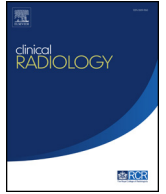**ELSEVIER**

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

Technical Report

A statistical framework to estimate diagnostic test performance for COVID-19

R. Symons [a,b,]*, K. Beath [c], A. Dangis [a], S. Lefever [d], A. Smismans [d], Y. De Bruecker [a], J. Frans [d]

[a] Department of Radiology, Imelda Hospital Bonheiden, Belgium
[b] Department of Imaging and Pathology, University Hospital Leuven, Belgium
[c] Department of Statistics, Macquarie University, Sydney, Australia
[d] Department of Microbiology, Imelda Hospital Bonheiden, Belgium

Introduction

Diagnostic test performance is traditionally evaluated by comparison to a perfect reference standard test assumed to have 100% sensitivity and specificity. Reverse transcriptase polymerase chain reaction (RT-PCR) detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA in throat swabs is currently the most common used diagnostic test for coronavirus disease 2019 (COVID-19). Chest computed tomography (CT) has been suggested by some as an alternative test for initial diagnosis[1]; However, both RT-PCR and CT are imperfect tests for COVID-19 diagnosis with reports suggesting RT-PCR sensitivity to be only 70%[1]. As no perfect tests are available for COVID-19 diagnosis, traditional diagnostic accuracy assessments using either RT-PCR or chest CT as the reference standard are inherently biased. Latent class analysis can be used to estimate the accuracy of diagnostic tests in the absence of a reference standard.[2] The aim of this study was to use latent class analysis to estimate the diagnostic performance of RT-PCR and CT for the diagnosis of COVID-19 and to provide researchers with access to a statistical framework to calculate the diagnostic accuracy of different COVID-19 diagnostic tests for their own local institution.

Materials and methods

From 15 March to 1 June 2020, 1,201 consecutive symptomatic patients with possible COVID-19 infection who underwent both RT-PCR and chest CT at presentation at Imelda Hospital Bonheiden were included retrospectively. This study was approved by the institutional review board, and informed consent was waived. Two PCR methods were used to detect SARS-CoV-2 in nasopharyngeal swabs (eSwab, Copan Diagnostics), both using the E-gene as target: ARIES system (Luminex, Austin, USA) and Rotorgene Q (Qiagen, Hilden, Germany). No cross reactivity for other human coronaviruses, influenza, or respiratory syncytial virus has been shown. Chest CT was scored as suggestive for or inconsistent with COVID-19 infection

* Guarantor and correspondent: R. Symons, Department of Radiology, Imelda Hospital, Imeldalaan 9, 2820, Bonheiden, Belgium. Tel.: +32 15 50 49 83.
  E-mail address: rolf.symons@imelda.be (R. Symons).

Table 1
Estimated performance for the diagnosis of coronavirus disease 2019 (COVID-19) for Reverse transcriptase polymerase chain reaction (RT-PCR), computed tomography (CT), clinical, and laboratory parameters.

	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
CT	82.5 (77.9,86.9)	98.2 (96.6,100.0)	95.4 (90.6,100.0)	92.8 (90.7,94.8)
RT-PCR	85.3 (79.6,91.0)	100.0	100.0	94.0 (91.3,96.4)
Fever	59.4 (53.7,65.2)	64.6 (59.7,69.7)	42.3 (37.0,47.7)	78.4 (74.7,82.3)
Dyspnoea	54.8 (48.9,60.4)	57.0 (51.4,61.6)	35.6 (30.9,40.2)	74.2 (69.6,78.5)
Chest pain	15.0 (10.9,19.7)	71.9 (67.5,76.9)	18.9 (13.5,24.9)	65.9 (62.6,69.6)
Cough	60.0 (54.1,65.5)	52.6 (47.8,57.7)	35.6 (31.0,40.2)	75.1 (70.7,79.1)
Diarrhoea	17.3 (13.2,21.9)	84.4 (80.5,88.1)	32.8 (24.4,41.7)	70.0 (67.0,73.3)
Myalgia	15.0 (11.0,19.6)	92.3 (89.6,94.9)	45.8 (34.2,58.2)	71.3 (68.0,74.6)
Anorexia	44.3 (38.4,50.3)	68.8 (64.2,73.4)	38.2 (32.5,44.3)	73.9 (70.2,77.5)
Syncope	3.7 (1.7,6.2)	97.8 (96.2,99.2)	42.0 (21.1,68.7)	69.9 (66.7,73.1)
Anosmia	4.7 (2.5,7.4)	99.5 (98.6,100.0)	81.4 (56.0,100.0)	70.5 (67.5,73.7)
Saturation ≤92%	20.2 (15.8,25.2)	94.9 (92.4,97.1)	63.7 (50.8,76.8)	73.2 (70.2,76.3)
D-dimer >500	84.8 (78.9,89.9)	39.2 (32.8,46.8)	37.7 (33.4,42.8)	85.4 (80.1,90.6)
CRP ≥5 mg/l	87.7 (83.5,91.6)	50.6 (45.5,55.4)	43.7 (38.9,47.9)	90.3 (87.0,93.6)
Lymphocytes <1,000×10^6/ml	49.5 (43.4,55.8)	77.5 (73.0,81.7)	48.9 (41.8,56.3)	77.8 (74.5,81.4)
LDH ≥214 U/l	81.6 (76.1,86.2)	56.9 (50.9,63.0)	45.2 (40.2,50.9)	87.7 (83.8,91.1)
Total bilirubin ≥0.7 mg/dl	25.4 (19.5,31.6)	73.1 (67.5,78.0)	29.2 (21.8,36.5)	69.1 (65.3,72.8)
ALT ≥31 U/l	39.2 (33.8,45.8)	78.9 (74.1,82.9)	44.7 (37.8,51.8)	74.8 (71.4,78.5)
GFR ≤90 ml/min/1.73m^2	75.5 (70.1,80.2)	42.1 (36.8,47.3)	36.2 (31.9,40.8)	79.7 (74.8,84.2)
Creatine kinase ≥67 U/l	26.2 (21.0,31.6)	85.0 (80.7,88.6)	43.2 (34.2,52.5)	72.6 (69.0,75.8)
Troponins ≥14 ng/l	29.3 (21.2,38.2)	67.5 (60.2,73.7)	28.2 (20.5,36.2)	68.7 (64.0,73.1)

CRP, C-reactive protein; LDH, lactate dehydrogenase; ALT, alanine transferase; GFR, glomerular filtration rate.
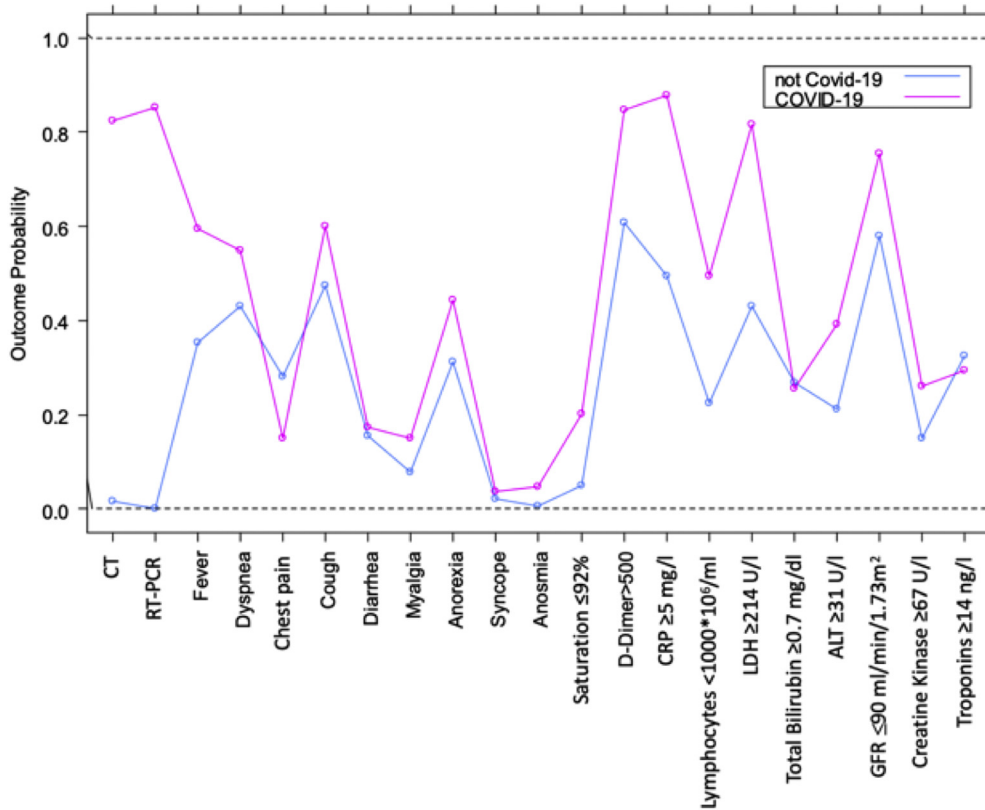

Figure 1 Outcome probabilities based on the presence of the different diagnostic parameters. The pink curve shows the percentage of COVID-19 positive cases with positive findings on the different diagnostic parameters (e.g., findings suggestive for COVID-19 on CT, RT-PCR, fever, etc.). The blue curve shows the percentage of COVID-19 negative cases with positive findings on the different diagnostic parameters (e.g., findings suggestive for COVID-19 on CT, RT-PCR, fever, etc.).

based on the presence of typical findings such as multiple ground-glass opacities, bilateral/multifocal involvement, peripheral distribution, crazy paving, consolidation, and reversed halo sign.[3] RT-PCR results were not available at the moment of CT evaluation. The data were analysed using a random effects latent class model with RT-PCR specificity constrained to 100%.[4] Analysis was performed using Latent Gold with Syntax Module and R (Foundation for Statistical Computing). Confidence intervals (CI) were calculated using a non-parametric bootstrap.

Results

RT-PCR demonstrated excellent sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV; 85%, 100%, 100%, and 94%, respectively) for COVID-19 diagnosis, similar to chest CT (83%, 98%, 95%, and 93%, respectively). Both RT-PCR and CT significantly outperformed other clinical and laboratory parameters for COVID-19 diagnosis (Table 1, Fig 1). RMarkdown code is available online with this article to allow researchers to calculate the diagnostic performance of different COVID-19 tests for their own institution (See Electronic Supplementary Material Appendix S1).

Discussion

Most estimates of accuracy for RT-PCR and CT are based on the false assumption that one or the other could be used as a standard reference. The use of an imperfect reference test leads to an underestimation of the diagnostic accuracy of other diagnostic measurements. Latent class analysis indicated both RT-PCR and chest CT are highly sensitive and specific for COVID-19 diagnosis. At Imelda Hospital Bonheiden, RT-PCR has the edge over CT as it has 100% specificity, rapid availability, and does not expose the patient to ionising radiation. CT could be used as a problem solver for patients with a high clinical suspicion for COVID-19 but negative RT-PCR, and in institutions where RT-PCR is not readily available. Additionally, CT could be used as an adjunct after COVID-19 diagnosis as the extent of lung involvement has prognostic value.[5,6]

Limitations of this study include its retrospective and single-centre design.

A correct interpretation of the reliability of COVID-19 test results is essential for optimal protection of patients, healthcare workers, and the general population. We provide researchers with a tool to calculate the diagnostic performance of COVID-19 test results for their own clinical practice.

Conflict of interest

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crad.2020.10.004.

References

1. Ai T, Yang Z, Hou H, *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.
2. Black MA, Craig BA. Estimating disease prevalence in the absence of a gold standard. *Stat Med* 2002;**21**(18):2653—69.
3. Dangis A, Gieraerts C, Bruecker Y De, *et al*. Accuracy and reproducibility of low-dose submillisievert chest CT for the diagnosis of COVID-19. *Radiol Cardiothorac Imaging* 2020;**2**(2):e200196.
4. Qu Y, Tan M, Kutner MH. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics* 1996:797—810.
5. Colombi D, Bodini FC, Petrini M, *et al*. Well-aerated lung on admitting chest CT to predict adverse outcome in COVID-19 pneumonia. *Radiology* 2020:201433.
6. Gieraerts C, Dangis A, Janssen L, *et al*. Prognostic value and reproducibility of AI-assisted analysis of lung involvement in COVID-19 on low-dose submillisievert chest CT: sample size implications for clinical trials. *Radiol Cardiothorac Imaging* 2020;**2**(5):e200441.