medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FALSE-NEGATIVE RESULTS OF INITIAL RT-PCR ASSAYS FOR COVID-19: A SYSTEMATIC REVIEW

AUTHORS

Ingrid Arevalo-Rodriguez, MSc, PhD[1]; Diana Buitrago-Garcia, MSc[2], Daniel Simancas-Racines, MD, PhD[3], Paula Zambrano-Achig, MSc[4], Rosa Del Campo, MSc, PhD[5], Agustín Ciapponi, MD, PhD[6], Omar Sued, MD[7], Laura Martínez-García, MSc[8], Anne Rutjes, PhD[9], Nicola Low, PhD[10], Jose A Perez-Molina, MD, PhD[11], Javier Zamora, MSc, PhD[12,13]

1. Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal, IRYCIS, CIBER of Epidemiology and Public Health, Madrid, Spain.
2. Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland.
3. Centro de investigación en Salud Pública y Epidemiología Clínica (CISPEC). Facultad de Ciencias de la Salud "Eugenio Espejo", Universidad UTE, Ecuador.
4. Centro de investigación en Salud Pública y Epidemiología Clínica (CISPEC). Facultad de Ciencias de la Salud "Eugenio Espejo", Universidad UTE, Ecuador.
5. Department of Microbiology, Ramón y Cajal University Hospital, Ramón y Cajal Health Research Institute (IRYCIS), Madrid, Spain.
6. Instituto de Efectividad Clínica y Sanitaria (IECS-CONICET), Buenos Aires, Argentina
7. Fundación Huésped, Buenos Aires, Argentina.
8. Department of Microbiology, Ramón y Cajal University Hospital, Ramón y Cajal Health Research Institute (IRYCIS), CIBER of Epidemiology and Public Health, Madrid, Spain.
9. Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland.
10. Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland.
11. National Referral Centre for Tropical Diseases, Infectious Diseases Department, Hospital

FINAL DRAFT

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Universitario Ramón y Cajal, Insituto Ramón y Cajal de Investgación Sanitaria, Madrid, Spain.

12. Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal, IRYCIS, CIBER of Epidemiology and Public Health, Madrid, Spain.
13. Barts and the London School of Medicine and Dentistry, Queen Mary University London (UK)

Corresponding author:

Ingrid Arevalo-Rodriguez

Clinical Biostatistics Unit, Hospital Universitario Ramón y Cajal, IRYCIS

CIBER of Epidemiology and Public Health

Madrid, Spain

Email: ingrid.arevalo@salud.madrid.org

ORCID ID: 0000-0002-7326-4504

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ABSTRACT

Background: Cases with negative reverse transcription-polymerase chain reaction (RT-PCR) results at initial testing for suspicion of SARS-CoV-2 infection, and found to be positive in a subsequent test, are considered as RT-PCR false-negative cases. False-negative cases have important implications for COVID-19 management, isolation, and risk of transmission. We aimed to review and critically appraise evidence about the proportion of RT-PCR false-negatives at initial testing for COVID-19.

Methods: We performed a systematic review and critical appraisal of literature with high involvement of stakeholders in the review process. We searched on MEDLINE, EMBASE, LILACS, the WHO database of COVID-19 publications, the EPPI-Centre living systematic map of evidence about COVID-19, and the living systematic review developed by the University of Bern (ISPM). Two authors screened and selected studies according to the eligibility criteria and collected data of included studies (no-independent verification). Risk of bias was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool. We calculated the false-negative proportion with the corresponding 95% CI using a multilevel mixed-effect logistic regression model using STATA 16®. Certainty of the evidence about false-negative cases was rated using the GRADE approach for tests and strategies. The information is current up to 6 April 2020.

Findings: Five studies enrolling 957 patients were included. All studies were affected by several biases and applicability concerns. Pooled estimation of false-negative proportion was 0.085 (95% CI= 0.034 to 0.196; tau-squared = 1.08; 95% CI= 0.27 to 8.28; p<0.001); however, this estimation is highly affected by unexplained heterogeneity, and its interpretation should be avoided. The certainty of the evidence was judged as very low, due to the risk of bias, indirectness, and inconsistency issues.

Conclusions: The collected evidence has several limitations, including risk of bias issues, high heterogeneity, and concerns about its applicability. Nonetheless, our findings reinforce the need for

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

repeated testing in patients with suspicion of SARS-Cov-2 infection given that up to 29% of patients could have an initial RT-PCR false-negative result.

Systematic review registration: Protocol available on OSF website: https://osf.io/gp38w/

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

BACKGROUND

On December 31, 2019, the World Health Organization (WHO) was alerted about a cluster of pneumonia patients in the city of Wuhan, in China's Hubei province [1]. Chinese authorities confirmed a week later the outbreak of a novel coronavirus currently called Severe Acute Respiratory Coronavirus 2 (SARS-CoV-2) [2]. This new virus is the underlying cause of Coronavirus Disease 2019 (COVID-19), which has become a worldwide public health emergency and reached pandemic status [3]. By the time of this article's writing, the virus has spread to 212 countries and territories and has caused over 85,837 deaths worldwide [4].

Patients with COVID-19 exhibit respiratory symptoms such as fever, cough, and shortness of breath as primary manifestations [5, 6]. Although most of the cases present mild symptoms, some cases have developed pneumonia, severe respiratory diseases, kidney failure and even death [7-9]. SARS-CoV-2 mainly spreads through person-to-person contact via respiratory droplets from coughing and sneezing, and through surfaces that have been contaminated with these droplets.[10] Recent studies have suggested the presence of asymptomatic cases in cluster families, possibly transmitting the virus before a virus-carrying person displays any symptom [11].

Because the signs of infection mentioned above are non-specific, confirmation of cases is currently based on the detection of a viral sequence by reverse transcription-polymerase chain reaction (RT-PCR). Different RT-PCR schemes have been proposed; all of them include the N gene that codes for the viral nucleocapsid. Other alternative targets are the E gene, for the viral envelope, or the S gene for the spike, and the Hel gene for the RNA polymerase gene (RdRp/Helicase) [12, 13]. Molecular criteria for *in vitro* diagnosis of COVID-19 disease are heterogeneous, and usually require the detection of two or more genes of SARS-CoV-2 [14].

RT-PCR repeated testing might be required to confirm a clinical diagnosis, especially in the presence of symptoms close related to COVID-19 disease [15]. Cases with negative RT-PCR results at initial

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

testing and later found to be positive in a subsequent test are commonly considered cases with an initial false-negative result. Some researchers have suggested that these failures in SARS-CoV-2 detection are related to multiple pre-analytical and analytical factors, such as lack of standardisation to collect specimens, the time and conservation of samples until to be received in the laboratory, the use of non-adequately validated assays, contamination during the procedure, insufficient viral specimens and load, the incubation period of the disease, and the risk of active recombination and mutation [14, 16].

The availability of accurate laboratory tools for COVID-19 is essential for case identification, contact tracing, and optimization of infection control measures, as it was shown by previous epidemics caused by SARS-CoV and Middle East respiratory syndrome coronavirus (MERS-CoV) [17-19]. Due to the COVID-19 pandemic causing an important burden on health systems around the globe, and considering that a missing COVID 19 case might have severe consequences at several levels, we aimed to estimate through a systematic review of the literature the proportion of false-negatives related to the detection of SARS-CoV-2 using RT-PCR assays at the initial laboratory test.

METHODS

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) for diagnostic test accuracy (DTA) to perform this report [20]. For the development of this systematic review of literature, we used selected methods for the development of rapid reviews, such as a high involvement of stakeholders in the review process (including the definition of the review question), a non-independent verification of data selection and extraction, and parallelisation of tasks (that is, to perform selected activities simultaneously instead of consecutively). Other review shortcuts and omission of review tasks were not applied. A protocol of this review was published in the Open Science Framework repository for public consultation (https://osf.io/gp38w/).

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Criteria for considering studies for this review

We included observational studies (including accuracy studies, cohorts, and case series) reporting the initial use of RT-PCR to the detection of SARS-CoV-2 RNA in patients under suspicion of infection by clinical or epidemiological criteria. Specially, we prioritised studies enrolling consecutive patients who were receiving RT-PCR as initial testing with further confirmation of SARS-CoV-2 infection and/or COVID-19 diagnosis (positive/negative). We did not impose limits by age, gender, or study location.

We aimed to include all types of RT-PCR kits, regardless of the brand/manufacturer, the RNA extraction method used, the number of target gene assays assessed and cycle threshold value for positivity. Studies comparing the accuracy of two or more tests for COVID-19 diagnosis were also considered if we could abstract the fraction of negative test results as defined by an initial RT-PCR assay.

We excluded studies without clear information about false-negative cases, the number of final confirmed cases, or an unclear verification of negative cases. Case reports, studies based on laboratory samples, and literature reviews were also excluded.

Search methods for identification of studies

We carried out a comprehensive and sensitive search strategy based on the proposal for the living systematic review developed by the University of Bern's Institute of Social and Preventive Medicine-ISPM in the following databases:

- MEDLINE (Ovid SP, 1946 to April 6th, 2020)
- Embase (Ovid SP, 1982 to April 6th, 2020)
- LILACS (iAH English) (BIREME, 1982 to April 6th, 2020)

We did not apply any language restrictions to electronic searches (S1 Appendix). As additional sources of potential studies, we searched in repositories of preprint articles (such as Medrxiv),

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

clinical trials registries for ongoing or recently completed trials (clinicaltrials.gov; the World Health Organization's International Trials Registry and Platform, and the ISRCTN Registry), and the reference lists of all relevant papers. Finally, we also screened the following resources for additional information:

- The WHO Database of publications on coronavirus disease (COVID-19). Available on https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov.
- The Living systematic map of the evidence about COVID-19 produced by EPPI-Centre. Available on: http://eppi.ioe.ac.uk/cms/Projects/DepartmentofHealthandSocialCare/Publishedreviews/COVID19Livingsystematicmapoftheevidence/tabid/3765/Default.aspx
- The Living systematic review developed by the Institute of Social and Preventive Medicine-ISPM from the University of Bern available on https://ispmbern.github.io/covid-19/

Data collection and analysis

For the selection of potential studies, one reviewer screened the search results based on the title and abstract, with additional verification by a second reviewer (no-independent verification). We retrieved the full-text copy of each study assessed as potentially eligible, and pairs of reviewers confirmed eligibility according to the selection criteria (non-independent verification). In case of disagreements we reached consensus by discussion. For data extraction one reviewer extracted qualitative and quantitative data from eligible studies. An additional reviewer checked all the extracted information for accuracy (non-independent verification of data extraction).

Assessment of methodological quality

We assessed the methodological quality of accuracy studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool [21]. Due to the lack of tools to assess the risk of bias

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

associated with case series, we decided to apply the QUADAS-2 tool in case of inclusion of this type of report.

Statistical analysis and data synthesis

For all included studies, we extracted data about the number of cases initially considered as negative (i.e. false-negative cases) as well as the total of confirmed cases in further investigations. We presented the results of estimated proportions (with 95% CIs) in a forest plot, in order to assess the between-study variability. We aimed to calculate the false-negative rate with the corresponding 95% CI using a multilevel mixed-effect logistic regression model implemented in Stata 16®'s metaprop_one command. This allowed us to estimate the between-study heterogeneity from the variance of study-specific random intercepts. We assessed the heterogeneity between the results of the primary studies using the Tau-square statistic. A probability value less than 0.1 ($p<0.1$) was considered to suggest statistically significant heterogeneity and preclude a pooled result of numerical data.

We planned to investigate the potential sources of heterogeneity using a descriptive approach and performing a random-effects meta-regression analysis. Anticipated sources of heterogeneity included the type of specimen collected, the presence or not of clinical findings, the number of RNA targets genes under assessment, and the time of symptom evolution.

Summary of findings and certainty of the evidence

We rated the certainty of the evidence about false-negative cases following the GRADE approach for tests and strategies [22, 23]. We assessed the quality of evidence as high, moderate, low or very low, depending on several factors including risk of bias, imprecision, inconsistency, indirectness, and publication bias. We illustrate the consequences of the numerical findings in a population of 100 tested, according to three different prevalence estimates of the disease provided by the stakeholders involved in this review.

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Patient and public involvement

We involved several stakeholders in the design, conduct, and reporting of our research, including general and family physicians, specialists on infectious disease and microbiologists currently attending patients under suspicion of COVID 19 disease. The study protocol and preliminary results are publicly available on https://osf.io/gp38w/.

RESULTS

Electronic searches yielded 662 references from selected databases. In addition, we obtained 186 additional references searching in other resources (Figure 1). Our initial screening of titles and abstracts identified 61 references to assess in full text. We excluded 54 studies due to: a) case reporting fewer than five patients; b) unclear information about the results of initial RT-PCR and/or false-negatives; c) reviews and state-of-art; d) other reasons (S2 Appendix). Two studies were not available in full-text despite requests to their authors. We included five studies in qualitative and quantitative synthesis [24-28] which included 957 patients.

The sample size ranged from 36 to 601 confirmed cases (median 102 patients). All included studies were in pre-print status. Three studies were focused on accuracy estimations [24, 26, 27], while two additional studies reported information of a case series [25, 28]. Data collection of cases ranged from January 6 to February 8-2020. All studies were performed in institutions based in China (Figure 2). The age of participants ranged from 44 to 51 years (information derived from three studies) [24, 25, 27]. There were 577 men versus 213 women included (Table 1). Three studies included patients under suspicion of COVID-19 due to clinical findings and/or epidemiological criteria [24, 26, 27]. Confirmation of infection was performed after isolation of SARS-CoV-2 in any real-time RT-PCR assay for 2019-nCoV, including repeated RT-PCR after negative results (two or more). Three studies provided information about the proportion of confirmed cases with positive chest CT findings,

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ranging from 74 to 98%. One study provided information about the time from the symptom onset to CT scan as a proxy for the duration of disease [25], and a second one reported duration of fever [27].

Regarding RT-PCR testing, the RT-PCR brand/manufacturer was reported by two studies [24, 25], No studies reported criteria for positivity. Most of the studies based their assessment on throat samples, such as pharyngeal, nasal and oropharyngeal swabs. Four studies provided information about the time since the initial RT-PCR to repeated testing (Table 1).

FINAL DRAFT

217 **Table 1.** Characteristics of included studies

ID	Data collection	Setting	Age (years)	N total (male: female)	Type of specimen	RT-PCR Brand	Days from symptoms onset (days)	Interval between initial RT-PCR to additional RT-PCRs
Ai T 2020	January 6 to February 6	Tongji Hospital of Tongji MedicalCollege of Huazhong University of Science and Technology, Wuhan, Hubei, China	Mean 51 ± 15 Range of 2 to 95	601 (467:134)	Throat swab	TaqMan One-Step RT-PCR Kits from Shanghai Huirui Biotechnology Co., Ltd or Shanghai BioGerm Medical Biotechnology Co., Ltd,	Not reported	Median: 4 days (range from 4-8 days)
Bernheim A 2020	January 18 to February 2	Hospitals from four provinces in China: Nanchang (Jiangxi Province), Zhuhai (Guangdong Province), Chengdu (Sichuan province) and Guilin (Guangxi province)	Mean 45 ±15,6	102 (61:41)	Bronchoalveolar lavage, endotracheal aspirate, nasopharyngeal swab, or oropharyngeal swab	Sansure Biotech Inc. (Changsha, China), Shanghai Zhijiang Biotechnology Co. (Shanghai, China), or Da An Gene Co. (Guangzhou, China).	Range to 0 to 12	Not reported
Fang Y 2020	January 19 to February 4	Taizhou Enze Medical Center (Group) Enze Hospital	Median 45 Interquartile range of 39-55	51 (29:22)	Throat swab, sputum	Not reported	Mean 3±3	Repeat testing was performed at intervals of 1 day or more; no further information provided
Long C 2020	January 20 to February 8	Yichang Yiling Hospital	Mean 44,8 ±18,2	36 (20:16)	Not reported	Not reported	Only duration of fever reported: 2,6 ± 1,7 days	Range from 2 to 8 days
Xie X 2020	January 16 to February 2	Database of Radiology Quality Control Centre, Hunan/ 3 cities in Hunan Province, China	Not reported	167 (Not reported)	swab test; no further details provided	Not reported	Not reported	Range from 2 to 8 days

218

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Quality of included studies

We applied the QUADAS-II tool to all included studies to reflect critical limitations in the validity of the findings (Figure 3 and S3 Appendix). The reference standard domain was the most affected by the potential risk of bias due to the lack of independence between the index test and the confirmation of cases (repeated RT-PCR testing). Details about the criteria for positivity were not provided by all included studies, and this domain was judged as under unclear risk and unclear applicability concerns. In addition, the applicability of patient selection was judged as with great concerns due to most of the studies selected patients who underwent both RT-PCR and Chest CT, excluding patients who can be candidates to receive the index test in the current clinical practice.

Findings

We analyse information from five studies collecting information from 957 patients confirmed to have SARS-CoV-2 infection and 53 cases with RT-PCR negative findings in their initial assessment (Figure 4). False-negative proportions ranged from 0.02 [24] to 0.29 [26]. Only one study provided subgroup information about time since illness onset to CT scans [25], as a proxy of the time of symptom evolution, with proportions ranged from 0.15 (≤2 days) to 0.08 (3 or more days) (Figure 4).

The pooled estimation of false-negative proportion was 0.085 (95% CI 0.034 to 0.196) estimated by a mixed-effects logistic regression model. However, pooled data is affected by a considerable between-study heterogeneity (tau-squared = 1.08; 95% CI= 0.27 to 8.28; p<0.001), since we are not able to warrant that the average estimation provided by the meta-analysis is a valid and representative estimation of the true value of the false-negative proportion in the current practice, we instead used the range of proportions in the analysis of the certainty of the evidence using the GRADE approach. A full exploration of heterogeneity was not possible given that: a) most of the

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

studies collected upper or lower respiratory specimens; b) all studies included patients with clinical findings suggestive of COVID 19 disease; c) subgroup information by the time of evolution of symptom was only provided by one study; and d) key information about the characteristics of the index test, such as positivity criteria, were not reported. The high variability of pooled estimation was not reduced with the separate estimation of false-negative proportion by type of study (accuracy versus case series).

Certainty of the evidence

We use the range of false-negative proportions to develop a summary of findings following the GRADE approach. The quality of the evidence was judged to be very low due to issues related to the risk of bias, indirectness, and inconsistency (Figure 5). We illustrate the consequences of the range of false-negative proportions in a population of 100 tested, according to three different prevalence seen in the current clinical practice for participant stakeholders (30%, 50%, and 80%) (Figure 5). Using a prevalence of 50%, we found that 1 to 14 cases would be misdiagnosed and then they could no receive adequate clinical management, and they could require repeated testing at some point of their hospitalization or even they could require other investigations for competitive diagnoses. This numerical approach should be interpreted with caution due to the multiple limitations of the evidence described above (Figure 5).

DISCUSSION

Our systematic review included five studies and 957 participants providing information about the proportion of false-negative cases related to the detection of SARS-CoV-2 by RT-PCR assays at first use. The included studies enrolled patients under suspicion of COVID 19 [24, 26, 27] or confirmed COVID 19 cases [25, 28]. Almost all studies enrolled a selected sample of patients (i.e. patients with

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

findings for RT-PCR and chest CT) from several provinces of China and collected between January to February 2020. We considered all studies to be affected by several sources of bias, especially related to the independence between the index test and the reference standard and the unclear report of key RT-PCR characteristics. A meta-analysis of the proportions using Stata® showed a considerable heterogeneity not explained by the collected data, and this variability is a limitation for the full interpretation of averaged proportion. As an alternative, we preferred to provide an analysis of the range of false-negative proportions derived from included studies in a cohort of 100 patients tested and using three different prevalence of the disease derived from the current clinical practice of our participant stakeholders. Using a prevalence of 80%, we found that 2 to 23 cases would be misdiagnosed and then they could no receive adequate clinical management. However, we emphasized that this numerical approach should be interpreted with caution due to the multiple limitations of the evidence described above (Quality of evidence: Very low).

Although we did not impose restrictions on population characteristics such as age, setting or publication status, we noticed that our findings are limited due to all the studies were performed in one country (China), and they reported data only for the beginning of the pandemic (January 2020), in addition to the lack of reporting about the index test previously mentioned. RT-PCR kits in use for included studies were likely the first kits developed by detection of SARS-CoV-2, and then the tests currently in use might have a great technological evolvement and different characteristics to those of the initial tools.

Despite the scarcity of information to answer the review question, our study carried out a comprehensive literature search to identify all relevant studies, including several sources of unpublished literature such as pre-print repositories. Our assessment also includes a rigorous assessment of potential sources of bias, a formal statistical analysis of results and a final assessment

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

of the certainty of the evidence under a well-known system (GRADE). We applied selected methods associated with rapid reviews to streamline the review process, such as the involvement of stakeholders in the development of the review, a non-independent verification of data selection and extraction, and parallelisation of tasks (that is, to conduct selected activities simultaneously instead of consecutively) [29]. We avoided the use of methods that potentially might affect the quality of the review process, such as those related to limiting the search strategies, the omission of quality assessment of the collected evidence and the narrative synthesis of results [29, 30].

Due to the permanent involvement of clinicians managing COVID 19 patients in the development of this review, we were able to define a review question that responds to a clinical inquiry relevant to current clinical practice [31-33]. In fact, the number of cases misdiagnosed as not having the target condition is a critical figure due to the severe consequences of not treatment of missing patients. This estimation also can help in the estimation of additional resources in the current clinical practices to confirm a suspicious case.

Implications for practice

Our findings reinforce the need for repeated testing in patients with suspicion of being infected, due to either clinical or epidemiological reasons, given that up to 29% of patients may have an initial negative RT-PCR (certainty of evidence: very low). The collected evidence has several limitations in terms of risk of bias and applicability; in addition, lack of reporting of several key factors remains a significant constraint for analysis of collected data. A false negative result during the recovering phase could have important implications for isolation and risk of transmission, although this risk is reduced by the documentation of at least two negative samples before the discharge. A consequent positive result could also be erroneously considered as reinfection. An update of this review when new studies would be available is warranted.

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Implications for research

Due to the multiple difficulties associated with the lack of reporting of included studies, and due to the high probability of new studies being published in the short-term, we provided some recommendations for future studies candidates to be included in an update of this review:

- Inclusion of a series of consecutive patients instead of selected groups, to avoid spectrum bias.
- Inclusion of a series of consecutive patients instead of selected groups, to avoid spectrum bias
- If samples /specimens are analysed, reporting of information by patient
- Description of RT-PCR scheme in use, including target genes under assessment and positivity criteria
- Description of pre-analytical steps (conservation of samples, time until being sent to the laboratory, training of personal)
- Clear reporting of the time since the onset of symptoms, especially for those patients with clinical findings at admission
- Reporting of the number of additional RT-PCR assays performed
- Details about the application of the reference standard, including the time of administration after the index test (initial RT-PCR)
- If possible, database sharing could allow re-analyses by independent researchers, including individual-patient data (IPD)-meta-analysis and increasing thus the confidence on the new evidence
- Add serological samples to a cohort of individuals with compatible symptoms and negative PCR to warrant an independent verification of infection.

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DECLARATIONS

Contributors: IAR, DBG, DS and JZ conceived the study. IAR, DBG, DSR, RDC, JAPM and JZ designed the study. IAR, DBG, DSR, PZA screened titles and abstracts for inclusion. IAR, DBG, DSR, PZA, ,AR and JZ extracted and analysed data. RDC, JAPM, AC, OS and NL assisted in the interpretation from a clinical viewpoint. IAR, DBG, DSR and JZ wrote the first draft, which all authors revised for critical content. All authors approved the final manuscript. IAR and JZ are the guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: Ingrid Arevalo-Rodriguez is funded by the Instituto de Salud Carlos III through the "Acción Estrategica en Salud 2013-2016 / Contratos Sara Borrell convocatoria 2017/CD17/00219" (Co-funded by European Social Fund 2014-2020, "Investing in your future").

Competing interests: All authors declare: no support from any organisation for the submitted work; no competing interests with regards to the submitted work.

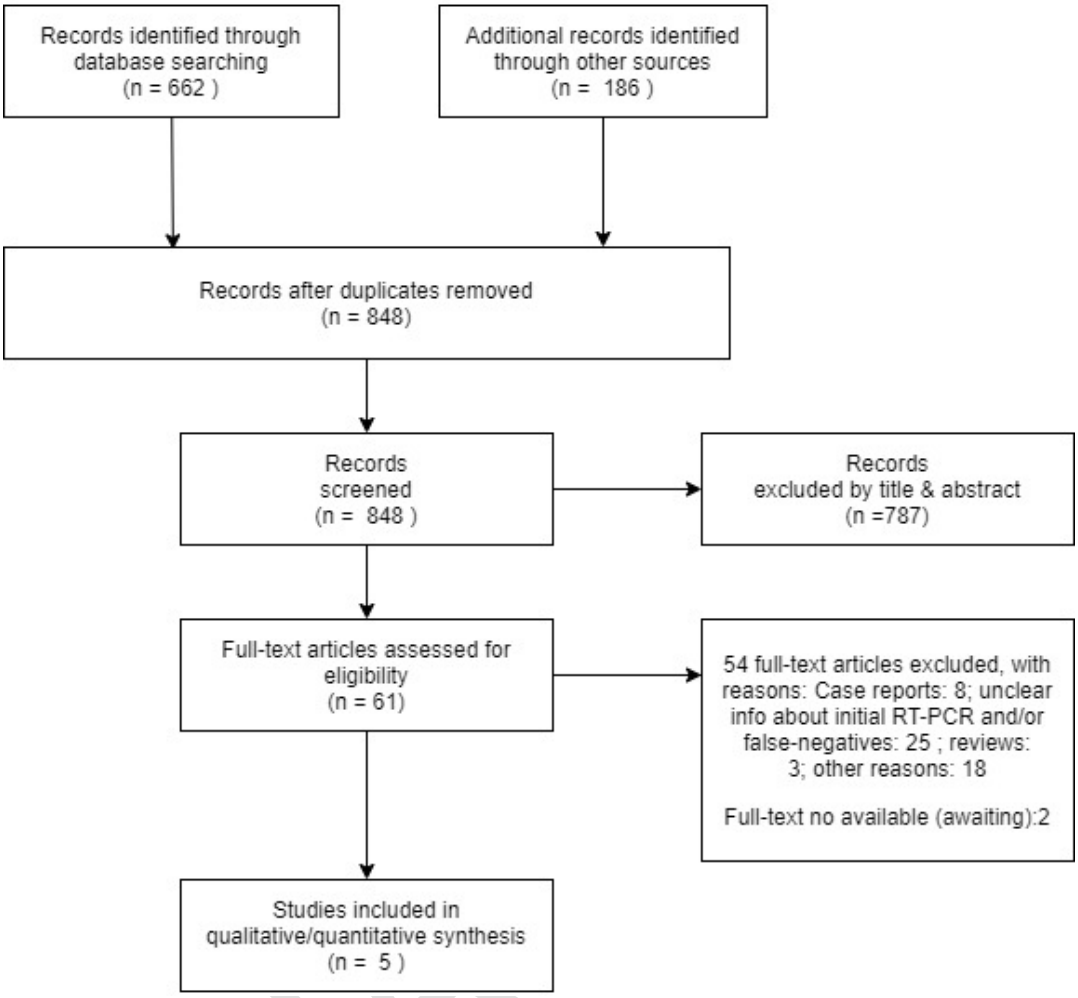
Ethical approval: Not required.

Data sharing: The study protocol is available online at https://tinyurl.com/vvbgqya. Most included studies are publically available. Additional data are available upon reasonable request.

FINAL DRAFT

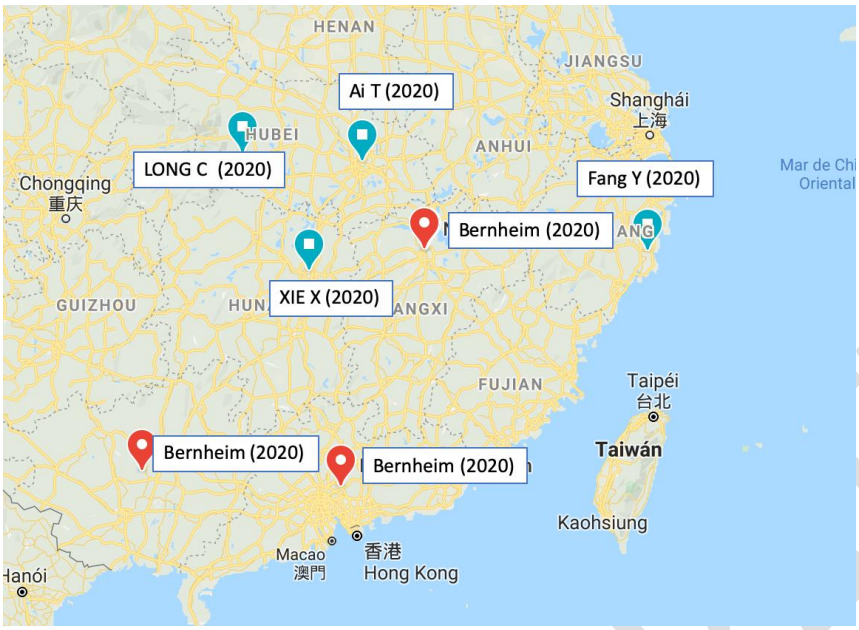
medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 1. PRISMA flow diagram



medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

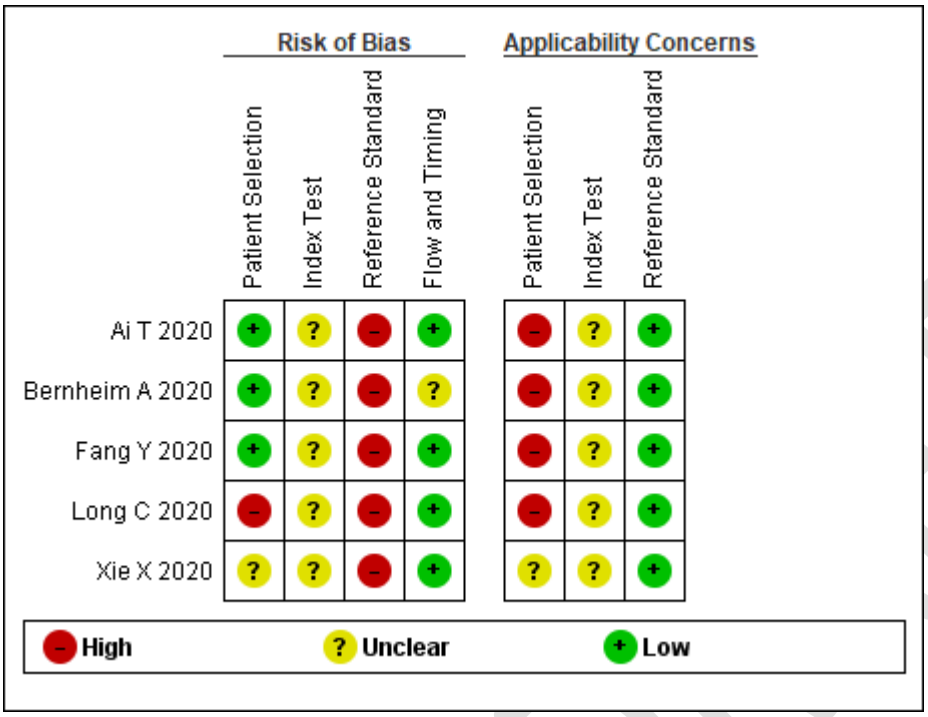
Figure 2. Setting- included studies



FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 3. Methodological quality- QUADAS-II tool

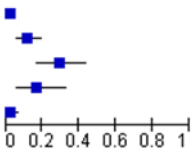


FINAL DRAFT

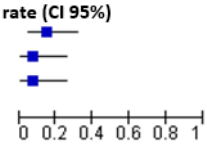
medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 4. Forest plot included studies

Study ID	Time from disease onset	FN	TP	False-negative rate (CI 95%)
Ai T 2020	Not reported	15	586	0.02 (0.01 to 0.04)
Bernheim A 2020	Ranged from to 12 days	12	90	0.12 (0.06 to 0.20)
Fang Y 2020	Mean 3 ± 3 days	15	36	0.29 (0.17 to 0.44)
Long C 2020	Mean 2,6 ±1,7 days	6	30	0.17 (0.06 to 0.33)
Xie X 2020	Not reported	5	162	0.03 (0.01 to 0.07)



Subgroups	Time from disease onset	FN	TP	False-negative rate (CI 95%)
Bernheim A 2020a	≤ 2 days	5	28	0.15 (0.05 to 0.32)
Bernheim A 2020b	3 to 5 days	2	23	0.08 (0.01 to 0.26)
Bernheim A 2020c	6 to 12 days	2	23	0.08 (0.01 to 0.26)



Note: Bernheim A 2020 reported time since the symptom onset to CT scan; Long C 2020 reported duration of fever.

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 5. Certainty of the evidence (GRADE assessment)

Proportion: ranged from 0.02 to 0.29	**Effect per 100 patient tested**			**Number of participants (studies)**	**Certainty of the evidence (GRADE)**
	Prevalence 30% Typically seen in	Prevalence 50% Typically seen in	Prevalence 80% Typically seen in		
False-negatives (patients incorrectly classified as not having COVID-19	1 to 9	1 to 14	2 to 23	957 (5 studies)	⊕OOO VERY LOW[1,2,3]

Notes= 1. Evidence downgraded one level due to risk of bias issues: index test as a part of the reference standard, unclear positivity thresholds for index test; 2. Evidence downgraded one level due to indirectness: selected population enrolled in studies; 3. Evidence downgraded one level due to inconsistency: Tau-square=1.08; 95% CI 0.27 to 8.28; p<0.001.

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

REFERENCES

1. Perrella A, Carannante N, Berretta M, Rinaldi M, Maturo N, Rinaldi L. Novel Coronavirus 2019 (Sars-CoV2): a global emergency that needs new approaches? European review for medical and pharmacological sciences. 2020;24(4):2162-4. Epub 2020/03/07. doi: 10.26355/eurrev_202002_20396. PubMed PMID: 32141586.
2. Pang J, Wang MX, Ang IYH, Tan SHX, Lewis RF, Chen JI, et al. Potential Rapid Diagnostics, Vaccine and Therapeutics for 2019 Novel Coronavirus (2019-nCoV): A Systematic Review. J Clin Med. 2020;9(3). doi: 10.3390/jcm9030623. PubMed PMID: 32110875.
3. Meo SA, Alhowikan AM, Al-Khlaiwi T, Meo IM, Halepoto DM, Iqbal M, et al. Novel coronavirus 2019-nCoV: prevalence, biological and clinical characteristics comparison with SARS-CoV and MERS-CoV. European review for medical and pharmacological sciences. 2020;24(4):2012-9. Epub 2020/03/07. doi: 10.26355/eurrev_202002_20379. PubMed PMID: 32141570.
4. World Health Organization. Coronavirus disease 2019 (COVID-19): Situation Report – 80. Geneve, Switzerland: 2020.
5. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497-506. doi: 10.1016/S0140-6736(20)30183-5. PubMed PMID: 31986264.
6. Chen Q, Quan B, Li X, Gao G, Zheng W, Zhang J, et al. A report of clinical diagnosis and treatment of nine cases of coronavirus disease 2019. J Med Virol. 2020. doi: 10.1002/jmv.25755. PubMed PMID: 32162699.
7. Paules CI, Marston HD, Fauci AS. Coronavirus Infections—More Than Just the Common Cold. JAMA. 2020;323(8):707-8. doi: 10.1001/jama.2020.0757.
8. Young BE, Ong SWX, Kalimuddin S, Low JG, Tan SY, Loh J, et al. Epidemiologic Features and Clinical Course of Patients Infected With SARS-CoV-2 in Singapore. JAMA. 2020. doi: 10.1001/jama.2020.3204.
9. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious Diseases. doi: 10.1016/S1473-3099(20)30243-7.
10. Azman AS, Luquero FJ. From China: hope and lessons for COVID-19 control. The Lancet Infectious Diseases. doi: 10.1016/S1473-3099(20)30264-4.
11. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. N Engl J Med. 2020;382(12):1177-9. doi: 10.1056/NEJMc2001737. PubMed PMID: 32074444.
12. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2020;25(3). Epub 2020/01/30. doi: 10.2807/1560-7917.es.2020.25.3.2000045. PubMed PMID: 31992387; PubMed Central PMCID: PMCPMC6988269.
13. Chu DKW, Pan Y, Cheng SMS, Hui KPY, Krishnan P, Liu Y, et al. Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. Clinical chemistry. 2020;66(4):549-55. Epub 2020/02/08. doi: 10.1093/clinchem/hvaa029. PubMed PMID: 32031583; PubMed Central PMCID: PMCPMC7108203.
14. Lippi G, Simundic AM, Plebani M. Potential preanalytical and analytical vulnerabilities in the laboratory diagnosis of coronavirus disease 2019 (COVID-19). Clinical chemistry and laboratory medicine. 2020. Epub 2020/03/17. doi: 10.1515/cclm-2020-0285. PubMed PMID: 32172228.

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

15. Organization WH. Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases. Geneve, Switzerland: 2020 2 March 2020. Report No.: Contract No.: WHO/COVID-19/laboratory/2020.4.

16. Li D, Wang D, Dong J, Wang N, Huang H, Xu H, et al. False-Negative Results of Real-Time Reverse-Transcriptase Polymerase Chain Reaction for Severe Acute Respiratory Syndrome Coronavirus 2: Role of Deep-Learning-Based CT Diagnosis and Insights from Two Cases. Korean journal of radiology. 2020;21(4):505-8. Epub 2020/03/17. doi: 10.3348/kjr.2020.0146. PubMed PMID: 32174053; PubMed Central PMCID: PMCPMC7082661.

17. Gostin LO. Public Health Emergency Preparedness: Globalizing Risk, Localizing Threats. JAMA. 2018;320(17):1743-4. doi: 10.1001/jama.2018.16491.

18. Lin C, Ye R, Xia YL. A meta-analysis to evaluate the effectiveness of real-time PCR for diagnosing novel coronavirus infections. Genetics and molecular research : GMR. 2015;14(4):15634-41. Epub 2015/12/05. doi: 10.4238/2015.December.1.15. PubMed PMID: 26634531.

19. Sharfstein JM, Becker SJ, Mello MM. Diagnostic Testing for the Novel Coronavirus. JAMA. 2020. doi: 10.1001/jama.2020.3864.

20. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement. JAMA. 2018;319(4):338-96.

21. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine. 2011;155(8):529-36. Epub 2011/10/19. doi: 10.7326/0003-4819-155-8-201110180-00009. PubMed PMID: 22007046.

22. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias and indirectness in rating the certainty across a body of evidence for test accuracy. Journal of Clinical Epidemiology. doi: 10.1016/j.jclinepi.2019.12.020.

23. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Inconsistency, Imprecision, publication bias and other domains for rating the certainty of evidence for test accuracy and presenting it in evidence profiles and summary of findings tables. Journal of Clinical Epidemiology. doi: 10.1016/j.jclinepi.2019.12.021.

24. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology. 2020:200642. doi: 10.1148/radiol.2020200642. PubMed PMID: 32101510.

25. Bernheim A, Mei X, Huang M, Yang Y, Fayad ZA, Zhang N, et al. Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection. Radiology. 2020:200463. doi: 10.1148/radiol.2020200463. PubMed PMID: 32077789.

26. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. Radiology. 2020:200432. doi: 10.1148/radiol.2020200432. PubMed PMID: 32073353.

27. Long C, Xu H, Shen Q, Zhang X, Fan B, Wang C, et al. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? European journal of radiology. 2020;126:108961. Epub 2020/04/02. doi: 10.1016/j.ejrad.2020.108961. PubMed PMID: 32229322; PubMed Central PMCID: PMCPMC7102545.

28. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for Typical 2019-nCoV Pneumonia: Relationship to Negative RT-PCR Testing. Radiology. 2020:200343. doi: 10.1148/radiol.2020200343. PubMed PMID: 32049601.

29. Arevalo-Rodriguez I, Moreno-Nunez P, Nussbaumer-Streit B, Steingart K, Gonzalez Peña L, Buitrago-Garcia D, et al. Rapid reviews of medical tests used many similar methods to systematic

FINAL DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20066787; this version posted April 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

reviews but key items were rarely reported: a scoping review. Journal of Clinical Epidemiology. 2019;116:98-105.

30. Tricco AC, Antony J, Zarin W, Strifler L, Ghassemi M, Ivory J, et al. A scoping review of rapid review methods. BMC medicine. 2015;13:224. Epub 2015/09/18. doi: 10.1186/s12916-015-0465-6. PubMed PMID: 26377409; PubMed Central PMCID: PMCPMC4574114.

31. Wilson MG, Lavis JN, Gauvin FP. Developing a rapid-response program for health system decision-makers in Canada: findings from an issue brief and stakeholder dialogue. Systematic reviews. 2015;4:25. Epub 2015/04/16. doi: 10.1186/s13643-015-0009-3. PubMed PMID: 25875495; PubMed Central PMCID: PMCPMC4373100.

32. Moore G, Redman S, Rudge S, Haynes A. Do policy-makers find commissioned rapid reviews useful? Health Res Policy Syst. 2018;16(1):17. Epub 2018/02/28. doi: 10.1186/s12961-018-0293-1. PubMed PMID: 29482643; PubMed Central PMCID: PMCPMC5828139.

33. Hartling L, Guise JM, Hempel S, Featherstone R, Mitchell MD, Motu'apuaka ML, et al. Fit for purpose: perspectives on rapid reviews from end-user interviews. Systematic reviews. 2017;6(1):32. doi: 10.1186/s13643-017-0425-7. PubMed PMID: 28212677; PubMed Central PMCID: PMCPMC5316162.

FINAL DRAFT