

# BMJ Open Meta-analysis of the robustness of COVID-19 diagnostic kit performance during the early pandemic

Chandrakumar Shanmugam <sup>1</sup>, Michael Behring,<sup>2</sup> Vishwas Luthra <sup>3</sup>, Sixto M Leal,<sup>4</sup> Sooryanarayana Varambally,<sup>4</sup> George J Netto,<sup>4</sup> Upender Manne <sup>5</sup>

**To cite:** Shanmugam C, Behring M, Luthra V, *et al*. Meta-analysis of the robustness of COVID-19 diagnostic kit performance during the early pandemic. *BMJ Open* 2022;**12**:e053912. doi:10.1136/bmjopen-2021-053912

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-053912>).

CS and MB contributed equally.

CS and MB are joint first authors.

Received 01 June 2021  
Accepted 29 March 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Upender Manne;  
[upendermanne@uabmc.edu](mailto:upendermanne@uabmc.edu)

## ABSTRACT

**Background** Accurate detection of SARS-CoV-2 is necessary to mitigate the COVID-19 pandemic. However, the test reagents and assay platforms are varied and may not be sufficiently robust to diagnose COVID-19.

**Methods** We reviewed 85 studies (21 530 patients), published from five regions of the world, to highlight issues involved in the diagnosis of COVID-19 in the early phase of the pandemic. All relevant articles, published up to 31 May 2020, in PubMed, BioRxiv, MedRxiv and Google Scholar, were included. We evaluated the qualitative (9749 patients) and quantitative (10 355 patients) performance of RT-PCR and serologic diagnostic tests for real-world samples, and assessed the concordance (5538 patients) between test performance in meta-analyses. Synthesis of results was done using random effects modelling and bias was evaluated according to QUADAS-2 guidelines.

**Results** The RT-PCR tests exhibited heterogeneity in the primers and reagents used. Of 1957 positive RT-PCR COVID-19 participants, 1585 had positive serum antibody (IgM±IgG) tests (sensitivity 0.81, 95% CI 0.66 to 0.90). While 3509 of 3581 participants RT-PCR negative for COVID-19 were found negative by serology testing (specificity 0.98, 95% CI 0.94 to 0.99). The chemiluminescent immunoassay exhibited the highest sensitivity, followed by ELISA and lateral flow immunoassays. Serology tests had higher sensitivity and specificity for laboratory approval than for real-world reporting data.

**Discussion** The robustness of the assays/platforms is influenced by variability in sampling and reagents. Serological testing complements and may minimise false negative RT-PCR results. Lack of standardised assay protocols in the early phase of pandemic might have contributed to the spread of COVID-19.

## INTRODUCTION

In December 2019, there was a cluster of unexplained pneumonia cases in Wuhan, China, and a novel coronavirus was identified as the causative agent.<sup>1</sup> The virus was named as SARS-CoV-2, and the disease as COVID-19.<sup>2</sup> The clinical spectrum ranges from asymptomatic forms to acute respiratory failure and multiorgan dysfunction syndrome, coagulopathy and death.<sup>3 4</sup> On 11 March 2020, the WHO described the spread of these infections

## Strengths and limitations of this study

- This study offers the first evaluation of COVID-19 test performance with consideration of the heterogeneity of RT-PCR primers.
- We compare the performance of manufacturer-based, laboratory/approval data to the performance of the same test kits in a real-world setting in the early phase of the pandemic.
- We perform a qualitative analysis of RT-PCR assays using 85 studies (21 530 patients), and a quantitative meta-analysis of RT-PCR versus serum antibody assays in a subset of 30 publications (10 355 patients).
- Much of the information in the early pandemic was reported from China, and often from non-peer reviewed, preprint sources.
- Data measuring duration of the infection was not available in majority of included studies.

as a pandemic, which persists as a global crisis. Robust diagnostic tests are required to mitigate the spread of this virus and thereby to minimise the impact of COVID-19 on the health, economy and social well-being of mankind.

The standard diagnosis of COVID-19 is based on clinical and radiologic evidence and viral genome detection by RT-PCR in respiratory samples.<sup>5</sup> Gene-specific primers are used in the RT-PCR assays; structural genes include *envelope (E)*, *nucleocapsid (N)* and/or *spike (S)*-genes; non-structural genes include *RNA-dependent RNA polymerase (RdRp)* or *open reading frame1ab (ORF1ab)*.<sup>6 7</sup> Some studies used only a single-gene-specific primer, and others used multiple-gene primers. Since studies published in the early phase of the pandemic reported a 3%–41% range of false negativity by RT-PCR, a repeat RT-PCR testing was suggested.<sup>8 9</sup> Furthermore, false negativity was attributed to either mutations in the regions to which the primers bind or to sampling and laboratory practices, including collection, transportation and handling.<sup>10</sup>

Timing of sample collection with respect to the course of infection and the sample type also influence test results.<sup>11</sup> Alternatively, the diagnosis can be made by detection of antigens (E, N or S) and/or antibodies (IgM or IgG or both) in blood samples.<sup>12</sup> However, these tests have the potential for false positives owing to cross-reactivity with other human coronaviruses.<sup>13 14</sup> Due to the unprecedented public health emergency, the U.S. Food and Drug Administration (FDA) authorised, on 1 June 2020, EUA requests for more than 15 diagnostic and serologic tests. Though serology testing can detect the false positives of RT-PCR tests in clinically suspected patients, its value in COVID-19 diagnosis as a complementary assay in the mitigation of the pandemic is not well defined. However, given the complexities in COVID-19 testing, there is a need for a review of performance for tests commonly used.

In this systematic review and meta-analysis, we examine testing for the diagnosis of COVID-19 in the early pandemic and evaluate the sensitivity and specificity of serological tests relative to RT-PCR tests. Our objectives were to assess the uniformity of primer usage in RT-PCR assays and evaluate whether primers used in gold-standard RT-PCR tests affect the validity of serological tests. Furthermore, we compared the performance of serological tests/platforms in approval contrived/laboratory versus real-world data.

## METHODS

### Literature search

This research was accomplished according to standards outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.<sup>15</sup> To find relevant studies, international databases, including PubMed, Embase, MedRxiv, BioRxiv and Google Scholar, were searched for articles published until 31 May 2020. The following search terms were used (selected using English MeSH keywords and Emtree terms): [SARS-CoV-2 AND diagnosis] OR [2019-nCoV AND diagnosis]" OR ["COVID-19 AND diagnosis] and [SARS-CoV-2 AND RT-PCR] OR, [2019-nCoV AND RT-PCR]" OR ["COVID-19 AND RT-PCR] and [SARS-CoV-2 AND serology] OR [2019-nCoV AND serology]" OR ["COVID-19 AND serology]. Additional searches were performed for references listed in the included studies.

### Eligibility criteria

Relevant articles that reported diagnostic information for infected patients were included in the analysis. Preprint articles with non-peer review were considered for inclusion. Articles were excluded if appropriate information was not reported or if they were in the Chinese language. Population sample sizes of <5 participants were not included; reviews and editorials were not considered. For meta-analysis and approval versus real-world performance, studies that reported percent sensitivity/specificity without including patient numbers were also excluded.

## Data extraction and report quality evaluation

Two authors (CS and VL) screened and evaluated the literature independently. Discrepancies were resolved by consensus after evaluation by a third author (MB). The following were extracted for review and meta-analysis: journal name, authors, period of publication (end of May 2020), location of study, total number of patients, tissue of origin for samples tested, whether samples were from upper or lower respiratory tract (or both), primers for RT-PCR, platforms for serology tests and antibodies tested for serology. Counts of true positives, false negatives, true negatives and false positives were used in the meta-analysis.

An author (MB) extracted and analysed the approved testing kit performance data from the following sources: FDA EUA Authorized Serology Test Performance,<sup>16</sup> the Foundation for Innovative New Diagnostics SARS-CoV-2 diagnostic pipeline<sup>17</sup> and package inserts provided on company websites for each product. Real-world sample testing data from kits in meta-analyses were compared against the performance of the same kits, or platforms, reported in approval documentation. Variables abstracted were study authors/test developer, name of test, test platform, and true positives, false negatives, true negatives, and false positives for each antibody or antibody combination measured (IgM, IgG, IgA, combined and Pan-Ig). Risk of bias within individual studies of meta-analysis was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for assessment of diagnostic studies.<sup>18</sup> QUADAS-2 has been developed specifically for evaluating bias in the meta-analyses of diagnostic test accuracy.

### Patient and public involvement

Since we performed a meta-analysis and systematic review, it was not appropriate or possible to involve patients or the public in the design, conduct, reporting or dissemination plans of our research.

### Statistical analysis

Statistical analyses were performed with R V.6.3.2 (12 December 2019).<sup>19</sup> The package 'meta' was used for meta-analyses.<sup>20</sup> Random effects models were used to measure sensitivity and specificity of outcomes across studies. Subgroup analysis was performed to evaluate the effect of assay, RT-PCR primer type and setting (laboratory vs real world) on serum test performance. Heterogeneity across studies and subgroups was evaluated using Cochrane's Q statistic, and residual heterogeneity was quantified as a percentage with the I<sup>2</sup> statistic. An I<sup>2</sup> measure of 0% shows no observed heterogeneity, with increasing values from 0%–100% indicating higher levels of heterogeneity.<sup>21</sup> An assumption of homogeneity was rejected for p values < 0.1. The evaluation of publication bias was not possible using FDA and EU reported approval data.

## RESULTS

### Search results and population characteristics

Our search generated 112 publications with potential relevance to the performance of COVID-19 diagnostic

**Table 1** Characteristics of the studies included in qualitative analysis

	Total studies	Total pop.	Structural gene primers		Structural and non-structural gene primers		Non-structural gene primers		Not reported	
			N studies	N pop.	N studies	N pop.	N studies	N pop.	N studies	N pop.
<b>Total</b>	<b>85</b>	<b>21 530</b>	<b>22</b>	<b>4265</b>	<b>31</b>	<b>5484</b>	<b>3</b>	<b>123</b>	<b>29</b>	<b>11 658</b>
<b>Location</b>										
Asia (excluding China)	6	378	2	53	4	325				
China	28	12 187	8	1802	17	3047	3	123	24	7215
Europe	12	5757	4	528	8	993			4	4236
North America	10	3001	8	1882	2	1119				
Global		207							1	207
<b>Primers</b>										
<i>N</i> —single	11	2016	11	2016						
<i>E</i> —single	4	759	4	759						
<i>S</i> —single	1	412	1	412						
<i>N, E</i>	2	226	2	226						
<i>S, N</i>	4	852	4	852						
<i>ORF1Ab</i> , single	2	59					2	59		
<i>RdRp</i> , single	1	64					1	64		
<i>E+ORF1Ab</i>	2	1119			2	1119				
<i>E+RdRp</i>	2	259			2	259				
<i>M, E</i>	1	48			1	48				
<i>N+ORF1Ab</i>	14	2703			14	2703				
<i>N+E+RdRp</i>	4	333			4	333				
<i>S, N, E, RdRp, ORF1ab</i>	1	13			1	13				
<i>N, E, ORF1ab</i>	1	33			1	33				
<i>N, RNase P</i>	1	190			1	190				
<i>S, N, RdRp, ORF1ab, E, M</i>	1	52			1	52				
<i>N, S, RdRp</i>	1	273			1	273				
<i>N, E, S, RdRp</i>	2	349			2	349				
<i>S, ORF1Ab</i>	1	112			1	112				
<b>Sample source</b>										
Upper respiratory	23	6748	3	575	9	2633	1	64	10	3476
Upper and lower respiratory	1	52			1	52				
Upper respiratory+other*	9	751	3	368	2	44	1	38	3	301
Lower respiratory+other*	1	273			1	273				
Upper respiratory+serum	20	6407	7	1473	9	1432			4	3502
Upper respiratory+serum+other*	4	941	2	840	1	80	1	21		
Upper and lower respiratory+other*	4	678	1	280	3	398				
Upper and lower respiratory+serum+other*	2	518			1	132			1	386
Serum	18	2376	6	729	4	440			8	1207
Other*	1	199							1	199
Not reported	2	2587							2	2587

\*Other=bronchioalveolar lavage, feces, urine, neonatal, amniotic fluid and breast milk.

E, envelope; N, nucleocapsid; N pop, patient population; ORF1ab, open reading frame 1ab; RdRp, RNA-dependent RNA polymerase; S, spike.

**Table 2** Characteristics of studies included in quantitative meta-analysis

	N studies	N pop.
<b>Total</b>	30	10355
<b>Location</b>		
Asia (excluding China)	2	261
China	19	6375
Europe	7	2900
North America	2	819
<b>PCR primers</b>		
<b>Structural</b>		
N, single	5	1084
E, single	1	49
N, E	1	201
N, S	2	408
<b>Structural and non-structural</b>		
with <i>ORF1Ab</i>	8	1115
with <i>RdRp</i>	2	186
N, <i>RNAse P</i>	1	190
not reported	10	7122
<b>Ab tested</b>		
IgG	2	220
IgM+IgG	25	7828
IgA+IgG+IgM	1	208
IgA+IgG	1	37
not reported	1	2062
<b>Serum Ab</b>		
CLIA	8	3705
ELISA	8	1908
LFIA	10	3800
CLIA+ELISA	2	548
LFIA+ELISA	1	80
Not reported	1	314

CLIA, chemiluminescent immunoassay; E, envelope; LFIA, lateral flow immunoassay; N, nucleocapsid; ORF1ab, open reading frame1ab; RdRp, RNA-dependent RNA polymerase; S, spike.

tests. After excluding duplicate publications, manuscripts that did not report numbers of patients used for sensitivity/specificity calculations and studies with a sample size of <5 patients, 85 studies were selected for qualitative synthesis of RT-PCR primer usage. From this set, a subset of 30 publications were selected for the quantitative meta-analysis of serologic versus RT-PCR diagnostic testing for COVID-19 (online supplemental table S1). Ancillary analysis compared the performance of these 30 real-world studies to that reported in laboratory approval data from 47 diagnostic serum-based tests. In all, our qualitative synthesis of RT-PCR studies included 85 studies and 21 530 patients. From this synthesis, a group of 30 studies with 10 355 patients from 5 regions of the world were selected

for meta-analysis and comparison to performance from laboratory approval data (online supplemental figure S1).

### Uniformity of primer usage in RT-PCR diagnostic tests

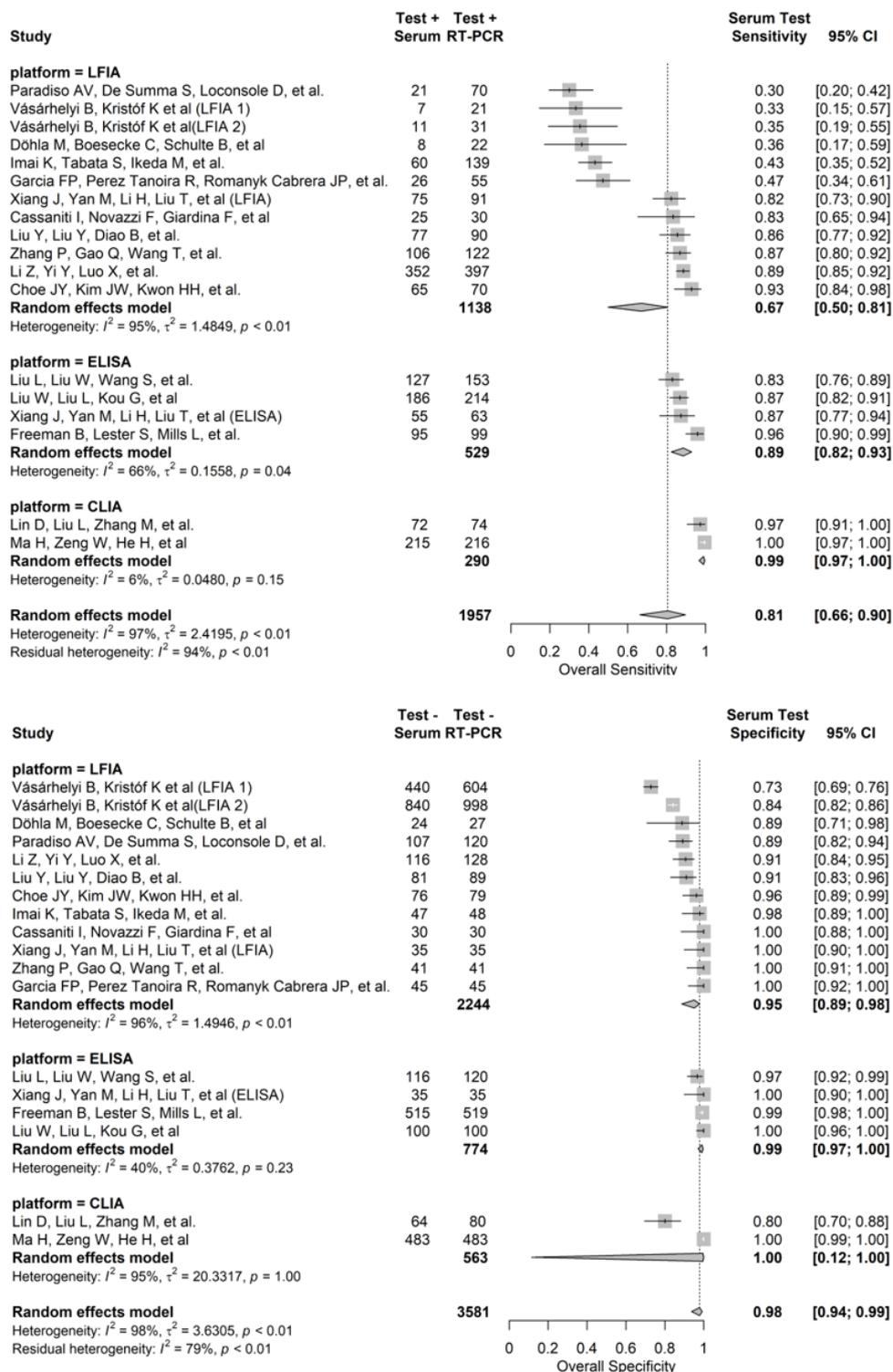
We reviewed use of single primer of structural genes as compared with use of both structural and non-structural gene primers in 56 population-based studies with 9872 participants. Overall, high proportions of studies employed both structural and non-structural gene primers in RT-PCR testing (55% (31 in 56) in studies and 56% (5484 in 9872) of total participants). Additionally, 29 studies (11 658 patients) did not report RT-PCR primer data. Single markers were most frequently tested in China and North American studies (table 1). In general, the most tested samples were from the upper respiratory tract, regardless of primer status. Sample source and location in the respiratory tract were not reported for 8%–20% of patients, and this was more common for studies using single-gene primer.

### Meta-analysis: RT-PCR versus serum antibody testing

In general, patient sera were tested for IgM and IgG antibodies. China was the region with the highest frequency of antibody testing, and lateral flow immunoassay (LFIA) and chemiluminescent immunoassay (CLIA) testing platforms were most often used. Of the 45 studies included in the qualitative synthesis, 30 manuscripts reported both serum antibody testing and RT-PCR testing for the same patients. Key characteristics of this population include: China as the regional location for research; lack of reporting of RT-PCR primer information for ~33% (10/30) of all studies; most studies used IgM and IgG serum-based antibody tests; and LFIA, CLIA and ELISA platforms were common across studies (table 2).

We used the IgM±IgG serum antibody test since it was most commonly used across studies. Of 1957 participants (pooled sensitivity 0.81, 95% CI 0.66 to 0.90) with a positive RT-PCR COVID-19 result, 1585 were also detected as positive with serum antibody tests. Of 3581 true negatives in RT-PCR, 3509 negatives were also found by serum antibody testing (pooled specificity 0.98, 95% CI 0.94 to 0.99). For both models, heterogeneity between studies was significant ( $p < 0.01$  for both,  $I^2 = 97\%$  and  $I^2 = 98\%$  for sensitivity and specificity, respectively).

Subanalyses of differences based on the testing platform found that sensitivity between groups differed ( $p < 0.0001$ ), with CLIA tests performing best (0.99, 95% CI 0.97 to 0.99); ELISA as next best (0.89, 95% CI 0.82 to 0.93); and LFIA as having the poorest sensitivity (0.67, 95% CI 0.50 to 0.81). LFIA test sensitivity also showed heterogeneity between studies ( $p < 0.01$ ,  $I^2 = 95\%$ ). For IgM/IgG tests, specificity did not differ significantly by platform ( $p = 0.06$ ). However, a performance trend followed sensitivity, with LFIA underperforming (figure 1 and online supplemental figure S2).



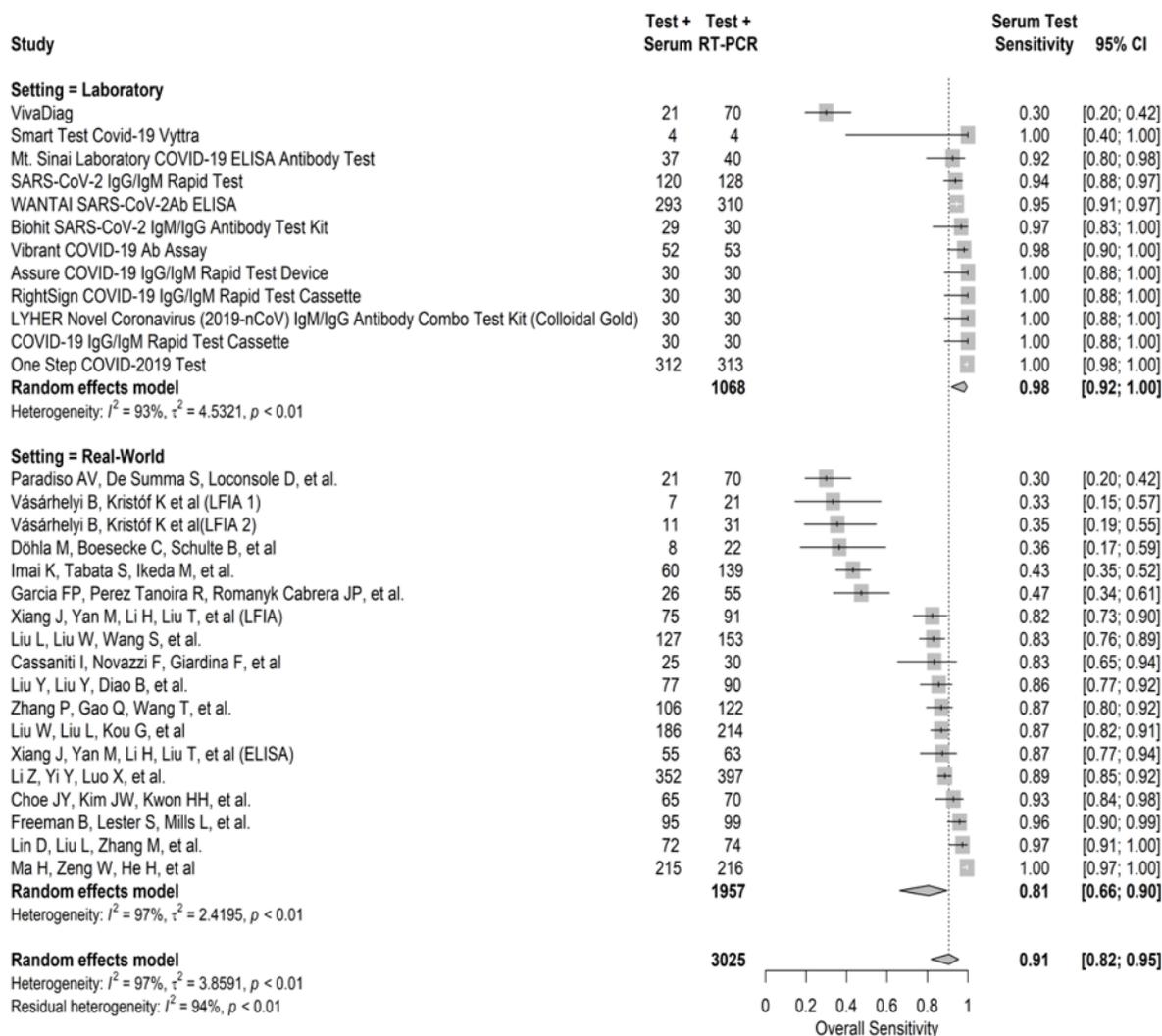
**Figure 1** Comparison of performance (sensitivity and specificity) of serology tests (IgM/IgG) based on assay platforms. CLIA, chemiluminescent immunoassay; LFIA, lateral flow immunoassay.

### Serum antibody testing performance: approval data versus real-world data

For manufacturer-based, laboratory approval data, IgM±IgG testing detected COVID-19 positivity for 1045 of 1068 RT-PCR-determined ‘true’ positive patients (sensitivity 0.98, 95% CI 0.92 to 1.0). In the same group, serum testing correctly identified 1928 of 1967 (specificity 0.98, 95% CI 0.95 to 0.99) true negatives by RT-PCR. For both models (sensitivity and specificity), there was evidence of

heterogeneity ( $p < 0.01$  for both and  $I^2 = 93\%$  and  $I^2 = 94\%$  for sensitivity and specificity, respectively).

We evaluated IgM±IgG serum test performance in subgroup analyses comparing laboratory approval performance data to real-world performance in study data. In manufacturer data presented for approval, serum antibody testing detected 1047 of 1068 ‘true positive’ cases of COVID-19 (sensitivity 0.98, 95% CI 0.92 to 1.0). Real-world use of serum IgM±IgG testing was evident for 2450



**Figure 2** Comparison of sensitivity of laboratory setting versus real-world setting of RT-PCR and serology (IgM/IgG) kits.

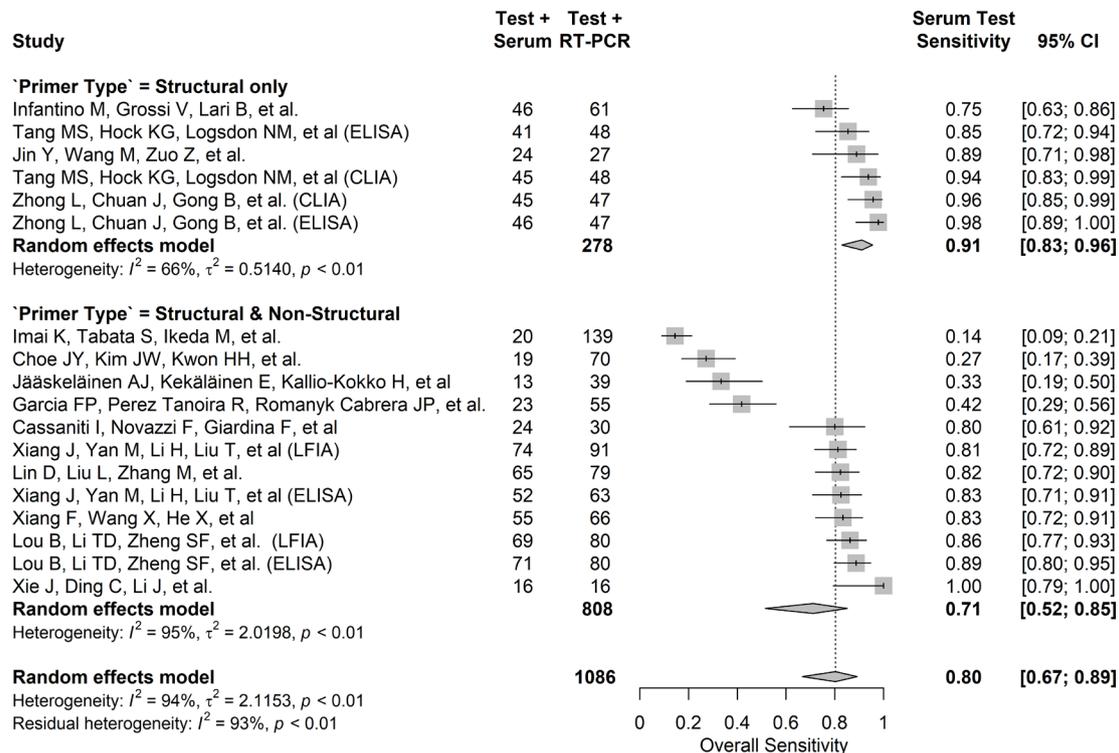
of 3025 participants diagnosed with COVID-19 by RT-PCR (sensitivity 0.81, 95% CI 0.66 to 0.90). For both groups, there was heterogeneity between studies ( $p < 0.01$  for both,  $I^2 = 93%$  and  $I^2 = 97%$  for approval and real-world specificity, respectively) (figure 2 and online supplemental figure S2). In addition, the overall sensitivity between approval and real-world testing groups differed significantly ( $Q = 8.37$ ,  $p = 0.004$ ). An analysis of specificity by the same subgroups found no significant difference between laboratory approval and real-world data. Laboratory data identified 1928 of 1967 participants with true COVID-19 negative status (specificity 0.98, 95% CI 0.95 to 0.99). Real-world data found 5437 of 5548 true negatives (specificity 0.98, 95% CI, 0.96 to 0.99) (analysis not shown).

Since, in IgM±IgG tests, there were differences in sensitivity between platforms, we evaluated the effect of approval-based data versus real-world data by the type of platform. In an analysis stratified for ELISA, CLIA and LFIA, there was no significant difference in specificity between approval and real-world data (data not shown). However, for ELISA tests, real-world capacity to detect true positives was lower than in laboratory-based analyses.

In real-world studies, the sensitivity of ELISA was 0.89 (95% CI, 0.82 to 0.93), different from laboratory sensitivity for the same platform (0.94, 95% CI 0.91 to 0.96,  $Q = 4.74$ ,  $p = 0.03$ ). The LFIA platform also showed a trend of lower real-world sensitivity (0.67, 95% CI, 0.50 to 0.81) compared with laboratory approval sensitivity (0.99, 95% CI 0.90 to 0.99,  $Q = 8.56$ ,  $p = 0.003$ ). Laboratory/real-world groups for CLIA platforms were too small to be tested reliably (1 and 2 groups, respectively).

### Serum antibody testing performance: effect of primer choice on test validity

To evaluate the reliability of RT-PCR as a gold standard for serum-based test performance, we tested the consequences of using structural and non-structural primers in RT-PCR reference testing of serum. Analyses were divided into three subgroups based on antibody targets: IgM, IgG and IgG±IgM combined. In IgM and combined IgG±IgM testing, the primer choice had no significant influence on sensitivity or specificity. However, for IgG antibody tests, use of both a structural and a non-structural gene-specific primers in RT-PCR resulted in reduced sensitivity for



**Figure 3** The effect of primer choice in RT-PCR referent on sensitivity of tests based on serum IgG.

serum testing (figure 3 and online supplemental figure S2),  $Q=6.17$ ,  $p=0.013$ ). Furthermore, although not statistically significant, the sensitivity of both IgM and IgG±IgM combined data sets was lower when using a referent RT-PCR test with both primer types.

## DISCUSSION

Because of the highly infectious nature of COVID-19, a prompt, accurate and early diagnosis is necessary to deal with the ongoing pandemic, for such diagnoses can help reduce the spread of infection and its associated risk for mortality. Currently, the COVID-19 diagnosis is generally based on RT-PCR assays.<sup>8</sup> Alternative methods such as antigen-based and antibody-based serology tests, although available, have uncertain value. The current systematic review and meta-analysis addresses the challenges encountered in the diagnosis of COVID-19 by various methods. It also analyses differences between the FDA-approved EUA data and real-world data. There is worldwide non-uniformity in the performance of RT-PCR, including the number and types of primers and reagents used for COVID-19 diagnosis, which raise questions about its generalised applicability. Similarly, the studies based on serological tests showed diagnostic inaccuracies owing to individual differences in mounting an immune response as well as dependency on the time duration after the onset of symptoms. Overall, the sensitivity between RT-PCR and serology tests was 0.81 (95% CI, 0.66 to 0.90), and specificity was 0.98 (95% CI, 0.94 to 0.99). Among the various platforms for serodiagnosis, the highest sensitivity was exhibited by ELISA, followed by CLIA and LFIA. Furthermore, use of primers (structural, non-structural or

both) had a variable effect on sensitivity based on antibody targets. Sensitivity was significantly higher for IgG serology tests using structural primer-only RT-PCR tests as a referent. Serology tests had higher sensitivity for approval-based data than for real-world reporting. This difference was significant for ELISA-based platforms, and a non-significant trend towards inflated approval-based sensitivity was evident for both CLIA and LFIA platforms. These observations highlight the inconsistencies/challenges in the COVID-19 diagnosis by RT-PCR, which is the current gold standard, as well as in serologic testing.

For RT-PCR assays, the targets in SARS-CoV-2 include structural genes such as *E*, *N* and *S*, and nonstructural genes, including that for *RdRp* or *ORF1ab*.<sup>22</sup> In the early phase of the pandemic, some studies used a two-step diagnosis that included an initial screening phase using structural genes followed by a confirmatory phase using nonstructural genes.<sup>6 7 23</sup> The test is considered positive when both structural and non-structural markers are positive.<sup>24 25</sup> However, currently both types of primers are used simultaneously to diagnose COVID-19. The viral load or copy number of the viral genome is expressed as a Ct value, which when  $<37$  is indicative of a positive test, and a value of  $\geq 40$  is considered negative. A Ct value between  $>37$  and  $<40$  requires repetition of RT-PCR analysis to confirm the diagnosis.<sup>24</sup> However, the Ct value range varies widely according to assays and laboratory practices. A COVID-19-RdRp/Hel assay has a higher sensitivity than a conventional RdRp-P2 assay irrespective of the type of sample.<sup>26</sup> Overall, higher proportions of studies (58%) employed both structural and non-structural gene primers in RT-PCR testing. Single markers were used

in some Chinese and North American studies. These findings are indicative of non-uniformity in the RT-PCR methodology. We note that half of the positive, symptomatic patients became negative by the second week, when they became asymptomatic. In contrast, the asymptomatic, positive patients became negative 2 days after hospital admission, indicating the importance of a temporal factor in COVID-19 diagnosis by RT-PCR.<sup>27 28</sup>

Published in the early phase of the pandemic, 11 of 85 studies had clinically suspected COVID-19 patients. In these studies, the average test positivity by RT-PCR, regardless of the sample source, was 44% (online supplemental table S1), and test sensitivity was influenced by sample source (upper vs lower respiratory vs other samples), issues related to testing performance, and delay after onset of symptoms.<sup>29</sup> In the early phase of the COVID-19 pandemic, for studies evaluating suspected COVID-19 cases, the total positive RT-PCR for throat swabs was in the range of 30%–60% at initial presentation.<sup>8 30</sup> One study reported a yield of 72%–93% positive cases for lower respiratory samples (bronchioalveolar lavage and sputum) as compared with 32%–63% positivity for upper respiratory samples (oral and nasopharyngeal swabs) and 29% for stool samples.<sup>29</sup> Hence, a negative COVID-19 test based only on an upper respiratory sample at a single time-point is questionable. For most studies, the testing sample was from the upper respiratory tract, regardless of primer type used. However, the sample source was not reported for 8%–20% of patients, which was more common for studies using only structural gene primers. For stool samples testing positive for COVID-19, 66.7% also tested positive on pharyngeal swabs. Of the stool samples, 64.3% remained positive after pharyngeal clearance of the virus.<sup>31</sup> In contrast, none of the patients showed a positive test on upper respiratory samples after the anal swabs tested negative.<sup>31</sup> These findings raise concerns about whether patients with negative respiratory swabs are truly virus-free, and sampling of additional body sites is needed. As determined by various studies, the performance of the RT-PCR depends on usage of comparable protocols, including primers and reagents.<sup>32</sup> Additionally, it is unknown whether the currently used RT-PCR primers detect all SARS-CoV-2 strains.

The specific immune response to SARS-CoV-2 can be measured by serological testing. Several rapid serological tests, including point-of-care tests, are being developed. Even though some of these tests have been approved by the FDA through EUA, their accuracy needs to be validated.<sup>33</sup> A minimum of 1–2 weeks after the onset of infection is needed for seroconversion. Hence, antibody testing is of no value in the early phase of infection. Additionally, its value is limited by its cross-reactivity with other coronaviruses.<sup>34 35</sup> The initial RT-PCR positivity during the early stages (<15 days) of SARS-CoV-2 infection declines to 66.7% in the later phase (15–39 days), during this period, the antibody test can supplement RT-PCR in the diagnosis of COVID-19.<sup>34 35</sup> Additionally, serology testing becomes valuable for clinically suspected and RT-PCR negative (false-negative) individuals.

This research has limitations. Due to the dynamic reporting of COVID-19 testing research and inconsistencies

in reporting of predictive variables across studies, bias in sampling may have some effect on our results. Patient flow analysis suggests that lack of consistent RT-PCR reference standard given to patients in the same study, as well as the unclear reporting of patient selection methods could contribute to bias in these results (online supplemental figure S3). In addition, the observed heterogeneity between studies in the meta-analysis suggests that we must consider the possibility that the differences in results may be due to chance. Lastly, it is questionable to compare two separate testing methods of RT-PCR and seroprevalence in sensitivity/specificity analysis. In particular, given the relationship between time since diagnosis and accuracy of serology testing, a contributor to the observed differences in performance is time. Furthermore, because of each diagnostic kit having differing cut points for positive/negative, threshold effect as a source of heterogeneity in sensitivity and specificity cannot be ruled out.

The effective containment of COVID-19 involves accurate diagnoses and isolation of SARS-CoV-2-infected persons. Robustness of the assays/platforms is determined by variability of the samples, primers and reagents used. Serological tests alone are of value only during the latter times of infection; however, they complement RT-PCR when used in conjunction and minimise false-negative RT-PCR results. Additionally, some of the approved serological assays/platforms, particularly those developed using contrived/laboratory data, perform poorly when applied to real-world samples. We are currently in a new phase of the pandemic, and there is a need for a reliable/robust diagnostic test to mitigate the spread of COVID-19.

Our analyses of studies published in the early phase of the pandemic have highlighted issues related to COVID-19 diagnosis that need to be addressed as follows: (1) The high mutational rate exhibited by the SARS-CoV-2 virus may lead to the generation of new strains. Therefore, like for influenza virus, the existing diagnostic kits need to be modified constantly to optimise the detection of new strains; (2) Though RT-PCR diagnosis of COVID-19 is the gold standard, its combination with a serologic test may increase the accuracy of SARS-CoV-2 detection; (3) Approval agencies must account for the type of data (contrived vs real world) presented by diagnostic kit developer; (4) Although agencies employed EUA processes for the approval of diagnostic kits, there is a need to monitor their performance and assess their robustness in real-world samples, to permit continued use of these kits; and (5) Standardised assay protocols need to be developed and continually updated to mitigate the COVID-19 pandemic.

#### Author affiliations

<sup>1</sup>Pathology, MNR Medical College and Hospital, Sangareddy, Telangana, India

<sup>2</sup>Department of Pathology at UAB, The University of Alabama at Birmingham School of Medicine, Birmingham, Alabama, USA

<sup>3</sup>Sri Guru Ram Das Institute of Medical Sciences and Research, Amritsar, Punjab, India

<sup>4</sup>Department of Medicine, The University of Alabama at Birmingham, Birmingham, Alabama, USA

<sup>5</sup>Pathology, The University of Alabama at Birmingham School of Medicine, Birmingham, Alabama, USA

**Acknowledgements** We thank Dr. Donald Hill for his editorial assistance.

**Contributors** CS, MB and UM made substantial contributions to the conception and design of the work. CS, MB and VL contributed to the acquisition of the study data. Data were analysed by CS, MB and UM. MB was responsible for the statistical analysis. CS, MB, SML, SV, GJN and UM contributed to the interpretation of the data, and to drafting and revising of the manuscript. CS, MB and VL have accessed and verified the data. UM serves as the guarantor of overall content. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work. All authors have read and approved the manuscript.

**Funding** This study was supported, in part, by grant 5U54CA118948 and institutional/departmental impact funds (UM), and by CA047888, the UAB Cancer Prevention and Control Training Program (MB).

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. All data used in this research was generated from publicly available published manuscripts.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Chandrakumar Shanmugam <http://orcid.org/0000-0002-4782-4512>

Vishwas Luthra <http://orcid.org/0000-0002-5867-7781>

Upender Manne <http://orcid.org/0000-0002-1545-3032>

#### REFERENCES

- Chan JF-W, Yuan S, Kok K-H, *et al.* A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 2020;395:514–23.
- Naming the coronavirus disease (COVID-19) and the virus that causes it, 2020. Available: [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it). [Accessed 10 Jun 2020].
- Huang C, Wang Y, Li X, *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497–506.
- Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J Autoimmun* 2020;109:102433.
- Cheng MP, Papenburg J, Desjardins M, *et al.* Diagnostic testing for severe acute respiratory syndrome-related coronavirus 2: a narrative review. *Ann Intern Med* 2020;172:726–34.
- Chu DKW, Pan Y, Cheng SMS, *et al.* Molecular diagnosis of a novel coronavirus (2019-nCoV) causing an outbreak of pneumonia. *Clin Chem* 2020;66:549–55.
- Corman VM, Landt O, Kaiser M, *et al.* Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 2020;25.
- Ai T, Yang Z, Hou H, *et al.* Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology* 2020;296:E32–40.
- Xie X, Zhong Z, Zhao W, *et al.* Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing. *Radiology* 2020;296:E41–5.
- Tahamtan A, Ardebili A. Real-Time RT-PCR in COVID-19 detection: issues affecting the results. *Expert Rev Mol Diagn* 2020;20:453–4.
- Kim JY, Ko JH, Kim Y, *et al.* Viral load kinetics of SARS-CoV-2 infection in first two patients in Korea. *J Korean Med Sci* 2020;35:e86.
- Guo L, Ren L, Yang S, *et al.* Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). *Clin Infect Dis* 2020;71:778–85.
- To KK-W, Tsang OT-Y, Leung W-S, *et al.* Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* 2020;20:565–74.
- Wang W, Xu Y, Gao R, *et al.* Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA* 2020;2:1843–4.
- Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010;8:336–41.
- EUA authorized serology test performance, 2020. Available: <https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/eua-authorized-serology-test-performance> [Accessed 18 Jun 2020].
- SARS-CoV-2 diagnostic pipeline SARS-CoV-2 diagnostic pipeline. Foundation for innovative new diagnostics, 2020. Available: <https://www.finddx.org/covid-19/pipeline/> [Accessed 5 Jul 2020].
- Whiting PF, Rutjes AWS, Westwood ME, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
- R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2019.
- Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;22:153–60.
- Higgins JPT, Thompson SG, Deeks JJ, *et al.* Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
- Yin C. Genotyping coronavirus SARS-CoV-2: methods and implications. *Genomics* 2020;112:3588–96.
- Detection of 2019 novel coronavirus (2019-nCoV) in suspected human cases by RT-PCR who, 2019. Available: [https://www.who.int/docs/default-source/coronaviruse/peiris-protocol-16-1-20.pdf?sfvrsn=af1aac73\\_4](https://www.who.int/docs/default-source/coronaviruse/peiris-protocol-16-1-20.pdf?sfvrsn=af1aac73_4)
- Liu R, Han H, Liu F, *et al.* Positive rate of RT-PCR detection of SARS-CoV-2 infection in 4880 cases from one hospital in Wuhan, China, from Jan to Feb 2020. *Clin Chim Acta* 2020;505:172–5.
- Li X, Geng M, Peng Y, *et al.* Molecular immune pathogenesis and diagnosis of COVID-19. *J Pharm Anal* 2020;10:102–8.
- Chan JF-W, Yip CC-Y, To KK-W, *et al.* Improved Molecular Diagnosis of COVID-19 by the Novel, Highly Sensitive and Specific COVID-19-RdRp/Hel Real-Time Reverse Transcription-PCR Assay Validated *In Vitro* and with Clinical Specimens. *J Clin Microbiol* 2020;58. doi:10.1128/JCM.00310-20. [Epub ahead of print: 23 04 2020].
- Sethuraman N, Jeremiah SS, Ryo A. Interpreting diagnostic tests for SARS-CoV-2. *JAMA* 2020;323:2249–51.
- Chen J, Qi T, Liu L, *et al.* Clinical progression of patients with COVID-19 in Shanghai, China. *J Infect* 2020;80:e1–6.
- Wang W, Xu Y, Gao R, *et al.* Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA* 2020;2:1843–4.
- Yang Y, Yang M, Shen C. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. *medRxiv* 2020:02.11.20021493.
- Chen Y, Chen L, Deng Q, *et al.* The presence of SARS-CoV-2 RNA in the feces of COVID-19 patients. *J Med Virol* 2020;92:833–40.
- NHS. Guidance and standard operating procedure: COVID-19 virus testing in NHS laboratories, 2020. Available: <https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/guidance-and-sop-covid-19-virus-testing-in-nhs-laboratories-v1.pdf> [Accessed 2 Jun 2020].
- Maxmen A. The researchers taking a gamble with antibody tests for coronavirus. *Nature* 2020.
- Zhao J, Yuan Q, Wang H, *et al.* Antibody responses to SARS-CoV-2 in patients with novel coronavirus disease 2019. *Clin Infect Dis* 2020;71:2027–34.
- WHO. *Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases: interim guidance*. WHO, 2020.