocontamination Control, COVID-19 NIPH-NIH team, National Institute of Public Health-National Institute of Hygiene, Warsaw, Poland), Raquel Guiomar (Instituto Nacional de Saúde Dr. Ricardo Jorge, I.P., Portugal.), Líbia Zé-Zé (Instituto Nacional de Saúde Dr. Ricardo Jorge, I.P., Portugal.), Inês Costa (Instituto Nacional de Saúde Dr. Ricardo Jorge, I.P., Portugal.), Johan Brynedal Öckinger (Department of Virology, Clinical Microbiology, Karolinska University Laboratory, Karolinska University Hospital, Stockholm, Sweden), Berit Hammam (Department of Virology, Clinical Microbiology, Karolinska University Laboratory, Karolinska University Hospital, Stockholm, Sweden), Katarina Prosenc (National Laboratory for Health, Environment and Food Slovenia, Laboratory for Public Health Virology), Nataša Berginc (National Laboratory for Health, Environment and Food Slovenia, Laboratory for Public Health Virology)

Authors' contributions

Ivo Van Walle: conceptualisation, methodology, data curation, formal analysis, writing-review, editing. Katrin Leitmeyer: conceptualisation, methodology, data curation, writing-review, editing. Eeva K. Broberg: conceptualisation, methodology, data curation, writing-review, editing. European COVID-19 microbiological laboratories group: conceptualization, investigation, data curation, methodology, writing-review.

References

1. United States Food and Drug Administration. Emergency Use Authorizations for Medical Devices. <https://www.fda.gov/medical-devices/emergency-situations-medical-devices/emergency-use-authorizations-medical-devices#COVID19ivd>, last consulted 20 July 2020.
2. World Health Organisation. Coronavirus disease (COVID-19) Pandemic – Emergency Use Listing Procedure (EUL) open for in vitro diagnostics. https://www.who.int/diagnostics_laboratory/EUL/en/ last consulted 20 July 2020.
3. Joint Research Centre of the European Commission. COVID-19 In Vitro Diagnostic Devices and Test Methods Database. <https://COVID-19-diagnostics.jrc.ec.europa.eu/>
4. Foundation for Innovative Diagnostics (FIND). FIND evaluation update: SARS-CoV-2 molecular diagnostics. <https://finddx.shinyapps.io/COVID19DxDData/>
5. United States Food and Drug Administration openFDA dataset on COVID-19 Serological Testing Evaluations. <https://open.fda.gov/apis/device/covid19serology/download/>
6. Deeks JJ, Dinnes J, Takwoingi Y, et al., Cochrane COVID-19 Diagnostic Test Accuracy Group. Antibody tests for identification of current and past infection with SARS-CoV-2. Cochrane Database Syst Rev2020;6:CD013652. doi:10.1002/14651858.CD013652. pmid:32584464
7. EUnetHTA RCRC01 Authoring Team. The current role of antibody tests for novel coronavirus SARS-CoV-2 in the management of the pandemic. Collaborative Assessment. Diemen (The Netherlands): EUnetHTA; 2020 23rd of June. N°144 pages. Report No.:RCR01. Available from: <https://www.eunetha.eu>

8. Lisboa Bastos M, Tavaziva G, Abidi SK, et al. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *BMJ*. 2020;370:m2516. Published 2020 Jul 1. doi:10.1136/bmj.m2516
9. Dinnes J, Deeks JJ, Adriano A, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev*. 2020;8:CD013705. Published 2020 Aug 26. doi:10.1002/14651858.CD013705
10. Attwood LO, Francis MJ, Hamblin J, Korman TM, Druce J, Graham M. Clinical evaluation of AusDiagnostics SARS-CoV-2 multiplex tandem PCR assay. *J Clin Virol*. 2020;128:104448. doi:10.1016/j.jcv.2020.104448
11. Pan Y, Li X, Yang G, et al. Serological immunochromatographic approach in diagnosis with SARS-CoV-2 infected COVID-19 patients. *J Infect*. 2020;81(1):e28-e32. doi:10.1016/j.jinf.2020.03.051
12. Wang P. Combination of serological total antibody and RT-PCR test for detection of SARS-COV-2 infections. *J Virol Methods*. 2020;283:113919. doi:10.1016/j.jviromet.2020.113919
13. Zhao J, Yuan Q, Wang H, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019 [published online ahead of print, 2020 Mar 28]. *Clin Infect Dis*. 2020;ciaa344. doi:10.1093/cid/ciaa344
14. Zou L, Ruan F, Huang M, et al. SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *N Engl J Med*. 2020;382(12):1177-1179. doi:10.1056/NEJMc2001737
15. Green DA, Zucker J, Westblade LF, et al. Clinical Performance of SARS-CoV-2 Molecular Tests. *J Clin Microbiol*. 2020;58(8):e00995-20. Published 2020 Jul 23. doi:10.1128/JCM.00995-20
16. Kim H, Hong H, Yoon SH. Diagnostic Performance of CT and Reverse Transcriptase Polymerase Chain Reaction for Coronavirus Disease 2019: A Meta-Analysis. *Radiology*. 2020;296(3):E145-E155. doi:10.1148/radiol.2020201343
17. World Health Organization. (2020). Laboratory testing for coronavirus disease (COVID-19) in suspected human cases: interim guidance, 19 March 2020. World Health Organization. <https://apps.who.int/iris/handle/10665/331501>. License: CC BY-NC-SA 3.0 IGO
18. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* 2010, 36(3), 1-48.
19. COVID-19 Target product profiles for priority diagnostics to support response to the COVID-19 pandemic v.0.1 of 31 July 2020. <https://www.who.int/publications/m/item/covid-19-target-product-profiles-for-priority-diagnostics-to-support-response-to-the-covid-19-pandemic-v.0.1>
20. Sethuraman N, Jeremiah SS, Ryo A. Interpreting Diagnostic Tests for SARS-CoV-2 [published online ahead of print, 2020 May 6]. *JAMA*. 2020;10.1001/jama.2020.8259. doi:10.1001/jama.2020.8259
21. La Marca A, Capuzzo M, Paglia T, Roli L, Trenti T, Nelson SM. Testing for SARS-CoV-2 (COVID-19): a systematic review and clinical guide to molecular and serological in-vitro diagnostic assays [published online ahead of print, 2020 Jun 14]. *Reprod Biomed Online*. 2020;S1472-6483(20)30318-7. doi:10.1016/j.rbmo.2020.06.001
22. Böger B, Fachi MM, Vilhena RO, Cobre AF, Tonin FS, Pontarolo R. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19 [published online ahead of print, 2020 Jul 10]. *Am J Infect Control*. 2020;S0196-6553(20)30693-3. doi:10.1016/j.ajic.2020.07.011
23. Caini S, Bellerba F, Corso F, 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 Surveill*. 2020;25(23):2000980. doi:10.2807/1560-7917.ES.2020.25.23.2000980
24. Moura DTH, McCarty TR, Ribeiro IB, et al. Diagnostic Characteristics of Serological-Based COVID-19 Testing: A Systematic Review and Meta-Analysis. *Clinics (Sao Paulo)*. 2020;75:e2212. doi:10.6061/clinics/2020/e2212
25. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*. 2016;6(11):e012799. Published 2016 Nov 14. doi:10.1136/bmjopen-2016-012799

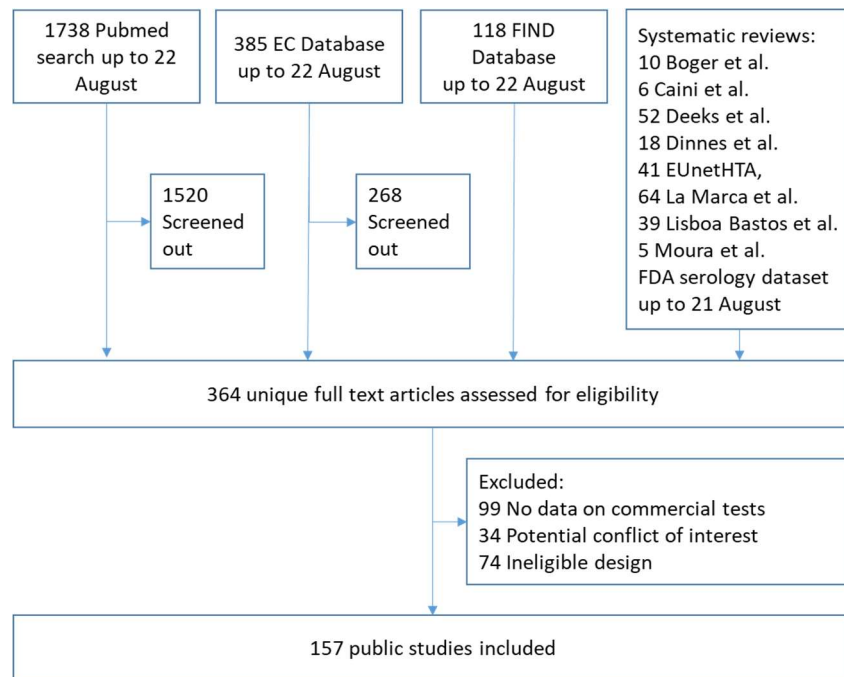


Figure 1. Selection of public studies.

Table 1: Pooled sensitivity and specificity results for antibody tests with confidence interval width $\leq 5\%$ for either or both and based on ≥ 2 studies. Only samples taken >14 days post onset of symptoms are included, and ≤ 28 days post onset for IgM only as target. Rows are sorted alphabetically by test, target and case population.

Category	Test	Target	Case population	Sensitivity ¹	Specificity ¹
Ab (CLIA)	Abbott, SARS-CoV-2 IgG assay on Architect	IgG	hospitalised	95.9 (93.4-97.5) n=368 BE, CA, NL, UK, US(3)	99.5 (99.3-99.6) n=8243 AT, BE(2), CA, DE(2), DK, FI, FR(3), IT, NL, SE, SG, UK(3), US(8)
Ab (CLIA)	Abbott, SARS-CoV-2 IgG assay on Architect	IgG	mild/asymptomatic	88.5 (84.6-91.5)x n=331 NL, UK(2), US	same as above
Ab (CLIA)	Abbott, SARS-CoV-2 IgG assay on Architect	IgG	unk	92.0 (90.4-93.3) n=1332 AT, BE, DE, DK, FI, FR(2), SE, SG, UK(2), US(4)	same as above
Ab (LFIA, POC)	Anhui Deep Blue Medical Technology, COVID-19 (SARS-CoV-2) IgG/IgM Antibody Test Kit	IgG	na	nd	99.4 (96.5-99.9) n=158 CA, US
Ab (ELISA)	Beijing Wantai Biological Pharmacy Enterprise, Wantai SARS-CoV-2 IgM ELISA	IgM	hospitalised	92.8 (88.3-95.7)x* n=195 CN(2), NL	98.7 (98.0-99.1) n=1505 CN(2), DK, NL(2)
Ab (ELISA)	Beijing Wantai Biological Pharmacy Enterprise, Wantai SARS-CoV-2 total Ab ELISA	total Ab	hospitalised	97.5 (95.9-98.5)* n=603 CN(2), DE, DK, NL	99.5 (99.2-99.7) n=3097 CN(2), DE, DK(2), FR(2), NL(3)
Ab (ELISA)	Beijing Wantai Biological Pharmacy Enterprise, Wantai SARS-CoV-2 total Ab ELISA	total Ab	unk	97.5 (94.9-98.8) n=279 AT, DK, FR	same as above
Ab (ELISA)	Bio-Rad, Platelia SARS-CoV-2 Total Ab	total Ab	na	nd	96.4 (93.3-98.1) n=250 BE, FR, LU, NL
Ab (LFIA, POC)	CTK Biotech, OnSite COVID-19 IgG/IgM Rapid Test	IgG	na	nd	98.6 (95.2-99.6) n=148 AU, NL
Ab (CLIA)	DiaSorin, Liaison XL S1/S2 IgG chemiluminescence immunoassay	IgG	hospitalised	92.9 (89.6-95.2)x* n=324 CA, DE, NL	97.7 (97.3-98.0)* n=5994 AT, BE(2), CA, DE(3), DK, FI, FR, NL(2), SE, UK, US(2)
Ab (CLIA)	DiaSorin, Liaison XL S1/S2 IgG chemiluminescence immunoassay	IgG	mild/asymptomatic	81.9 (76.3-86.3)x n=226 NL, UK	same as above
Ab (CLIA)	DiaSorin, Liaison XL S1/S2 IgG chemiluminescence immunoassay	IgG	unk	90.9 (88.9-92.6)** n=967 AT(2), BE(2), DK, SE, UK, US	same as above
Ab (CLIA)	Diazyme Laboratories, DZ-Lite SARS-CoV-2 IgM and IgG CLIA	IgG	unk	95.3 (84.5-98.7)x n=43 US(2)	99.0 (97.5-99.6) n=414 US(2)
Ab (CLIA)	Diazyme Laboratories, DZ-Lite SARS-CoV-2 IgM and IgG CLIA	IgG or IgM	unk	100.0 (91.8-100.0)x n=43 US(2)	98.6 (96.9-99.3) n=414 US(2)
Ab (CLIA)	Diazyme Laboratories, DZ-Lite SARS-CoV-2 IgM and IgG CLIA	IgM	unk	90.7 (78.4-96.3)x n=43 US(2)	99.5 (98.3-99.9) n=414 US(2)
Ab (LFIA, POC)	Dynamiker Biotechnology Tianjin, 2019 nCoV IgG/IgM Rapid test	IgG or IgM	hospitalised	100.0 (89.0-100.0)x n=31 BE, DK	97.6 (94.8-98.9) n=248 BE, DK, SE
Ab (LFIA, POC)	Dynamiker Biotechnology Tianjin, 2019 nCoV IgG/IgM Rapid test	IgG or IgM	unk	89.0 (79.8-94.3)x** n=73 SE, TW	same as above

Category	Test	Target	Case population	Sensitivity ¹	Specificity ¹
Ab (ELISA)	Epitope Diagnostics, EPI-KT-1032 Coronavirus COVID-19 IgG ELISA Kit	IgG	hospitalised	94.0 (86.7-97.4)x* n=83 CA, NL, US	97.6 (96.7-98.3)* n=1451 AT, CA, DE(2), NL, UK, US(3)
Ab (ELISA)	Epitope Diagnostics, EPI-KT-1032 Coronavirus COVID-19 IgG ELISA Kit	IgG	mild/asymptomatic	74.8 (65.8-82.0)x** n=107 NL, US	same as above
Ab (ELISA)	Epitope Diagnostics, EPI-KT-1032 Coronavirus COVID-19 IgG ELISA Kit	IgG	unk	96.0 (90.1-98.4)x* n=99 AT, DE, US	same as above
Ab (ELISA)	Epitope Diagnostics, EPI-KT-1033 Coronavirus COVID-19 IgM ELISA Kit	IgM	hospitalised	95.5 (78.2-99.2)x* n=22 CA, NL	98.1 (97.0-98.9) n=810 AT, CA, NL, US
Ab (ELISA)	Epitope Diagnostics, EPI-KT-1033 Coronavirus COVID-19 IgM ELISA Kit	IgM	unk	83.3 (70.4-91.3)x* n=48 AT, US	same as above
Ab (ELISA)	Euroimmun Medizinische Labordiagnostika, Anti-SARS-CoV-2 IgA S1 ELISA	IgA	hospitalised	96.0 (92.5-97.9)x n=224 BE(2), CA, DK, FI, FR, GR, NL	86.7 (84.9-88.3)** n=1459 AU, BE(2), CA, DK, ES, FI(2), FR(2), GR, LU, NL(2), US
Ab (ELISA)	Euroimmun Medizinische Labordiagnostika, Anti-SARS-CoV-2 IgA S1 ELISA	IgA	mild/asymptomatic	67.2 (55.0-77.4)x n=64 FI, NL	same as above
Ab (ELISA)	Euroimmun Medizinische Labordiagnostika, Anti-SARS-CoV-2 IgA S1 ELISA	IgA	unk	94.8 (90.9-97.1)x n=212 AU, BE, FR, US	same as above
Ab (ELISA)	Euroimmun Medizinische Labordiagnostika, Anti-SARS-CoV-2 IgG S1 ELISA	IgG	hospitalised	92.6 (89.7-94.7) n=431 BE(3), CA, CH(2), DE, DK, FI, FR, GR, NL, US	97.9 (97.4-98.3) n=3954 AU, BE(3), CA, CH(2), DE(6), DK(2), ES, FI(2), FR(3), GR, LU, NL(2), US(5)
Ab (ELISA)	Euroimmun Medizinische Labordiagnostika, Anti-SARS-CoV-2 IgG S1 ELISA	IgG	mild/asymptomatic	79.5 (71.9-85.5)x** n=132 CH, FI, NL, US	same as above
Ab (ELISA)	Euroimmun Medizinische Labordiagnostika, Anti-SARS-CoV-2 IgG S1 ELISA	IgG	unk	89.0 (86.7-91.0)* n=785 AT, AU, BE, DE(2), DK, FR, UK, US(2)	same as above
Ab (LFIA, POC)	Getein Biotech, One Step Test for Novel Coronavirus (2019-nCoV) IgM/IgG Antibody (Colloidal Gold)	IgG	na	nd	100.0 (96.9-100.0) n=120 CA, US
Ab (LFIA, POC)	Getein Biotech, One Step Test for Novel Coronavirus (2019-nCoV) IgM/IgG Antibody (Colloidal Gold)	IgG or IgM	na	nd	99.2 (95.4-99.9) n=120 CA, US
Ab (LFIA, POC)	Getein Biotech, One Step Test for Novel Coronavirus (2019-nCoV) IgM/IgG Antibody (Colloidal Gold)	IgM	na	nd	99.2 (95.4-99.9) n=120 CA, US
Ab (LFIA, POC)	Guangzhou Wondfo Biotech, Wondfo SARS-CoV-2 Antibody Test	IgG or IgM	unk	88.0 (82.6-92.0)x** n=184 AU, ES, TW, US	99.3 (98.3-99.7) n=605 AU, BR, ES, US(2)
Ab (LFIA, POC)	Hangzhou Alltest Biotech, 2019-nCoV IgG/IgM Rapid Test Cassette	IgG	unk	88.7 (81.6-93.3)x n=115 AU, ES	100.0 (98.5-100.0) n=254 AU, ES(2)
Ab (LFIA, POC)	Hangzhou Alltest Biotech, 2019-nCoV IgG/IgM Rapid Test Cassette	IgG or IgM	unk	92.3 (87.2-95.4)x n=168 AU, ES, TW	96.7 (93.8-98.2) n=269 AU, DK, ES(2)
Ab (LFIA, POC)	Hangzhou Alltest Biotech, 2019-nCoV IgG/IgM Rapid Test Cassette	IgM	unk	21.7 (15.2-30.1)x** n=115 AU, ES	97.2 (94.4-98.7) n=254 AU, ES(2)

Category	Test	Target	Case population	Sensitivity ¹	Specificity ¹
Ab (LFIA, POC)	Innovita Biological Technology, 2019-nCoV Ab Test (Colloidal Gold)	IgG	hospitalised	86.9 (76.2-93.2)x n=61 CA, JP	100.0 (98.5-100.0) n=258 CA, JP, US
Ab (LFIA, POC)	Innovita Biological Technology, 2019-nCoV Ab Test (Colloidal Gold)	IgM	hospitalised	75.4 (63.3-84.5)x** n=61 CA, JP	98.4 (96.1-99.4) n=258 CA, JP, US
Ab (ELISA)	Mikrogen Diagnostik, recomWell SARS-CoV-2 IgG	IgG	na	nd	96.4 (94.2-97.8) n=445 BE, DE, NL
Ab (ELISA)	NovaTec Immundiagnostica, NovaLisa SARS-CoV-2 IgA ELISA	IgA	hospitalised	88.7 (78.5-94.4)x n=62 BE(2)	95.2 (92.1-97.1)* n=293 BE(2), IT, NL
Ab (ELISA)	NovaTec Immundiagnostica, NovaLisa SARS-CoV-2 IgG ELISA	IgG	hospitalised	91.9 (82.5-96.5)x n=62 BE(2)	97.3 (94.7-98.6) n=293 BE(2), IT, NL
Ab (ELISA)	NovaTec Immundiagnostica, NovaLisa SARS-CoV-2 IgM ELISA	IgM	hospitalised	43.5 (31.9-55.9)x** n=62 BE(2)	99.0 (97.0-99.7) n=293 BE(2), IT, NL
Ab (CLIA)	Ortho Clinical Diagnostics, VITROS Immunodiagnostic Products Anti-SARS-CoV-2 IgG	IgG	unk	93.4 (89.4-96.0)x n=227 DK, UK	99.7 (99.3-99.9) n=1420 DK, UK, US
Ab (CLIA)	Ortho Clinical Diagnostics, VITROS Immunodiagnostic Products Anti-SARS-CoV-2 Total Ab	total Ab	na	nd	100.0 (99.5-100.0) n=732 DK, US
Ab (CLIA)	Roche, Elecsys Anti-SARS-CoV-2	total Ab	hospitalised	85.7 (75.7-92.1)x n=70 CA, DE, NL	99.8 (99.7-99.9) n=7833 AT, BE(3), CA, DE(5), DK, LU, NL, SE, SG, UK(2), US(5)
Ab (CLIA)	Roche, Elecsys Anti-SARS-CoV-2	total Ab	mild/asymptomatic	91.7 (84.4-95.7)x* n=96 NL, UK	same as above
Ab (CLIA)	Roche, Elecsys Anti-SARS-CoV-2	total Ab	unk	94.7 (93.3-95.7)* n=1351 AT(2), BE(3), DE(2), DK, SE, SG, UK(2), US(2)	same as above
Ab (LFIA, POC)	SD BioSensor, Standard Q COVID-19 IgM/IgG Duo	IgG	na	nd	99.8 (99.3-99.9)* n=1254 US(2)
Ab (LFIA, POC)	SD BioSensor, Standard Q COVID-19 IgM/IgG Duo	IgM	na	nd	98.8 (98.0-99.3) n=1256 US(2)
Ab (CLIA)	Shenzhen New Industries Biomedical Engineering (SNIBE), Maglumi 2019-nCoV (SARS-CoV-2) IgG/IgM kit	IgG	hospitalised	93.4 (85.5-97.2)x* n=76 BE(2)	97.6 (96.8-98.3)** n=1744 BE(2), CN(2), DK
Ab (CLIA)	Shenzhen New Industries Biomedical Engineering (SNIBE), Maglumi 2019-nCoV (SARS-CoV-2) IgG/IgM kit	IgG	unk	91.1 (89.2-92.6)** n=1084 CN, DK	same as above
Ab (CLIA)	Shenzhen New Industries Biomedical Engineering (SNIBE), Maglumi 2019-nCoV (SARS-CoV-2) IgG/IgM kit	IgG or IgM	hospitalised	96.1 (89.0-98.6)x n=76 BE(2)	98.6 (96.4-99.5) n=285 BE(3)
Ab (CLIA)	Shenzhen New Industries Biomedical Engineering (SNIBE), Maglumi 2019-nCoV (SARS-CoV-2) IgG/IgM kit	IgM	hospitalised	93.4 (85.5-97.2)x* n=76 BE(2)	99.2 (98.7-99.5)** n=1756 BE(2), CN(2), DK
Ab (CLIA)	Shenzhen New Industries Biomedical Engineering (SNIBE), Maglumi 2019-nCoV (SARS-CoV-2) IgG/IgM kit	IgM	unk	67.8 (65.0-70.5)x** n=1084 CN, DK	same as above
Ab (CLIA)	Shenzhen Yahuilong (YHLO) Biotech, SARS-CoV-2 IgG/IgM antibody detection kit	IgG	na	nd	99.0 (98.3-99.4) n=1313 CN(2), DK, IT

Category	Test	Target	Case population	Sensitivity ¹	Specificity ¹
Ab (CLIA)	Shenzhen Yahuilong (YHLO) Biotech, SARS-CoV-2 IgG/IgM antibody detection kit	IgM	na	nd	98.7 (97.9-99.2)** n=1314 CN(2), DK, IT
Ab (CLIA)	Siemens, Healthineers SARS-CoV-2 Total Assay on Atellica/ADVIA Centaur	total Ab	unk	96.7 (95.2-97.8)** n=757 DE, DK, UK	99.8 (99.5-99.9) n=2108 DE(2), DK, UK
Ab (LFIA, POC)	SureScreen Diagnostic, Covid-19 IgG/IgM Rapid Test Cassette	IgG	na	nd	99.0 (96.4-99.7) n=198 BE, NL
Ab (LFIA, POC)	VivaChek Biotech, VivaDiag COVID-19 IgM/IgG Rapid Test	IgG	unk	78.9 (69.7-85.9)x n=95 AU, US	98.2 (96.1-99.2) n=334 AU, BE, IT, NL, US
Ab (LFIA, POC)	VivaChek Biotech, VivaDiag COVID-19 IgM/IgG Rapid Test	IgG or IgM	hospitalised	100.0 (89.0-100.0)x n=31 BE, NL	97.5 (95.2-98.7) n=324 AU, BE, IT, US
Ab (LFIA, POC)	VivaChek Biotech, VivaDiag COVID-19 IgM/IgG Rapid Test	IgG or IgM	unk	80.0 (70.9-86.8)x n=95 AU, US	same as above
Ab (LFIA, POC)	VivaChek Biotech, VivaDiag COVID-19 IgM/IgG Rapid Test	IgM	unk	80.0 (70.9-86.8)x n=95 AU, US	97.8 (95.6-98.9) n=324 AU, BE, IT, US
Ab (LFIA, POC)	Xiamen Biotime Biotechnology, SARS-CoV-2 IgG/IgM Rapid Qualitative Test Kit	IgG	na	nd	98.0 (94.3-99.3) n=150 FI, US
Ab (CLIA)	Xiamen Innodx Biotech, Antibody test kit for 2019-nCoV	IgG or IgM	na	nd	99.3 (98.0-99.8) n=430 CN(2)
Ab (LFIA, POC)	Zhejiang Orient Gene Biotech, COVID-19 IgG/IgM Rapid Test Cassette	IgG	hospitalised	96.7 (91.7-98.7)x n=120 BE, CH, NL	97.7 (96.1-98.7) n=568 BE, CH, FR, NL, SE
Ab (LFIA, POC)	Zhejiang Orient Gene Biotech, COVID-19 IgG/IgM Rapid Test Cassette	IgG	unk	92.4 (85.1-96.3)x n=92 FR, SE	same as above
Ab (LFIA, POC)	Zhejiang Orient Gene Biotech, COVID-19 IgG/IgM Rapid Test Cassette	IgM	hospitalised	86.0 (77.5-91.6)x n=93 BE, NL	98.4 (96.3-99.3) n=308 BE, FR, SE
Ab (LFIA, POC)	Zhejiang Orient Gene Biotech, COVID-19 IgG/IgM Rapid Test Cassette	IgM	unk	82.6 (73.6-89.0)x* n=92 FR, SE	same as above
Ab (LFIA, POC)	Zhuhai Livzon Pharmaceutical Group, Diagnostic Kit for IgM / IgG Antibody to Coronavirus (SARS-CoV-2) (Lateral Flow)	IgG	hospitalised	86.4 (80.3-90.9)x n=162 CN(2), FR	98.0 (94.3-99.3) n=150 CN, FR, US
Ab (LFIA, POC)	Zhuhai Livzon Pharmaceutical Group, Diagnostic Kit for IgM / IgG Antibody to Coronavirus (SARS-CoV-2) (Lateral Flow)	IgM	hospitalised	75.9 (68.8-81.9)x n=162 CN(2), FR	99.3 (96.3-99.9) n=150 CN, FR, US

Abbreviations used: CLIA, chemiluminescence assay; ELISA, enzyme-linked immunosorbent assay, LFIA: lateral flow immunoassay; POC: point of care test; unk, unknown or unclearly defined; nd, not determined, either due to no data or due to data from only one country or study

¹Sensitivity and specificity values given as value (confidence interval), number of samples, list of countries (number of studies per country if >1). Value in bold if both confidence interval width ≤5% and value ≥95% (for sensitivity) or ≥98% (for specificity). *Confidence interval width >5%. *Moderate study heterogeneity (50.0≤I²<75.0%). **High study heterogeneity (I²≥75.0%).

Table 2: Pooled positive agreement and specificity results for nucleic acid tests with confidence interval width $\leq 5\%$ for either or both and based on ≥ 2 studies. Rows are sorted alphabetically by test, target and case population.

Category	Test	Target	Case population	Positive agreement ¹	Specificity ¹
PCR	Altona Diagnostics, RealStar SARS-CoV-2 RT-PCR Kit 1.0	E	unk	88.1 (80.4-93.1)x n=101 CH, FR, NL, US	100.0 (96.7-100.0) n=112 CH, NL
PCR	Altona Diagnostics, RealStar SARS-CoV-2 RT-PCR Kit 1.0	S	unk	87.1 (79.2-92.3)x n=101 CH, FR, NL, US	100.0 (96.7-100.0) n=112 CH, NL
PCR	Altona Diagnostics, RealStar SARS-CoV-2 RT-PCR Kit 1.0	S or E	unk	81.6 (75.8-86.3)x* n=207 FR(3), NL	100.0 (98.4-100.0) n=237 FR, NL, UK
PCR	AusDiagnostics, Coronavirus Typing Assay	ORF1ab	na	nd	100.0 (98.5-100.0) n=254 AU, UK
PCR	BGI, Real-time fluorescent RT-PCR kit for detecting 2019 nCoV	ORF1ab	unk	93.8 (88.7-96.7)x n=146 CH, JP, NL, PL	99.1 (95.1-99.8) n=112 CH, NL
PCR	CerTest Biotec, VIASURE SARS-CoV-2 Real Time PCR Detection Kit	N	unk	96.8 (89.1-99.1)x* n=63 CH, NL	100.0 (96.7-100.0) n=112 CH, NL
PCR	CerTest Biotec, VIASURE SARS-CoV-2 Real Time PCR Detection Kit	ORF1ab	unk	93.7 (84.8-97.5)x** n=63 CH, NL	100.0 (96.7-100.0) n=112 CH, NL
PCR	CerTest Biotec, VIASURE SARS-CoV-2 Real Time PCR Detection Kit	ORF1ab or N	na	nd	100.0 (98.2-100.0) n=207 NL, UK
PCR	DiaSorin, Simplexa COVID-19 Direct RT-PCR Kit	ORF1ab or S	unk	97.8 (94.4-99.1) n=180 US(3)	nd
PCR	Hologic, SARS-CoV-2 Assay (Panther Fusion System)	ORF1ab	unk	98.3 (96.8-99.1) n=525 FR, US(6)	nd
PCR	KH Medical, RADI COVID-19 Detection Kit and RADI COVID-19 Triple Detection Kit	RdRP	unk	96.8 (89.1-99.1)x* n=63 CH, NL	100.0 (96.7-100.0) n=112 CH, NL
PCR	KH Medical, RADI COVID-19 Detection Kit and RADI COVID-19 Triple Detection Kit	S	unk	98.4 (91.5-99.7)x n=63 CH, NL	100.0 (96.7-100.0) n=112 CH, NL
PCR	Primerdesign, genesig Real-Time PCR CoVID-19 kit	RdRP	unk	95.3 (89.4-98.0)x* n=106 CH, NL, PL	100.0 (98.8-100.0) n=307 CH, NL, UK
PCR	R-Biopharm, Ridagene SARS-CoV2	E	unk	100.0 (94.3-100.0)x n=63 CH, NL	100.0 (96.7-100.0) n=112 CH, NL
PCR	Roche, COBAS SARS-CoV-2 test	ORF1ab or E	unk	98.8 (97.9-99.3) n=1125 AT, CH, DE, FR, SI, US(5)	100.0 (90.8-100.0)x n=38 CH, FR
PCR	Seegene, Allplex 2019-nCoV assay	E	unk	85.0 (75.6-91.2)x** n=80 CH, FR, NL	100.0 (96.7-100.0) n=112 CH, NL
PCR	Seegene, Allplex 2019-nCoV assay	RdRP	unk	91.3 (83.0-95.7)x* n=80 CH, FR, NL	100.0 (96.7-100.0) n=112 CH, NL
PCR	Tibmolbiol, SARS-CoV (COVID19) E-gene	E	unk	100.0 (94.4-100.0)x n=65 CH, UK	100.0 (98.5-100.0) n=250 CH, UK

Abbreviations used: PCR, polymerase chain reaction; unk, unknown or unclearly defined; nd, not determined, either due to no data or due to data from only one country or study

¹Positive agreement and specificity values given as value (confidence interval), number of samples, list of countries (number of studies per country if >1). Value in bold if both confidence interval width $\leq 5\%$ and value $\geq 95\%$ (for positive agreement) or $\geq 98\%$ (for specificity).
*Confidence interval width >5%. *Moderate study heterogeneity ($50.0 \leq I^2 < 75.0\%$). **High study heterogeneity ($I^2 \geq 75.0\%$).

Table 3: Pooled positive agreement and specificity results for point of care antigen and nucleic acid tests. Rows are sorted alphabetically by category, test, target and case population.

Category (running time)	Test	Target	Case population	Positive agreement ¹	Specificity ¹
Ag (30')	Coris BioConcept, C-1023/TB Covid-19 Respi-Strip	Ag	unk	37.9 (33.1-43.0)x* n=369 BE(2), FR, PT	nd
Ag (15-20')	Liming Bio, COVID-19 Antigen Rapid Test Device	Ag	unk	0.0 (0.0-29.9)x*** n=9 CL	nd
Ag (5-8')	RapiGEN, BioCredit COVID-19 Ag	Ag	unk	41.8 (35.8-48.2)x** n=239 CL, HK	100.0 (77.2-100.0)x*** n=13 HK
Ag (14-16')	Savant Biotechnology, Huaketai New Coronavirus (SARS-CoV-2) N Protein Detection Kit (Fluorescence immunochromatography)	Ag	unk	83.3 (73.5-90.0)x*** n=78 CL	nd
Ag (10')	Shenzhen Bioeasy Biotechnology, Bioeasy 2019-nCoV Ag Fluorescence Rapid Test Kit	Ag	unk	86.2 (77.8-91.7)x** n=94 CL, PT	nd
LAMP (13')	Abbott, ID NOW COVID-19	RdRP	unk	79.7 (75.9-83.1)x** n=483 US(7)	nd
PCR (45')	Cepheid, GeneXpert Xpert Xpress SARS-CoV-2	E or N	unk	98.8 (97.3-99.5) n=427 BE, CH, CY, DE, FI, FR, NL, SE, US(5)	100.0 (82.4-100.0)x n=18 BE, CH, SE
PCR (30')	Mesa Biotech, Accula RSV Molecular POCT	N	unk	68.0 (54.2-79.2)x*** n=50 US	nd
PCR (unk)	MiCo BioMed, Coronavirus disease 2019(COVID-19) Detection Kit	ORF3a or N	hospitalised	66.7 (49.6-80.2)x*** n=33 EC	nd
PCR (75')	Mobidiag, Novodiag COVID-19	ORF1ab or N	unk	93.4 (84.3-97.4)x*** n=61 FI	nd

Abbreviations used: Ag: antigen; LAMP: loop-mediated isothermal amplification; PCR, polymerase chain reaction; unk, unknown or unclearly defined; nd, not determined, either due to no data or due to data from only one country or study

¹Positive agreement and specificity values given as value (confidence interval), number of samples, list of countries (number of studies per country if >1). Value in bold if both confidence interval width $\leq 5\%$ and value $\geq 95\%$ (for positive agreement) or $\geq 98\%$ (for specificity).
*Confidence interval width >5%. *Moderate study heterogeneity ($50.0 \leq I^2 < 75.0\%$). **High study heterogeneity ($I^2 \geq 75.0\%$). ***Study heterogeneity not determined (single study).

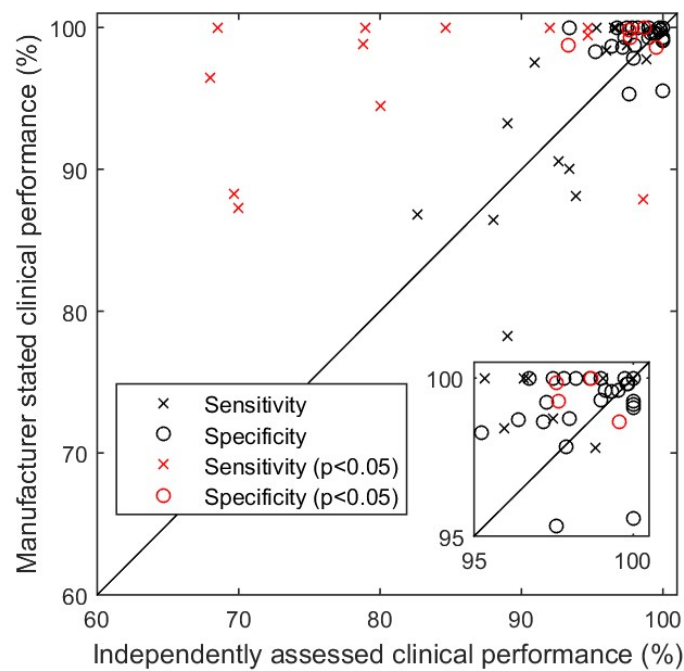


Figure 2. Independently assessed versus manufacturer reported clinical sensitivity and specificity per test, with significantly different ($p < 0.05$) results highlighted. Independently assessed results limited to those with 95% confidence interval width of $\leq 5\%$. The inset expands the 95-100% region.

SUPPLEMENTARY FIGURE S1: Pubmed search string

SUPPLEMENTARY TABLE S1: Minimum performance criteria proposed by different institutes. The data are summarised and translated when needed (only the original text is valid).

SUPPLEMENTARY TABLE S2: Studies included in the meta-analysis.

SUPPLEMENTARY TABLE S3: Clinical performance results.

SUPPLEMENTARY TABLE S4: Forest plots.