

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Chest X-Ray Has Poor Sensitivity and Prognostic Significance in COVID-19: A Propensity Matched Database Study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-042946
Article Type:	Original research
Date Submitted by the Author:	20-Jul-2020
Complete List of Authors:	Borakati, Aditya; Royal Free Hospital, Emergency Department; University College London, Division of Surgery and Interventional Science Perera, Adrian; Royal Free Hospital, Emergency Department Johnson, James; Royal Free Hospital, Emergency Department Sood, Tara; Royal Free Hospital, Emergency Department
Keywords:	COVID-19, Chest imaging < RADIOLOGY & IMAGING, ACCIDENT & EMERGENCY MEDICINE, GENERAL MEDICINE (see Internal Medicine), Diagnostic radiology < RADIOLOGY & IMAGING

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1
2
3 **Chest X-Ray Has Poor Sensitivity and Prognostic Significance in COVID-19: A**
4 **Propensity Matched Database Study**
5
6

7
8 **Authors:** Aditya Borakati (0000-0003-0457-4944), Adrian Perera (0000-0002-1753-
9 2586), James Johnson (0000-0001-7796-659X), Tara Sood (0000-0003-1620-916X)

10
11
12 Emergency Department, Royal Free Hospital, Pond Street, London, UK NW3 2QG

13
14
15 Aditya Borakati, Academic Foundation Doctor, Division of Surgery and Interventional
16 Science, Royal Free Campus, University College London

17
18
19 Adrian Perera, Emergency Medicine and Intensive Care Medicine Specialist Registrar,
20 Emergency Department, Royal Free Hospital

21
22
23 James Johnson, Foundation Doctor, Emergency Department, Royal Free Hospital

24
25
26 Tara Sood, Consultant in Emergency Medicine, Emergency Department, Royal Free
27 Hospital
28
29

30
31 **Correspondence:**

32
33 Dr Aditya Borakati

34
35 Academic Foundation Doctor

36
37
38 Emergency Department, Royal Free Hospital, London, UK NW3 2QG

39
40
41 a.borakati@doctors.org.uk
42

43
44 **Author contribution (CRediT) statement:**

45
46 **Aditya Borakati:** Conceptualization, Methodology, Validation, Formal Analysis,
47 Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project
48 Administration
49
50

51
52 **Adrian Perera:** Conceptualization, Methodology, Investigation, Writing- Review &
53 Editing, Supervision, Project Administration
54
55
56

1
2
3 **James Johnson:** Investigation
4

5
6 **Tara Sood:** Conceptualization, Methodology, Writing – Review & Editing, Supervision,
7
8 Project Administration
9

10 Aditya Borakati is the overall guarantor of this work.
11

12
13 **Word count:** 4236
14
15

16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Abstract

Objectives: To identify the diagnostic accuracy of common imaging modalities, chest X-ray (CXR) and computed tomography (CT) for diagnosis of COVID-19 in the general emergency population in the UK and to find the association between imaging features and outcomes in these patients.

Design: Retrospective analysis of electronic patient records

Setting: Tertiary academic health science centre and designated centre for high consequence infectious diseases in London, UK.

Participants: 1,198 patients who attended the emergency department with paired RT-PCR swabs for SARS-CoV 2 and CXR between 16th March and 16th April 2020

Main outcome measures: Sensitivity and specificity of CXR and CT for diagnosis of COVID-19 using the British Society of Thoracic Imaging reporting templates. Reference standard was any reverse transcriptase polymerase chain reaction (RT-PCR) positive naso-oropharyngeal swab within 30 days of attendance. Odds ratios of CXR in association with vital signs, laboratory values and 30-day outcomes were calculated.

Results: Sensitivity and specificity of CXR for COVID-19 diagnosis were 0.56 (95% CI 0.51-0.60) and 0.60 (95% CI 0.54-0.65), respectively. For CT scans these were 0.85 (95% CI 0.79-0.90) and 0.50 (95% CI 0.41-0.60), respectively. This gave a statistically significant mean increase in sensitivity with CT compared with CXR, of 29% (95% CI 19%-38%, $p < 0.0001$). Specificity was not significantly different between the two modalities.

Chest X-ray findings were not statistically significantly or clinical meaningfully associated with vital signs, laboratory parameters or 30-day outcomes.

Conclusions: Computed tomography has substantially improved diagnostic performance over CXR in COVID-19. CT should be strongly considered in the initial assessment for

1
2
3
4 suspected COVID-19. This gives potential for increased sensitivity and considerably
5 faster turnaround time, where capacity allows and balanced against excess radiation
6 exposure risk.
7
8

9
10 **Key words:** X-Rays, Computed Tomography, COVID-19, severe acute respiratory
11 syndrome coronavirus 2, Emergency Medicine, Diagnostic Imaging
12
13

14 **Strengths and limitations**

15
16
17 -Large, appropriately powered, study population consisting of all patients attending the
18 emergency department rather than those solely with confirmed COVID-19; this allowed
19 assessment of specificity for the imaging modalities and applicability to the general
20 population who may attend medical personnel with other complaints, but have
21 underlying SARS-CoV 2 infection
22
23
24
25

26
27 -Comprehensive statistical analyses were conducted to address confounding in
28 reporting of X-rays including propensity score matching and logistic regression to give a
29 'doubly robust' model
30
31
32

33 -Low amount of missing data and for secondary covariates only; multiple imputation was
34 performed with a good fit, however, observed data would be preferable to imputed data
35
36

37 -Single centre, retrospective study; potential for inter-reporter and inter-centre variability
38 in reporting
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 **Statistical review:** The statistical methods in this manuscript and associated code have
5 been reviewed by Dr Federico Ricciardi of the Department of Statistical Science at
6 University College London and confirmed as robust and accurate.
7
8

9
10 **Ethical approval:** This study was registered with the local institutional review board as a
11 service evaluation using anonymised data only. No formal ethics committee review was
12 required.
13
14

15
16 **Declarations of Interests:** The authors have no relevant conflicts of interest to declare.
17 All authors have completed the [Unified Competing Interest form](#) (available on request
18 from the corresponding author) and declare: no support from any organisation for the
19 submitted work; no financial relationships with any organisations that might have an
20 interest in the submitted work in the previous three years, no other relationships or
21 activities that could appear to have influenced the submitted work.
22
23
24
25
26

27
28 **Transparency declaration:** The lead author (AB) affirms that this manuscript is an
29 honest, accurate, and transparent account of the study being reported; that no important
30 aspects of the study have been omitted; and that any discrepancies from the study as
31 planned (and, if relevant, registered) have been explained.
32
33
34
35

36 **Funding:** No funding was received for this study.
37
38

39 **Copyright:** The Corresponding Author has the right to grant on behalf of all authors and
40 does grant on behalf of all authors, a worldwide licence to the Publishers and its
41 licensees in perpetuity, in all forms, formats and media (whether known now or created
42 in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii)
43 translate the Contribution into other languages, create adaptations, reprints, include
44 within collections and create summaries, extracts and/or, abstracts of the Contribution,
45 iii) create any other derivative work(s) based on the Contribution, iv) to exploit all
46 subsidiary rights in the Contribution, v) the inclusion of electronic links from the
47
48
49
50
51
52
53
54
55
56
57

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above

For peer review only

Introduction

SARS-CoV 2 and its resulting disease, COVID-19, have propagated exponentially worldwide, with over 10 million cases in 188 countries at the time of writing [1,2].

The gold standard for diagnosis of the virus is the detection of viral RNA through reverse transcriptase polymerase chain reaction (RT-PCR) of respiratory tract samples. However, this method has several limitations including: (1) low sensitivity at 59-71% [3,4], (2) relatively slow turnaround times ranging from a few hours to several days [5], (3) high expense and (4) limited capacity for testing in many countries.

Computed tomography (CT) has been shown to be more sensitive than RT-PCR for diagnosis of COVID-19 [3,4], while being significantly faster and cheaper. This comes with a large radiation dose and capacity is still lacking in many countries.

Plain film chest X-ray (CXR) is ubiquitous worldwide, with a 30-70x lower dose of radiation[6] and is commonly performed as an initial investigation in COVID-19.

Studies have so far only evaluated imaging in those with confirmed infection, it is therefore, not possible to calculate the specificity of these modalities. In the context of the global pandemic, infection may be widespread in the community, often with subclinical infection [7,8]. A reliable and rapid method to detect infection in the general population, who may present to medical personnel with other complaints, is needed.

Despite its extensive use, the specificity and sensitivity of CXR in the general emergency population for diagnosis of COVID-19 is unknown, nor how imaging features correlate with severity.

This study evaluated the performance of CXR in diagnosing COVID-19 in the emergency department (ED) of a tertiary care hospital.

Methods

This study was conducted at the Royal Free Hospital, London, UK, an academic health science centre and nationally designated centre for High Consequence Infectious Diseases [9].

All individuals attending the emergency department who had paired posterior-anterior chest radiographs and RT-PCR nasopharyngeal swabs for COVID-19 at the time of initial attendance between 16th March 2020 and 16th April 2020 were included.

All chest radiographs were reported by a Consultant Radiologist and rated on an ordinal scale for probability of COVID-19: Alternative pathology identified, not COVID-19; Clear chest, unlikely COVID; Indeterminate findings for COVID-19; Classical findings of COVID-19, based on the British Society of Thoracic Imaging's (BSTI) reporting templates (table 1) [10]. These were reported prior to RT-PCR results being available.

RT-PCR of swabs were performed in laboratories either at our centre or at a public health laboratory (PHE Collindale, UK), according to published national standard operating procedures [11]. Subsequent RT-PCR swabs taken within 30 days of initial ED attendance were also included.

CT scans performed within 30 days of attendance were retrieved. These were also reported according to the BSTI template. CT pulmonary angiogram was performed in the ED if the D-dimer was >5000 to exclude pulmonary emboli as per the locally agreed protocol. Subsequent CT chest imaging (whether pulmonary angiogram, contrast or non-contrast) was performed on the basis of clinical suspicion.

Prospectively recorded data was extracted from the Cerner Millennium electronic patient record system (Cerner Corp., Kansas City, MO).

Primary Outcome

The primary outcome is sensitivity and specificity of initial CXR, where it is reported as having classic COVID-19 features in the ED. This is compared with RT-PCR swab as the reference standard for diagnosis of COVID-19.

1
2
3
4 In the event of multiple RT-PCR swabs during one attendance, a single positive swab was taken
5 as an overall positive test during one admission.
6

7 **Secondary Outcomes**

8
9
10 In those patients who also had CT scans of the thorax, the diagnostic accuracy was compared
11 with CXR, with RT-PCR again as the reference standard. Sensitivity and specificity of CXR
12 when X-rays reported as indeterminate or atypical for COVID-19 were classed as positive was
13 also calculated.
14
15

16
17 Chest x-ray findings were correlated with vital signs at attendance and blood results, including:
18 neutrophil counts, D-dimer and C-reactive protein, which have been associated with poor
19 prognosis in COVID-19 [12]. Hazard ratios for clinical outcomes including direct admission to
20 the intensive treatment unit (ITU) from ED and 30-day mortality rates were also calculated for
21 CXR reporting categories.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Statistical Analysis

In the event of missing data, multiple imputation was conducted using a Predictive Mean Matching algorithm, via the MICE R package, as described previously [13]. Briefly, this uses a linear regression model (or logistic regression model for categorical data), to find a random value based on already observed data, to replace missing fields [14]. Variables without missing data fields were not modified. The number of imputed datasets was similar in number to the percentage of missing data as suggested by White and colleagues [15]. Balance diagnostics with density plots are available in supplementary file 1, adequate balance was assessed via visual inspection of imputed distributions with respect to the original dataset.

The propensity for a CXR being reported as positive or negative for COVID-19 was calculated for several plausible covariates that may influence image characteristics such as Age, Gender, Ethnicity, pre-existing morbidities and the respiratory rate of the patient using a generalised linear model [16]. X-ray positive and negative groups were then matched in each imputed dataset using the nearest neighbour algorithm, with a calliper of 0.2 of the propensity score standard deviation, without replacement and in random sequential order to obtain a 1:1 match as described elsewhere [17].

The balance of the match data was assessed quantitatively with mean differences of covariates in each of the X-ray groups pre- and post-matching, with a difference of less than 0.1% considered a good match (supplementary figure 2). Visual inspection of matches was also conducted to ensure balance (supplementary figures 2, 3 and 4).

After matching, outcome data were adjusted for covariates including age, gender, ethnicity and presence of co-morbidities as well as C-reactive protein, D-dimer, troponin and vital signs. This was achieved by generalised linear regression for continuous outcome data, binomial logistic regression for binary categorical outcomes, or ordinal logistic regression in the case of CXR where it is the outcome variable.

These regression models were run on each imputed dataset and outcomes were pooled together across each imputed data set according to Rubin's rules [18] to give an overall estimate.

Diagnostic Accuracy Statistics

1
2
3
4 Chest X-rays reported as classical for COVID-19 as per the BSTI guidelines were considered a
5 positive test in the primary analysis. In a secondary analysis X-rays reported as 'Indeterminate'
6 or 'Atypical' for COVID-19 were also considered positive. All other reports were classified as a
7 negative test. These were compared to nasopharyngeal aspirate RT-PCR results, which were
8 taken as the gold standard for diagnosis of COVID-19. Where more than one swab was taken
9 during the study period (up to 30 days after initial attendance), a single positive result was taken
10 as a positive result for calculation of diagnostic accuracy statistics.
11
12
13
14

15
16 Sensitivity, specificity, predictive values and diagnostic accuracy were calculated using the
17 propensity matched data after imputation and pooled across imputed datasets with 95%
18 confidence intervals. Apparent and true prevalence based on this dataset are also given for
19 interpretation of the predictive values.
20
21
22

23
24 Chest CTs were also reported according to the BSTI guidelines as with X-ray. Diagnostic
25 statistics were calculated on raw, unmatched and non-imputed data (due to a low volume of
26 data for imputation and matching) in the same manner as X-ray. Mean differences and 95%
27 confidence intervals between CT and X-ray for each of the diagnostic statistics are given, with a
28 p-value calculated from the confidence intervals.
29
30
31

32
33 Agreement between the modalities was assessed on the unmatched dataset, in the sample
34 where CT, CXR and RT-PCR were all available using Cohen's (for two group agreement) and
35 Fleiss' Kappa (when all 3 are compared).
36
37

38 39 **Data Presentation**

40
41 Descriptive statistics are given as means and standard deviations for normally distributed data
42 and as medians and interquartile ranges for non-normally distributed data, before and after
43 matching and multiple imputation (for the latter these statistics are pooled across imputations).
44
45

46
47 Association of explanatory variables with SARS-CoV 2 and Chest X-ray findings are given as
48 odds ratios in uni- and multi-variate configurations.
49
50

51
52 Data was considered statistically significant if $p < 0.05$. Given the large number of analyses in
53 this paper, data is separately highlighted if $p < 0.001$ as a secondary threshold to address the
54 potential for false positives with multiple testing.
55
56

1
2
3
4 Analyses were conducted using R 4.0.0 (R Foundation for Statistical Computing, Vienna,
5 Austria) and code for the analyses is given in supplementary file 2.
6

7 **Sample size calculation**

8
9
10 In this study, the lower confidence interval for sensitivity of CXR as reported by Wong et al.[19]
11 (56%) was used as an estimate of likely sensitivity for COVID-19 . A power of 80% at an alpha
12 of 0.05 was used to calculate the sample size for sensitivities and specificities of 56%. This
13 gave an estimated sample size of 165 in each of the COVID-19 negative and positive groups by
14 RT-PCR (total 330).
15
16
17
18

19 **Ethical approval**

20
21 This study was registered with the local institutional review board as a service evaluation using
22 anonymised data only. No formal ethics committee review was required.
23
24

25 **Reporting Guidelines**

26
27 This study is reported according to the STARD guidelines [20] for diagnostic accuracy studies.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Results

1,198 eligible patients with both CXR and RT-PCR were identified in the study period (figure 1). Their characteristics, stratified by positivity for SARS-CoV 2 infection by RT-PCR is summarized in table 2. This showed that those with confirmed SARS-CoV 2 infection were more likely to be male, older (mean age 66.2 vs 62.7), have lower saturations, higher respiratory rates, whilst being more likely to be admitted and die within 30 days. There was a significant association with X-ray images and SARS-CoV 2 at baseline, with 59.6% having classic imaging features of COVID-19 in those with positive swabs versus 39.1% in those with negative swabs. There was 8.6% missing data overall in the dataset when variables with >50% missing data were removed and 15 imputations were performed on these remaining variables only.

After multiple imputation for missing data and pooled propensity score matching for plausible covariates that may affect CXR reporting, there were 430 patients in each of the X-ray positive and X-ray negative groups, for a total of 860 patients. Adequate balance was achieved for relevant covariates with a mean difference of <0.1 between groups (supplementary table 2).

Computed tomography (CT) was performed in 302 patients with paired RT-PCR during the same time period, with a median serial interval of 4.5 days (inter quartile range 0-17) after the initial attendance in ED and of these 30.1% were within one day of attendance.

Diagnostic Accuracy

The pooled sensitivity and specificity of CXR was 0.56 (95% CI 0.51-0.60) and 0.60 (95% CI 0.54-0.65), respectively (table 4). This gave an overall diagnostic accuracy of 0.57 (95% CI 0.54-0.61) for CXR.

In comparison, sensitivity and specificity for CT was 0.85 (95% CI 0.79-0.90) and 0.50 (95% CI 0.41-0.60), respectively. This gave a statistically significant mean increase in sensitivity with CT compared with CXR by 29% (95% CI 19%-38%, $p < 0.0001$). Specificity was not significantly different between the two modalities. Diagnostic accuracy and negative predictive values were also significantly increased with CT at 0.15 and 0.22, respectively, while the negative likelihood ratio was significantly decreased at -0.44. This shows that the post-test odds of being negative for SARS-CoV 2 by RT-PCR with a negative CT is significantly lower.

1
2
3
4 Taking X-rays reported as indeterminate as positive increased the sensitivity of CXR to 0.80
5 (95% CI 0.77-0.84), however reduced specificity to 0.40 (95% CI 0.35-0.46). When CT scans
6 reported as indeterminate are also considered positive the sensitivity of CT increased to 0.93
7 (95% CI 0.89-0.96), whilst mean specificity reduced to 0.37 (95% CI 0.28-0.47), although this
8 was not statistically different from when indeterminate CTs are considered negative. Sensitivity
9 of CT remained significantly higher than CXR (when indeterminates are considered positive for
10 both) by 0.13 (95% CI 0.05-0.19, $p < 0.001$), specificity was not significantly different between the
11 two.
12

13
14
15
16
17
18 When comparing only the unimputed, unmatched subset of data where CT, RT-PCR and CXR
19 were all performed ($n=287$), the agreement between CT and CXR was poor (Cohen's kappa
20 0.406). Agreement between all three modalities was also poor (Fleiss' kappa 0.361).
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Association of CXR with Markers of Severity and Outcomes

Association of covariates with RT-PCR results is shown in table 4 and figure 2. Those who tested positive for SARS-CoV 2 by RT-PCR were significantly more likely to have a classical X-ray (OR 1.79 95% CI 1.25-2.56, $p < 0.002$) as would be expected by the diagnostic accuracy statistics (table 4). When the CXR report is considered as an ordered scale, worsening grades of report were associated more strongly with RT-PCR positivity, with a 1.94 x increase in odds for each grade.

Positive chest X-rays for COVID-19 were significantly associated with lower oxygen saturations (OR 0.94 95% CI 0.92-0.97, $p < 0.001$) and temperatures (2.30 95% CI 1.46-3.63, $p < 0.001$) in the ED following propensity score matching and multivariate regression (table 5 and figure 3).

They also had higher rates of admission to a general ward from the ED (OR 2.30 95% CI 1.46-3.63, $p < 0.001$) but no significant association with 30 day outcomes. There was a statistically significant increase in C-reactive protein with a positive X-ray, however, this is unlikely to be clinically meaningful due to the minimal association (OR 1.00 95% CI 1.00-1.01).

Discussion

This study is the first to report the diagnostic accuracy of CXR and CT in the general emergency population during the COVID-19 pandemic.

We show that CXR has poor sensitivity and specificity for diagnosis of COVID-19, whilst CT has 29% higher sensitivity. Many international radiological guidelines advise against CT scanning for the initial assessment of COVID-19 [21–23] or where there are equivocal CXRs, whilst in other countries CT scanning is performed as a routine first line investigation. Our results suggest that CT should be considered in the initial assessment of COVID-19 and that CXR findings poorly correlate with CT findings in this setting. We also show that indeterminate and non-classical features of COVID-19 significantly increase the sensitivity of these imaging modalities, without a significant decrease in specificity. Further, we demonstrate the limited prognostic value of CXR in COVID-19.

These findings mirror what has previously been reported in the literature on individuals with confirmed COVID-19. Wong et al. [19] showed a sensitivity of 59% for initial X-ray in confirmed COVID-19 infection, similarly initial case series in China also reported a sensitivity of 59.1% [12].

A recent in press article from Italy reported a much higher sensitivity of 89% for CXR in a smaller general emergency population (n=535) without confirmed COVID-19 at attendance [24]. However, this used telephone follow up for clinical symptoms of COVID-19 as a reference standard in individuals with an initial negative RT-PCR swab and appeared to classify any abnormal X-ray as positive, which may inflate this figure. When indeterminate CXRs are counted as positive in this study, the sensitivity would be in line with this Italian data. In the US, a study of patients attending an urgent care centre with confirmed COVID-19, showed a much lower sensitivity at 41.7% for CXR where any abnormality was found on the images [25]. In this study 97/636 reports were re-classified from 'possible pneumonia' to 'normal' on second reading from a radiologist, highlighting the importance of inter-rater agreement and possibly explaining this low estimate.

Computed tomography has been reported in previous studies as being up to 98% sensitive for the diagnosis of COVID-19 in confirmed patients, when RT-PCR is used as the reference standard in confirmed patients [3,4]. These studies used any potential features of COVID-19

1
2
3 (e.g. ground glass opacification, crazy paving) as a positive scan, regardless of spatial
4 distribution or features more characteristic of alternate pathology, unlike the BSTI guidelines
5 used in this study. When we classified indeterminate CTs as positive like these latter studies,
6 our estimates match their sensitivity values.
7
8
9

10
11 Consequently, a much lower specificity of 25% was found with initial RT-PCR in previous
12 literature; however, it is reported that 10 out of 15 (67%) of these negatives subsequently tested
13 positive. This would give an adjusted specificity of 75%, considering subsequent swabs as a
14 reference standard, which combined with the wider CIs in these smaller studies, would bring
15 estimates in line with the specificity in this paper. More recent meta-analyses have placed the
16 pooled sensitivity of CT in populations with confirmed COVID-19 only, at 89.76% (95% CI
17 84.42%-93.84%) [26], in line with the estimates identified here.
18
19
20
21
22

23
24 There is limited coverage in the literature on association of X-ray findings with clinical and
25 laboratory parameters and outcomes in the COVID-19 pandemic. This study demonstrates that
26 classic appearances of COVID-19 were associated with initial lower saturations and lower
27 temperature. Volume opacification of the lung fields were not quantified as a surrogate of
28 severity; however, the use of the BSTI grading templates does this somewhat. When the X-ray
29 report is considered as a graded scale from low likelihood of COVID-19 and severity to high
30 likelihood and severity of disease there was no significant difference in association with vital
31 signs or laboratory parameters compared with when the X-ray report is merely considered as a
32 binary positive and negative outcome for COVID-19.
33
34
35
36
37
38
39

40 Borghesi and colleagues have devised a X-ray grading system, the Brixia score, for severity in
41 admitted patients with confirmed SARS-CoV 2 infection [27]. They further found a significant
42 increase in the severity of CXR by this scoring system in those who were discharged versus
43 those who died [28,29].
44
45
46

47 Here, there were no relevant associations between CXR and laboratory values. This analysis
48 also found no association with positive X-rays and 30 day outcomes after multivariate analyses,
49 unlike Borghese et al. This is also in contrast to Guan et al. who found higher rates of ITU
50 admission and death in those with positive imaging findings. However, these studies analysed
51 only those with confirmed SARS-CoV 2 infection. The divergence observed in this study may be
52
53
54
55
56

1
2
3 due to classifying those with 'Alternate pathology/ Indeterminate' or 'CVXC3/ CVXC2' as per the
4 BSTI templates, negative for COVID-19 in these analyses. Other studies classified X-rays with
5 any abnormality as a positive for COVID-19. These alternate distributions may still be reflective
6 of underlying COVID-19 and we show significantly higher sensitivity for both CT and CXR when
7 these are classed as positive. It may be that correlating indeterminate X-rays (in addition to
8 classical images) with vitals, laboratory markers and 30 day outcomes would yield significant
9 associations. However this may be unlikely, Xu and Zhang et al. found that those with classical
10 bilateral and diffuse involvement in upper and lower lobes had more severe disease than those
11 without [30,31].

12
13
14
15
16
17
18
19
20 There were a total of 70 confirmed pulmonary emboli (PEs) in our dataset out of 114 CT
21 pulmonary angiograms (61.0%, 5.84% of all patients attending) performed in the emergency
22 department. The incidence of venous thromboembolism is reported as ranging from 20-30% in
23 admitted confirmed SARS-CoV 2 positive patients [32]. Although we have not focused on this
24 cohort of patients in this paper for the sake of brevity and simplicity, this high incidence
25 represents a further advantage for CT over CXR.

26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
CT, even with the absence of contrast has been shown to have strong accuracy in the diagnosis
of pulmonary emboli and many imaging features correlate with the presence of pulmonary
emboli. Sensitivities of non-contrast CT for diagnosis of PE have been reported at 96.9% and
specificity at 71.9% [33,34].

We therefore see the advantages of CT scanning in COVID-19 as threefold over other
diagnostic techniques: 1) The rapid turnaround; 2) Increased sensitivity and 3) The possibility to
identify pulmonary emboli in COVID-19, which are a significant burden in this group.

This must be balanced against the excess radiation exposure with CT. Radiation from CT and
its association with carcinogenesis is difficult to quantify and no definitive epidemiological
studies have confirmed excess risk of cancer[35]. Modern CT scanners and software
reconstruction techniques continue to minimise radiation exposure and many ways of shielding
parts of the body from radiation also exist. Nevertheless, the excess risk of lifetime cancer is
estimated at 1 per 5,000 CT examinations[36].

Strengths and Limitations

1
2
3
4 This study is the largest conducted on imaging in the COVID-19 pandemic and one of the only
5 studies conducted in the general population during the pandemic rather than only in confirmed
6 patients. This enables greater applicability to the clinical setting where the diagnosis is
7 uncertain, in addition to being able to calculate specificity, which is not possible in most studies.
8 This study was planned to be powered to detect a sensitivity and specificity of 56% for CXR and
9 greatly exceeded the sample size necessary for this.
10
11
12
13

14 Comprehensive statistical analyses were conducted to account for confounders in both factors
15 influencing reporting of CXR and in factors affecting outcomes. The data was collected from
16 prospectively maintained electronic records; however, the retrieval took place retrospectively
17 with its inherent disadvantages. We were not able to collect data on several relevant covariates
18 such as specific comorbidities or markers of severity such as lymphocytes. Furthermore, there
19 was a significant amount of missing data that required multiple imputation to replace, although
20 the fit of this imputed data was good, actual, observed data would be ideal.
21
22
23
24
25
26

27 Inter-rater reliability of imaging reports was not analysed in this paper and there was the
28 potential for individual radiologists to have greater or lesser accuracy in the diagnosis of COVID-
29 19. The literature has so far suggested a strong degree of agreement between radiologists in
30 reporting of COVID-19 images [28].
31
32
33

34 The single centre nature of this study further limits generalisability and the potential for inter-
35 hospital disagreement in imaging, in addition to inter-rater disagreement.
36
37
38

39 Finally, the median time for patients to receive a CT scan was 4.5 days following initial
40 attendance to ED. Thus, the scans may not have been directly comparable to the initial CXR,
41 both because of the progression of disease and because the SARS-CoV 2 status may have
42 been confirmed at this point, biasing the reporting of these scans.
43
44
45

46 **Future Research**

47

48 Although this study used RT-PCR of nasopharyngeal swabs as a reference standard, newer
49 methods exist for diagnosis of the disease. Serological assays for antibodies against SARS-
50 CoV 2 are increasingly available and may represent a better gold standard in diagnosis for
51 future research [37]. RT-PCR is limited by swabbing technique for nasopharyngeal samples and
52
53
54
55
56
57
58
59
60

1
2
3 the fact that the virus is more avid in the lower respiratory tract [38]. However, many patients
4 may not seroconvert prior to death limiting this test to survivors only.
5
6

7 Point of care lung ultrasound is a new technique for diagnosis of COVID-19 which may mitigate
8 many of the issues noted with the modalities discussed so far. It has no radiation, is fast, cheap
9 and may be able to detect lower respiratory tract disease unlike nasopharyngeal swab.
10
11 However, there is limited evidence beyond small case series on its diagnostic accuracy [39–41].
12
13 Further, like other ultrasound techniques accuracy will likely be operator dependent [42] and
14
15 experience will need to be built up for robust results in evaluating suspected COVID-19.
16
17

18 Finally, much research has been conducted in the use of artificial intelligence techniques to
19 correctly diagnose COVID-19 based on imaging [43–45]. These techniques would obviate
20 capacity limitations in reporting imaging as well as eliminate inter-reporter variability. However,
21
22 as with any supervised machine learning technique, large, generalisable datasets, with correctly
23
24 pre-classified positive and negative cases (which in turn will depend on a truly accurate
25
26 reference standard) are needed [46].
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Conclusion

Chest X-ray has poor sensitivity and specificity in diagnosing COVID-19 in the general population during the pandemic. CT scanning has demonstrated excellent sensitivity and should strongly be considered during the pandemic in the initial assessment of COVID-19. This needs to be balanced against the risk of excess radiation with CT, where capacity allows.

Summary box

What is already known on this topic

- Small observational studies, predominantly in China, have reported on imaging features in COVID-19 after a confirmed RT-PCR swab test
- These studies have shown limited sensitivity for chest X-ray, but excellent sensitivity for CT scans, it is not possible to calculate the specificity of these modalities as they only included patients with confirmed COVID-19, therefore it is not possible to assess their utility in the general population who may or may not have COVID-19
- Literature on this general population attending emergency departments and the accuracy of these imaging techniques is limited
- International guidelines including from the British Society of Thoracic Imaging and American College of Radiology do not recommend the use of CT in initial evaluation of suspected COVID-19, largely due to capacity concerns

What this study adds

- This study shows that Chest x-ray has poor sensitivity and specificity in patients with suspected COVID-19 attending the emergency department, whilst CT has excellent sensitivity and is 29% more sensitive than CXR in our study cohort; there was also poor agreement between CT and CXR findings in COVID-19
- Patients with indeterminate imaging without classical distribution of COVID-19 should still be considered at high risk of having the disease

1
2
3 -Our data suggest that CT should be employed more widely as an initial investigation, where
4 capacity allows and balanced against the risk of excess radiation exposure
5
6
7
8
9

10 **Acknowledgements**

11
12 We would like to thank Scott Wilson from the Royal Free Hospital's clinical practice group
13 analytics department for retrieving the data from the hospital's data warehouse.
14
15

16 We would like to thank Dr Federico Ricciardi of the Department of Statistical Science and
17 PRIMENT Clinical Trials Unit at University College London for reviewing the statistical methods
18 in this study.
19
20
21
22
23

24 **Data availability**

25
26 Anonymised data is available on reasonable request from the corresponding author. Analysis
27 scripts are available from DOI: [10.6084/m9.figshare.12674099](https://doi.org/10.6084/m9.figshare.12674099)
28
29

30 **Declarations of Interest**

31
32 The authors declare no conflicts of interest.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- 1 COVID-19 Map. Johns Hopkins Coronavirus Resour. Cent. <https://coronavirus.jhu.edu/map.html> (accessed 30 Jun 2020).
- 2 Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020;**0**. doi:10/ggnsjk
- 3 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. doi:10/ggmw6p
- 4 Fang Y, Zhang H, Xie J, *et al*. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* 2020;:200432. doi:10/ggnnkj
- 5 Konrad R, Eberle U, Dangel A, *et al*. Rapid establishment of laboratory diagnostics for the novel coronavirus SARS-CoV-2 in Bavaria, Germany, February 2020. *Eurosurveillance* 2020;**25**. doi:10/ggp6bw
- 6 Lin EC. Radiation Risk From Medical Imaging. *Mayo Clin Proc* 2010;**85**:1142–6. doi:10/c445mk
- 7 Mizumoto K, Kagaya K, Zarebski A, *et al*. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance* 2020;**25**:2000180. doi:10/ggn4bd
- 8 Arons MM, Hatfield KM, Reddy SC, *et al*. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. *N Engl J Med* 2020;**382**:2081–90. doi:10/ggszfg
- 9 High consequence infectious diseases (HCID). GOV.UK. <https://www.gov.uk/guidance/high-consequence-infectious-diseases-hcid> (accessed 24 May 2020).
- 10 Desai S. COVID-19 BSTI Reporting templates | The British Society of Thoracic Imaging. Br. Soc. Thorac. Imaging. 2020. <https://www.bsti.org.uk/covid-19-resources/covid-19-bsti-reporting-templates/> (accessed 29 Apr 2020).

- 1
2
3
4 11 NHS England. Guidance and Standard Operating Procedure: COVID-19 virus testing in
5 NHS Laboratories. 2020. [https://www.england.nhs.uk/coronavirus/wp-](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)
6 [content/uploads/sites/52/2020/03/guidance-and-sop-covid-19-virus-testing-in-nhs-](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)
7 [laboratories-v1.pdf](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 24 May 2020).
8
9
- 10
11 12 Guan W, Ni Z, Hu Y, *et al*. Clinical Characteristics of Coronavirus Disease 2019 in China. *N*
12 *Engl J Med* 2020;0:null. doi:10/ggm6dh
13
- 14
15 13 Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. *J Stat Softw*
16 2011;**45**. doi:10/gdqc9c
17
- 18
19 14 Ginkel JR van, Linting M, Rippe RCA, *et al*. Rebutting Existing Misconceptions About
20 Multiple Imputation as a Method for Handling Missing Data. *J Pers Assess* 2020;**102**:297–
21 308. doi:10/gftj5w
22
23
- 24
25 15 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and
26 guidance for practice. *Stat Med* 2011;**30**:377–99. doi:10.1002/sim.4067
27
28
- 29
30 16 He H, McDermott MP. A robust method using propensity score stratification for correcting
31 verification bias for binary tests. *Biostat Oxf Engl* 2012;**13**:32–47. doi:10/c4jzn6
32
- 33
34 17 Ho DE, Imai K, King G, *et al*. MatchIt: Nonparametric Preprocessing for Parametric Causal
35 Inference. *J Stat Softw* 2011;**42**. doi:10/gdwtnq
36
- 37
38 18 Marshall A, Altman DG, Holder RL, *et al*. Combining estimates of interest in prognostic
39 modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res*
40 *Methodol* 2009;**9**:57. doi:10.1186/1471-2288-9-57
41
42
- 43
44 19 Wong HYF, Lam HYS, Fong AH-T, *et al*. Frequency and Distribution of Chest Radiographic
45 Findings in COVID-19 Positive Patients. *Radiology* 2020;:201160.
46 doi:10/ggqbp4
47
48
- 49
50 20 Cohen JF, Korevaar DA, Altman DG, *et al*. STARD 2015 guidelines for reporting diagnostic
51 accuracy studies: explanation and elaboration. *BMJ Open* 2016;**6**:e012799.
52 doi:10.1136/bmjopen-2016-012799
53
54
55
56
57
58
59
60

- 1
2
3
4 21 Rubin GD, Ryerson CJ, Haramati LB, *et al.* The Role of Chest Imaging in Patient
5 Management during the COVID-19 Pandemic: A Multinational Consensus
6 Statement from the Fleischner Society. *Radiology* 2020;:201365.
7 doi:10/ggrmg4
8
9
10
11 22 ACR Recommendations for the use of Chest Radiography and Computed Tomography (CT)
12 for Suspected COVID-19 Infection. [https://www.acr.org/Advocacy-and-Economics/ACR-](https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection)
13 [Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-](https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection)
14 [COVID19-Infection](https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection) (accessed 5 Jun 2020).
15
16
17
18 23 British Society of Thoracic Imaging. COVID-19: BSTI STATEMENT AND GUIDANCE.
19 2020;:1. https://www.bsti.org.uk/media/resources/files/COVID11.3.20_2.pdf (accessed 5 Jun
20 2020).
21
22
23
24 24 Schiaffino S, Tritella S, Cozzi A, *et al.* Diagnostic Performance of Chest X-Ray for COVID-
25 19 Pneumonia During the SARS-CoV-2 Pandemic in Lombardy, Italy. *J Thorac Imaging*
26 2020; **Publish Ahead of Print**. doi:10/ggx268
27
28
29
30 25 Weinstock MB, Echenique A, Russell JW, *et al.* Chest X-Ray Findings in 636 Ambulatory
31 Patients with COVID-19 Presenting to an Urgent Care Center: A Normal Chest X-Ray Is no
32 Guarantee. ;:10.
33
34
35
36 26 Bao C, Liu X, Zhang H, *et al.* Coronavirus Disease 2019 (COVID-19) CT Findings: A
37 Systematic Review and Meta-analysis. *J Am Coll Radiol* 2020; **17**:701–9. doi:10/ggr28p
38
39
40 27 Borghesi A, Zigliani A, Masciullo R, *et al.* Radiographic severity index in COVID-19
41 pneumonia: relationship to age and sex in 783 Italian patients. *Radiol Med (Torino)*
42 2020; **125**:461–4. doi:10/ggtvwp
43
44
45
46 28 Borghesi A, Maroldi R. COVID-19 outbreak in Italy: experimental chest X-ray scoring system
47 for quantifying and monitoring disease progression. *Radiol Med (Torino)* 2020; **125**:509–13.
48 doi:10/ggtvwn
49
50
51 29 Borghesi A, Zigliani A, Golemi S, *et al.* Chest X-ray severity index as a predictor of in-
52 hospital mortality in coronavirus disease 2019: A study of 302 patients from Italy. *Int J Infect*
53 *Dis* 2020; **96**:291–3. doi:10.1016/j.ijid.2020.05.021
54
55
56
57
58
59
60

- 1
2
3
4 30 Xu Y-H, Dong J-H, An W-M, *et al.* Clinical and computed tomographic imaging features of
5 novel coronavirus pneumonia caused by SARS-CoV-2. *J Infect* 2020;**80**:394–400.
6 doi:10/ggqwf3
7
8
9 31 Zhang J-J, Dong X, Cao Y-Y, *et al.* Clinical characteristics of 140 patients infected with
10 SARS-CoV-2 in Wuhan, China. *Allergy* Published Online First: 19 February 2020.
11 doi:10/ggpx6g
12
13
14 32 Lodigiani C, Iapichino G, Carenzo L, *et al.* Venous and arterial thromboembolic
15 complications in COVID-19 patients admitted to an academic hospital in Milan, Italy. *Thromb*
16 *Res* 2020;**191**:9–14. doi:10/ggvcft
17
18
19 33 Chien C-H, Shih F-C, Chen C-Y, *et al.* Unenhanced multidetector computed tomography
20 findings in acute central pulmonary embolism. *BMC Med Imaging* 2019;**19**:65.
21 doi:10/ggzg85
22
23
24 34 Mohamed N, Othman MoustafaHM, Hassan L, *et al.* The accuracy of non-contrast chest
25 computed tomographic Scan in the detection of pulmonary thromboembolism. *J Curr Med*
26 *Res Pract* 2019;**4**:61. doi:10/ggzg83
27
28
29 35 McCollough CH, Bushberg JT, Fletcher JG, *et al.* Answers to Common Questions About the
30 Use and Safety of CT Scans. *Mayo Clin Proc* 2015;**90**:1380–92. doi:10/f3jgqx
31
32
33 36 Moser JB, Sheard SL, Edyvean S, *et al.* Radiation dose-reduction strategies in thoracic CT.
34 *Clin Radiol* 2017;**72**:407–20. doi:10/f95q7p
35
36
37 37 Long Q-X, Liu B-Z, Deng H-J, *et al.* Antibody responses to SARS-CoV-2 in patients with
38 COVID-19. *Nat Med* 2020;:1–4. doi:10.1038/s41591-020-0897-1
39
40
41 38 Wang W, Xu Y, Gao R, *et al.* Detection of SARS-CoV-2 in Different Types of Clinical
42 Specimens. *JAMA* 2020;**323**:1843–4. doi:10/ggpp6h
43
44
45 39 Smith MJ, Hayward SA, Innes SM, *et al.* Point-of-care lung ultrasound in patients with
46 COVID-19 – a narrative review. *Anaesthesia*;n/a. doi:10/ggr2p7
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4 40 Haaksma ME, Heldeweg MLA, Matta JEL, *et al.* Lung ultrasound findings in patients with
5 novel SARS-CoV2. *medRxiv* 2020;:2020.05.18.20105775.
6
7 doi:10.1101/2020.05.18.20105775
8
- 9 41 Benchoufi M, Bokobza J, Chauvin AA, *et al.* Lung injury in patients with or suspected
10 COVID-19 : a comparison between lung ultrasound and chest CT-scanner severity
11 assessments, an observational study. *medRxiv* 2020;:2020.04.24.20069633.
12
13 doi:10.1101/2020.04.24.20069633
14
15
- 16 42 Fine D, Perring S, Herbetko J, *et al.* Three-dimensional (3D) ultrasound imaging of the
17 gallbladder and dilated biliary tree: reconstruction from real-time B-scans. *Br J Radiol*
18 1991;**64**:1056–7. doi:10/fqr9mh
19
20
21
- 22 43 Shi F, Wang J, Shi J, *et al.* Review of Artificial Intelligence Techniques in Imaging Data
23 Acquisition, Segmentation and Diagnosis for COVID-19. *IEEE Rev Biomed Eng* 2020;:1–1.
24
25 doi:10/ggs2km
26
27
- 28 44 Li L, Qin L, Xu Z, *et al.* Artificial Intelligence Distinguishes COVID-19 from Community
29 Acquired Pneumonia on Chest CT. *Radiology* 2020;:200905. doi:10/ggpdgp
30
31
- 32 45 Wang L, Wong A. COVID-Net: A Tailored Deep Convolutional Neural Network Design for
33 Detection of COVID-19 Cases from Chest X-Ray Images. *ArXiv200309871 Cs Eess*
34 Published Online First: 11 May 2020.<http://arxiv.org/abs/2003.09871> (accessed 13 Jun
35 2020).
36
37
38
39
- 40 46 Kotsiantis SB. Use of machine learning techniques for educational proposes: a decision
41 support system for forecasting students' grades. *Artif Intell Rev* 2012;**37**:331–44.
42
43 doi:10/fmbng4
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Tables

Ordinal scale for study	BSTI grade	Features on X-ray
0	CVCX3- Non-COVID-19	Alternative pathology such as pneumothorax with no features of COVID-19 identified
1	CVCX0- Normal	No pathology seen
2	CVCX2- Indeterminate for COVID-19 or atypical features	Poor quality film or central/ basal consolidation
3	CVCX1- Classic findings of COVID-19	Peripheral ground glass opacities

Table 1- Ordinal scale used in this study based on the British Society of Thoracic Imaging (BSTI)

Reporting Template [10]

	SARS-CoV 2 RT-PCR		p-value	Missing (%)
	Negative	Positive		
n (%)	435 (36.3)	763 (63.7)		
Number of Swabs (%)	810 (48.3)	868 (51.7)		
Age (mean (SD))	62.74 (17.72)	66.18 (17.58)	0.001*	0
Ethnicity			0.097	19
Other- Asian (%)	29 (8.0)	72 (11.8)		
South- Asian (%)	27 (7.5)	38 (6.2)		
Black (%)	41 (11.4)	91 (14.9)		
Mixed (%)	6 (1.7)	6 (1.0)		
Other (%)	56 (15.5)	105 (17.2)		
White (%)	202 (56.0)	297 (48.8)		
Sex – Male (%)	233 (53.6)	480 (62.9)	0.002*	0
Oxygen Saturation (median (IQR))	95 (6)	93 (8)	<0.001**	6.3
Respiratory Rate (median (IQR))	22 (8)	26 (12)	<0.001**	6.3
Glasgow Coma Scale (median (IQR))	15 (0)	15 (0)	0.043*	6.6
Systolic BP (median (IQR))	134 (32)	130 (30)	0.009*	15.8
Heart Rate (median (IQR))	96 (27)	94 (27)	0.092	6.4
Temperature (median (IQR))	37.1 (1.4)	37.7 (1.4)	<0.001**	6.7
Chest X-ray report			<0.001**	0
Alternative pathology (%)	4 (0.9)	3 (0.4)		
No abnormalities (%)	178 (40.9)	136 (17.8)		
Indeterminate (%)	83 (19.1)	169 (22.1)		
Classic COVID-19 (%)	170 (39.1)	455 (59.6)		
Presence of comorbidities (%)	297 (79.0)	482 (80.3)	0.669	18.5
Dyspnoea (%)	274 (69.4)	497 (75.5)	0.034	12.1
Neutrophils (median (IQR))	6.42 (4.56)	5.25 (3.92)	<0.001**	2.3
D-Dimer (median (IQR))	1250 (2440)	1105 (1803)	0.204	23.2
Albumin (median (IQR))	39 (7)	37 (6)	<0.001**	10
C-Reactive Protein (median (IQR))	91.0 (115)	146.5 (264.8)	<0.001**	3
Creatine Kinase (median (IQR))	51 (104)	145 (260)	<0.001**	23.3
Troponin (median (IQR))	19 (46)	20 (44)	0.278	19.1
Admitted (%)	331 (76.0)	635 (83.2)	0.003*	0.1
Admitted to ITU (%)	5 (1.3)	32 (4.8)	0.005*	12.4
Thirty Day Follow Up Status			<0.001**	24
Discharged (%)	219 (78.2)	367 (58.3)		
On Ambulatory Follow Up (%)	14 (5.0)	49 (7.8)		
Admitted (%)	18 (6.4)	60 (9.5)		
Died (%)	29 (10.4)	154 (24.4)		
CT report			<0.001**	0
No pathology identified (%)	23 (22.1)	6 (3.3)		
Classic COVID-19 findings (%)	52 (50.0)	157 (85.8)		

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Indeterminate for COVID-19 (%)	14 (13.5)	14 (7.7)		
Alternative pathology identified (%)	15 (14.4)	6 (3.3)		
Day of Symptoms (mean (SD))	9.84 (9.63)	8.56 (15.80)	0.368	69.2

Table 2- Baseline characteristics of dataset stratified by overall SARS-CoV 2 RT-PCR status, including subsequent swabs during the study period- NB there were 480 additional swabs on 399 unique patients with a median of 2 and mean of 3.5 per patient; *significant at $p < 0.05$; **significant at $p < 0.001$

For peer review only

	Chest X-ray	CT Chest	Mean Difference	p-value
Total (n)	860	302		
True Positives (n)	305	162	-	-
False Positives (n)	125	55	-	-
True Negatives (n)	187	56	-	-
False Negatives (n)	243	29	-	-
Apparent prevalence (95% CI)	0.50 (0.47-0.53)	0.72 (0.66-0.77)	0.22 (0.04-0.21)	<0.0001**
True prevalence (95% CI)	0.64 (0.60-0.67)	0.63 (0.58-0.69)	-0.00 (-0.09-0.03)	0.111
Sensitivity (95% CI)	0.56 (0.51-0.60)	0.85 (0.79-0.90)	0.29 (0.19-0.38)	<0.0001**
Specificity (95% CI)	0.60 (0.54-0.65)	0.50 (0.41-0.60)	-0.10 (-0.25-0.04)	0.119
Positive Predictive Value (95% CI)	0.71 (0.66-0.75)	0.75 (0.68-0.80)	0.04 (-0.06-0.14)	0.492
Negative Predictive Value (95% CI)	0.43 (0.39-0.48)	0.66 (0.55-0.76)	0.22 (0.06-0.37)	0.005*
Positive Likelihood Ratio (95% CI)	1.39 (1.19-1.62)	1.71 (1.41- 2.08)	0.32 (-0.22-0.89)	0.258
Negative Likelihood Ratio (95% CI)	0.74 (0.64-0.84)	0.30 (0.21-0.44)	-0.44 (-0.64 - -0.21)	0.022*
Diagnostic Accuracy (95% CI)	0.57 (0.54-0.61)	0.72 (0.66-0.77)	0.15 (0.06-0.23)	<0.0001**

Table 3- Diagnostic Accuracy Metrics for CXR and CT Chest with RT-PCR for SARS-CoV 2, as the reference standard; *significant difference at the <0.05 level; **significant difference at the <0.0001 level

		SARS-CoV 2 RT-PCR		OR (univariable)	OR (multivariable)		
		Negative	Positive				
1							
2							
3							
4	n	312	548				
5							
6	Chest X-ray report	Alternative pathology (%)	3 (0.8)	3 (0.5)	-	-	
7		No abnormalities (%)	123 (39.6)	104 (19.1)	0.76 (0.08-6.82, p=0.801)	0.48 (0.03-8.82, p=0.620)	
8		Indeterminate/ atypical findings (%)	61 (19.5)	136 (4.8)	1.99 (0.22-17.81, p=0.535)	0.92 (0.05-16.88, p=0.952)	
9		Classic COVID (%)	125 (40.1)	305 (55.6)	2.17 (0.24-19.19, p=0.484)	1.14 (0.06-20.98, p=0.927)	
10		Age	Mean (SD)	61.8 (17.9)	67.0 (17.7)	1.02 (1.01-1.02, p<0.001)**	1.02 (1.00-1.03, p=0.028)*
11		Sex	Female (%)	138 (44.3)	212 (38.7)	-	-
12			Male (%)	174 (55.7)	336 (61.3)	1.26 (0.93-1.70, p=0.137)	1.19 (0.83-1.71, p=0.340)
13		Ethnicity	Other Asian (%)	31 (9.9)	66 (12.0)	-	-
14			White (%)	164 (52.7)	270 (49.2)	0.76 (0.44-1.31, p=0.326)	0.73 (0.38-1.40, p=0.339)
15			Black (%)	39 (12.4)	84 (15.3)	1.01 (0.52-1.98, p=0.974)	0.92 (0.43-1.97, p=0.827)
16			Mixed (%)	6 (1.8)	4 (0.8)	0.36 (0.08-1.62, p=0.184)	0.74 (0.11-4.94, p=0.754)
17			South Asian (%)	22 (7.0)	36 (6.6)	0.77 (0.34-1.76, p=0.531)	0.68 (0.28-1.65, p=0.390)
18			Other (%)	51 (16.2)	89 (16.2)	0.82 (0.43-1.55, p=0.535)	0.88 (0.45-1.74, p=0.716)
19		Comorbidity	No (%)	65 (20.8)	95 (17.4)	-	-
20			Yes (%)	247 (79.2)	453 (82.6)	1.25 (0.82-1.89, p=0.296)	1.00 (0.53-1.88, p=0.993)
21		Dyspnoea on attendance	No (%)	90 (28.8)	139 (25.4)	-	-
22			Yes (%)	222 (71.2)	409 (74.6)	1.19 (0.82-1.73, p=0.356)	0.84 (0.53-1.32, p=0.447)
23		Oxygen Saturation	Median (IQR)	96 (6)	93 (8)	0.94 (0.91-0.97, p<0.001)**	0.97 (0.93-1.00, p=0.072)
24		Respiratory rate	Median (IQR)	23 (8)	25 (8)	1.04 (1.01-1.07, p=0.002)*	1.01 (0.98-1.05, p=0.462)
25		Glasgow Coma Scale	Median (IQR)	15 (0)	15 (0)	1.02 (0.89-1.17, p=0.819)	1.21 (0.98-1.48, p=0.073)
26		Temperature	Mean (SD)	37.2 (1.4)	37.7 (1.1)	1.48 (1.26-1.73, p<0.001)**	1.44 (1.20-1.74, p<0.001)**
27		Heart Rate	Mean (SD)	96.7 (20.5)	94.9 (21.5)	1.00 (0.99-1.00, p=0.305)	1.00 (0.99-1.01, p=0.702)
28							
29		Systolic Blood Pressure	Mean (SD)	136.2 (25.8)	132.6 (24.5)	0.99 (0.99-1.00, p=0.086)	0.99 (0.98-1.00, p=0.097)
30							
31		Neutrophils	Median (IQR)	6.26 (4.52)	5.05 (3.93)	0.92 (0.89-0.96, p<0.001)**	0.87 (0.82-0.91, p<0.001)**
32		D-Dimer	Median (IQR)	1220 (2343)	1061 (1814)	1.00 (1.00-1.00, p=0.403)	1.00 (1.00-1.00, p=0.419)
33		C-Reactive Protein	Median (IQR)	45 (100)	77 (107)	1.00 (1.00-1.01, p<0.001)**	1.00 (1.00-1.01, p=0.021)*
34		Troponin	Median (IQR)	20 (55)	21 (46)	1.00 (1.00-1.00, p=0.890)	1.00 (1.00-1.00, p=0.667)
35		Albumin	Median (IQR)	39 (7)	37 (6)	0.97 (0.94-1.00, p=0.071)	1.02 (0.98-1.06, p=0.432)
36		Creatine Kinase	Median (IQR)	94 (131)	145 (263)	1.00 (1.00-1.00, p=0.119)	1.00 (1.00-1.00, p=0.152)
37		Admitted from ED	Admitted (%)	235 (75.2)	453 (82.7)	-	-
38			Discharged (%)	77 (24.8)	95 (17.3)	1.56 (1.06-2.33, p=0.022)**	1.35 (0.79-2.30, p=0.272)
39		Admitted To ITU from ED	No (%)	307 (98.5)	532 (97.1)	-	-
40			Yes (%)	5 (1.5)	16 (2.9)	1.92 (0.60-6.18, p=0.274)	1.06 (0.25-4.40, p=0.940)
41							
42							
43							
44							
45							
46							
47							
48							
49							
50							
51							
52							
53							
54							
55							
56							
57							
58							
59							
60							

Thirty Day Follow up Status	Discharged (%)	259 (83.0)	368 (67.1)	-	-
	Admitted (%)	22 (6.9)	47 (8.5)	1.53 (0.82-2.87, p=0.181)	1.64 (0.77-3.51, p=0.198)
	Dead (%)	31 (10.1)	133 (24.4)	3.00 (1.86-4.84, p<0.001)**	2.81 (1.22-6.50, p=0.017)*

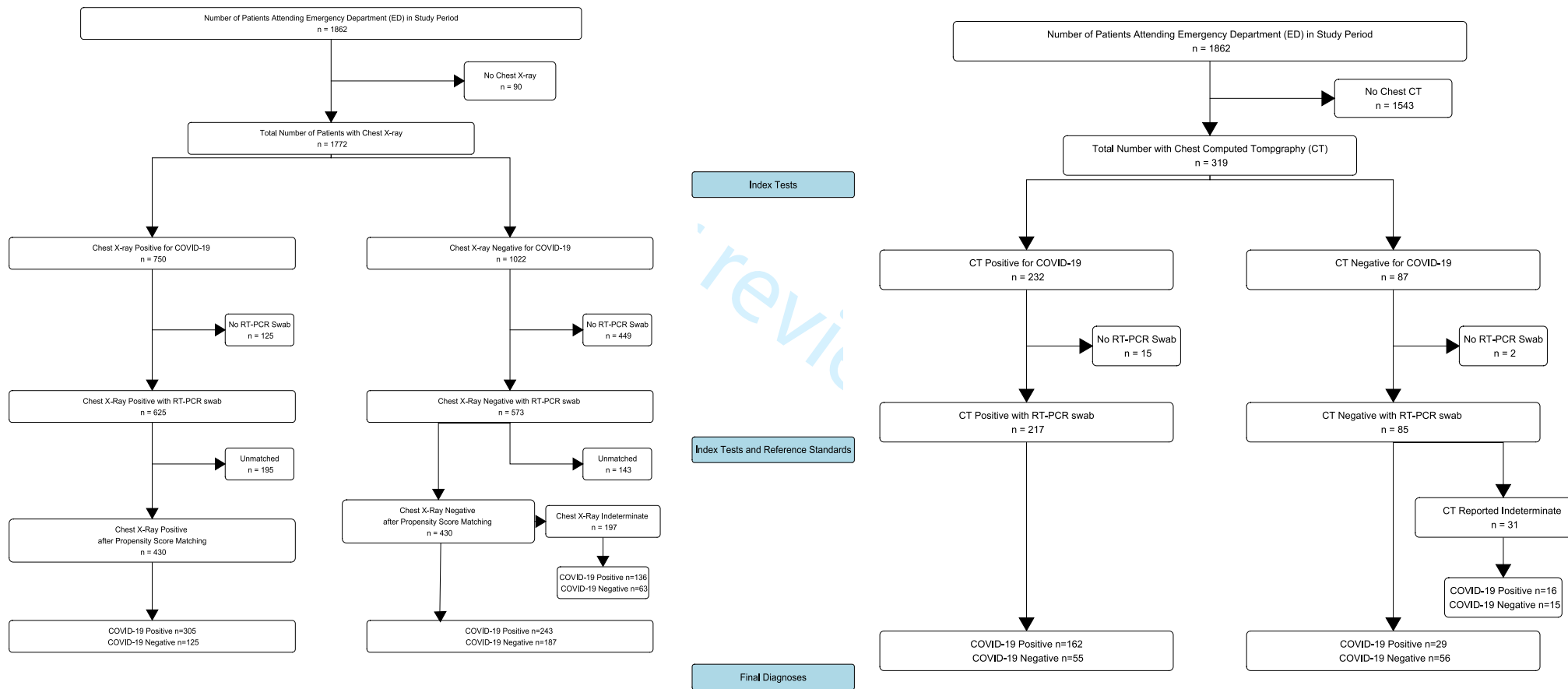
Table 4- Association of covariates with RT-PCR status for SARS-CoV 2, following propensity score matching and binomial logistic regression; SD- Standard deviation; IQR- Interquartile Range; *p<0.05; **p<0.001

		X-ray report		OR (univariable)	OR with XR as binary outcome (multivariable)	OR with XR as ordinal variable (multivariable)
		Other X-ray Findings	Classical COVID-19			
n		430	430			
RT-PCR for SARS-CoV 2	Negative (%)	187 (43.4)	125 (29.1)	-	-	-
	Positive (%)	243 (56.6)	305 (70.9)	1.85 (1.36-2.56, p<0.001)**	1.79 (1.25-2.56, p<0.002)*	1.94 (1.37-2.76, p<0.001)**
Age	Mean (SD)	65.0 (18.9)	65.3 (16.9)	1.00 (0.99-1.01, p=0.849)	0.99 (0.98-1.00, p=0.164)	1.00 (0.99-1.01, p=0.542)
Sex	Female (%)	176 (40.9)	175 (40.6)	-	-	-
	Male (%)	254 (59.1)	255 (59.3)	1.01 (0.75-1.37, p=0.940)	0.87 (0.63-1.20, p=0.400)	1.02 (0.49-2.09, p=0.967)
Ethnicity	Other Asian (%)	49 (11.4)	48 (11.2)	-	-	-
	South Asian (%)	29 (6.7)	29 (6.7)	1.04 (0.52-2.04, p=0.912)	1.02 (0.47-2.17, p=0.965)	1.02 (0.49-2.09, p=0.967)
	Black (%)	61 (14.2)	61 (14.2)	1.02 (0.55-1.85, p=0.957)	0.88 (0.46-1.69, p=0.719)	0.92 (0.52-1.65, p=0.789)
	Mixed (%)	5 (1.2)	5 (1.2)	0.92 (0.21-4.00, p=0.911)	0.86 (0.18-4.17, p=0.853)	0.85 (0.17-4.30, p=0.838)
	Other (%)	70 (16.3)	70 (16.3)	1.02 (0.58-1.79, p=0.943)	0.98 (0.52-1.82, p=0.942)	0.93 (0.53-1.64, p=0.810)
	White (%)	216 (50.2)	217 (50.5)	1.03 (0.63-1.67, p=0.913)	0.97 (0.57-1.67, p=0.926)	0.90 (0.55-1.47, p=0.666)
Comorbidity	No (%)	82 (19.1)	78 (18.1)	-	-	-
	Yes (%)	348 (80.9)	352 (81.9)	0.95 (0.66-1.36, p=0.777)	0.93 (0.59-1.49, p=0.782)	0.88 (0.57-1.37, p=0.592)
Dyspnoea	No (%)	191 (29.3)	103 (24.0)	-	-	-
	Yes (%)	304 (70.7)	327 (76.0)	1.31 (0.92-1.88, p=0.123)	1.20 (0.80-1.82, p=0.380)	1.22 (0.83-1.80, p=0.301)
Oxygen Saturation	Median (IQR)	95 (7)	93 (7)	0.94 (0.91-0.96, p<0.001)**	0.94 (0.92-0.97, p<0.001)**	0.94 (0.91-0.97, p<0.001)**
Respiratory rate	Median (IQR)	24 (10)	24 (10)	1.01 (0.99-1.02, p=0.570)	0.97 (0.94-1.00, p=0.063)	0.98 (0.96-1.01, p=0.157)
Glasgow Coma Scale	Median (IQR)	15 (0)	15 (0)	1.04 (0.92-1.19, p=0.524)	1.05 (0.90-1.23, p=0.503)	1.05 (0.92-1.21, p=0.464)
Temperature	Mean (SD)	37.6 (1.1)	37.5 (1.3)	0.93 (0.83-1.06, p=0.297)	0.79 (0.67-0.93, p=0.006)*	0.85 (0.73-0.99, p=0.031)*
Heart Rate	Mean (SD)	95.7 (21.4)	95.5 (21.0)	1.00 (0.99-1.01, p=0.888)	1.00 (0.99-1.01, p=0.864)	1.00 (0.99-1.01, p=0.872)
Systolic Blood Pressure	Mean (SD)	133.8 (25.0)	134.0 (25.6)	1.00 (0.99-1.01, p=0.907)	1.00 (0.99-1.01, p=0.335)	1.00 (1.00-1.01, p=0.478)
Neutrophils	Median (IQR)	5.44 (4.54)	5.67 (4.03)	1.00 (0.97-1.04, p=0.892)	0.96 (0.92-1.01, p=0.143)	0.96 (0.92-1.01, p=0.115)
D-Dimer	Median (IQR)	1119 (2221)	1119 (1850)	1.00 (1.00-1.00, p=0.513)	1.00 (1.00-1.00, p=0.568)	1.00 (1.00-1.00, p=0.385)

1							
2							
3	C-Reactive	Median (IQR)	46 (93)	88 (110)	1.00 (0.99-1.00,	1.00 (1.00-1.01,	1.00 (1.00-1.01,
4	Protein				p<0.001)**	p<0.001)**	p<0.001)**
5							
6	Troponin	Median (IQR)	23 (54)	20 (46)	1.00 (1.00-1.00, p=0.231)	1.00 (1.00-1.00, p=0.277)	1.00 (1.00-1.00, p=0.059)
7	Albumin	Median (IQR)	39 (7)	37 (6)	0.93 (0.90-0.96,	0.93 (0.90-0.97, p=0.001)*	0.94 (0.91-0.97, p=0.001)*
8					p<0.001)**		
9							
10	Creatine Kinase	Median (IQR)	110 (183)	134 (239)	1.00 (1.00-1.00, p=0.535)	1.00 (1.00-1.00, p=0.242)	1.00 (1.00-1.00, p=0.186)
11	Admitted from	Admitted (%)	315 (73.3)	373 (86.7)	2.37 (1.63-3.46,	2.30 (1.46-3.63,	2.22 (1.47-3.33,
12	ED				p<0.001)**	p<0.001)**	p<0.001)**
13		Discharged (%)	115 (26.7)	57 (13.3)	-	-	-
14							
15	Admitted to ITU	No (%)	423 (98.4)	416 (96.7)	-	-	-
16	from ED						
17		Yes (%)	7 (1.6)	14 (3.3)	2.17 (0.69-6.67, p=0.181)	1.27 (0.32-5.00, p=0.732)	1.34 (0.36-5.00, p=0.653)
18							
19	30 Day Follow	Discharged (%)	316 (73.5)	311 (72.3)	-	-	-
20	Up Status						
21		Admitted (%)	34 (7.9)	34 (7.9)	1.31 (0.81-2.13, p=0.282)	1.32 (0.69-2.53, p=0.392)	1.43 (0.78-2.63, p=0.653)
22							
23		Dead (%)	80 (18.6)	85 (19.8)	1.03 (0.73-1.45, p=0.886)	1.38 (0.80-2.37, p=0.247)	1.41 (0.87-2.27, p=0.157)

Table 5- Association of covariates with CXR report following propensity score matching and either binomial or ordinal logistic regression; SD- Standard deviation; IQR- Interquartile Range; *p<0.05; **p<0.001

Figures



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Figure 1- Inclusion and exclusion of patients during study period with test results

For peer review only

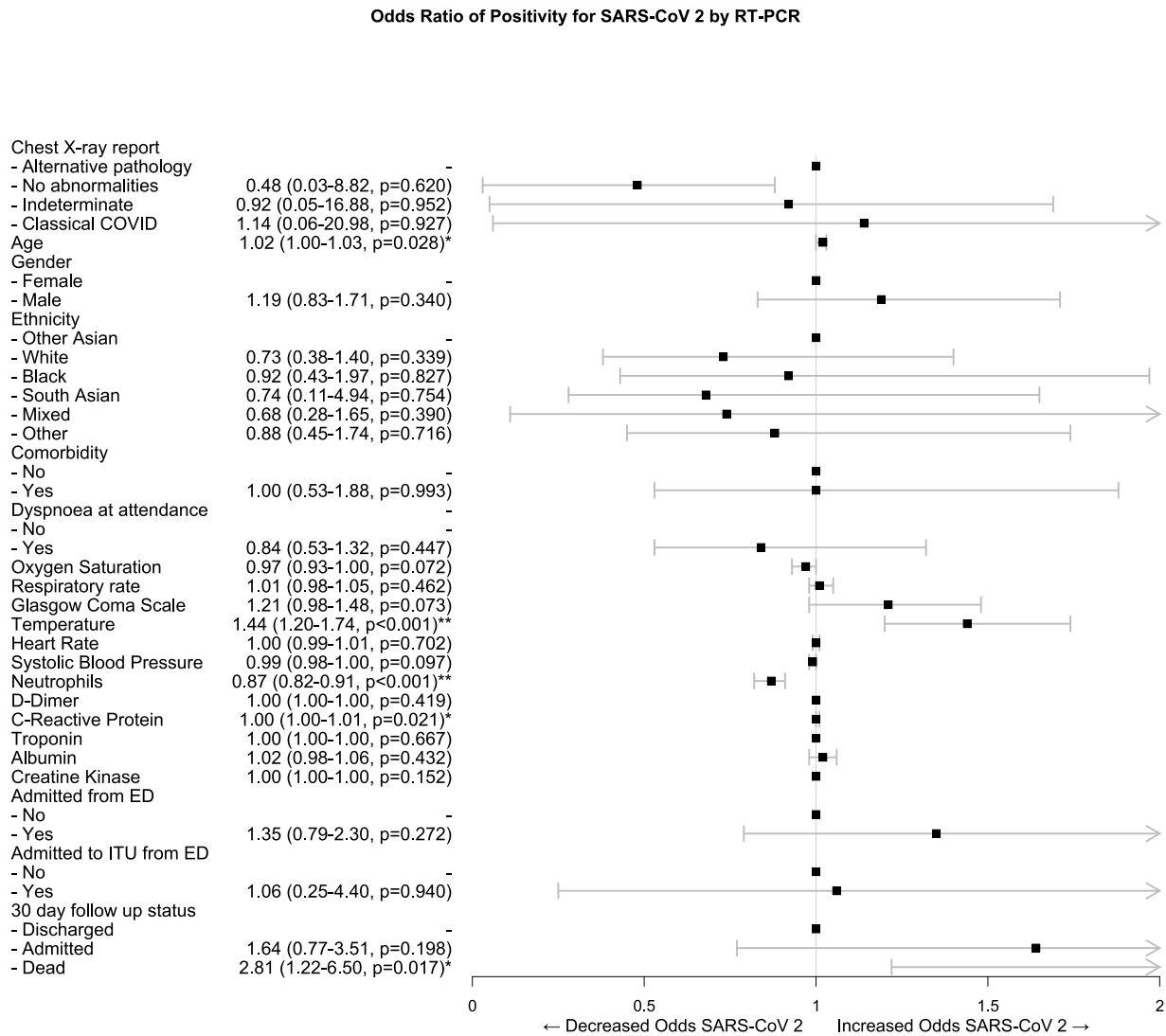


Figure 2- Forest plot of odds ratios of variables associated with RT-PCR positivity for SARS-CoV 2, following multiple imputation, propensity score matching and binomial logistic regression; *significant difference at the <0.05 level; **significant difference at the <0.001 level

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Odds Ratio of Classical COVID-19 Findings on Chest X-Ray

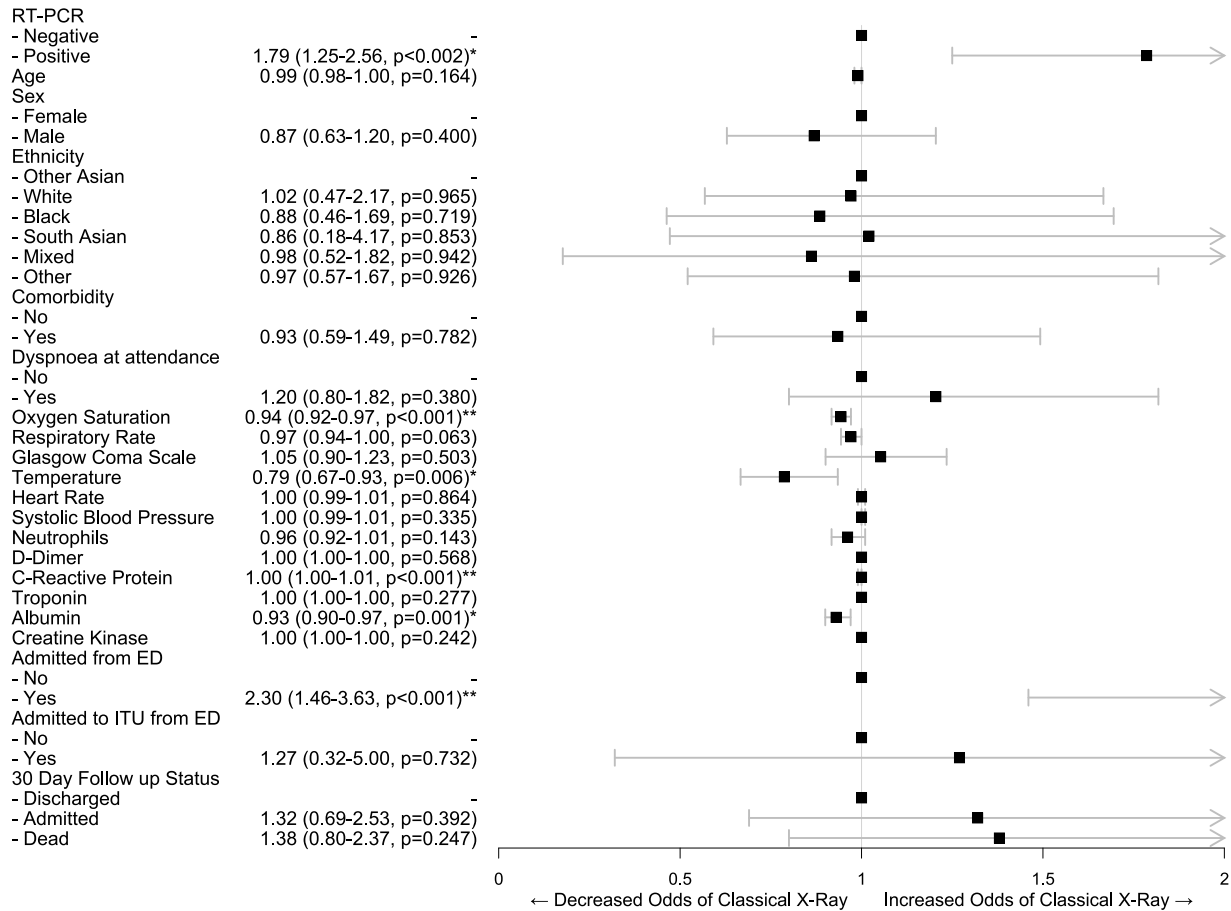
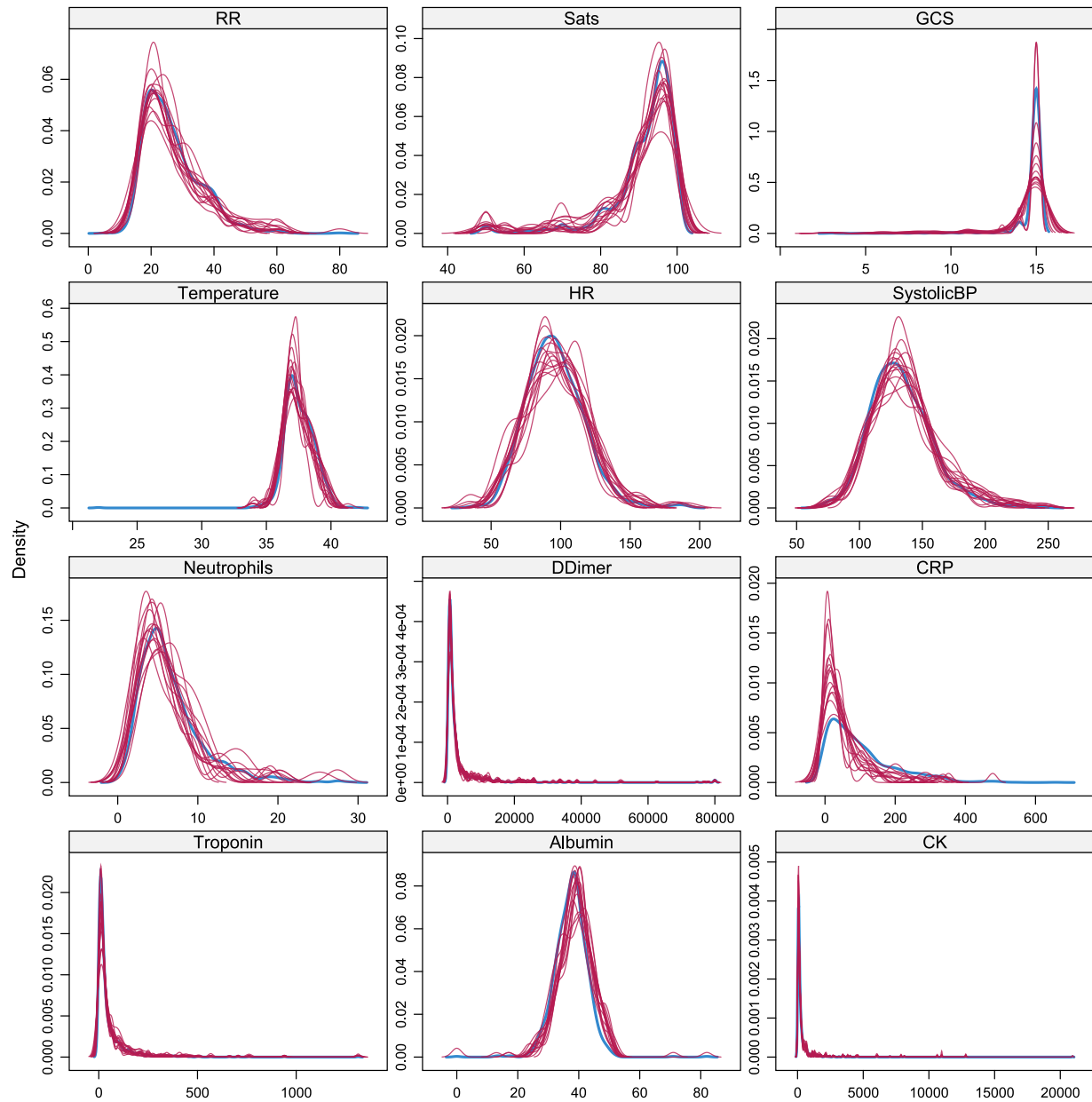


Figure 3- Forest plot of odds ratios of variables associated with classical Chest X-ray features COVID-19 following propensity score matching and binomial logistic regression; *significant difference at the <0.05 level; **significant difference at the <0.001 level

Supplementary file 1



Supplementary figure 1- Density plots of imputed datasets; Blue represents original dataset; other colours are individual imputed datasets (n=15)

Covariate:	Means Treated	Means Control	Standard Deviation Control	Mean Difference
Overall Propensity Score	0.422997940	0.53935303	0.1449627	-0.1163550897
Female	36.3782051	45.026178	0.4979547	-8.64797288
Male	63.6217949	54.973822	0.4979547	8.64797288
Age	63.796474359	66.19022688	18.5893357	-23.937525171
Comorbidity- Yes	76.1217949	84.467714	0.3625287	-8.34591892
Ethnicity- South Asian	6.5705128	6.631763	0.2490539	-0.06124983
Ethnicity- Black	16.1858974	11.518325	0.3195219	4.66757283
Ethnicity- Mixed	0.9615385	1.396161	0.1174340	-0.43462210
Ethnicity- Other	18.9102564	13.263525	0.3394765	5.64673110
Ethnicity- White	46.6346154	57.766143	0.4943635	-11.13152772
Respiratory Rate	29.214743590	24.01745201	7.2639816	5.1972915828

Supplementary table 1- Means of data before multiple imputation and propensity score matching

	Type	Minimum Difference Adjusted	Mean Difference Adjusted	Maximum Difference Adjusted
Distance	Distance	0.016988	0.027107	0.040963
Sex = Male	Binary	-0.03917	-0.0028	0.015982
Age	Contin.	-0.04586	-0.01371	0.027589
Comorbidity = Yes	Binary	-0.02331	-0.00778	0.004598
Ethnicity = Other Asian	Binary	-0.01392	0.002362	0.016471
Ethnicity = South Asian	Binary	-0.01399	-0.00136	0.011905
Ethnicity = Black	Binary	-0.01852	0.000443	0.015982
Ethnicity = Mixed	Binary	-0.00464	0.001403	0.007042
Ethnicity = Other	Binary	-0.01152	4.30E-06	0.00939
Ethnicity = White	Binary	-0.02353	-0.00285	0.018433
Respiratory Rate	Contin.	-0.06157	-0.03478	-0.00442

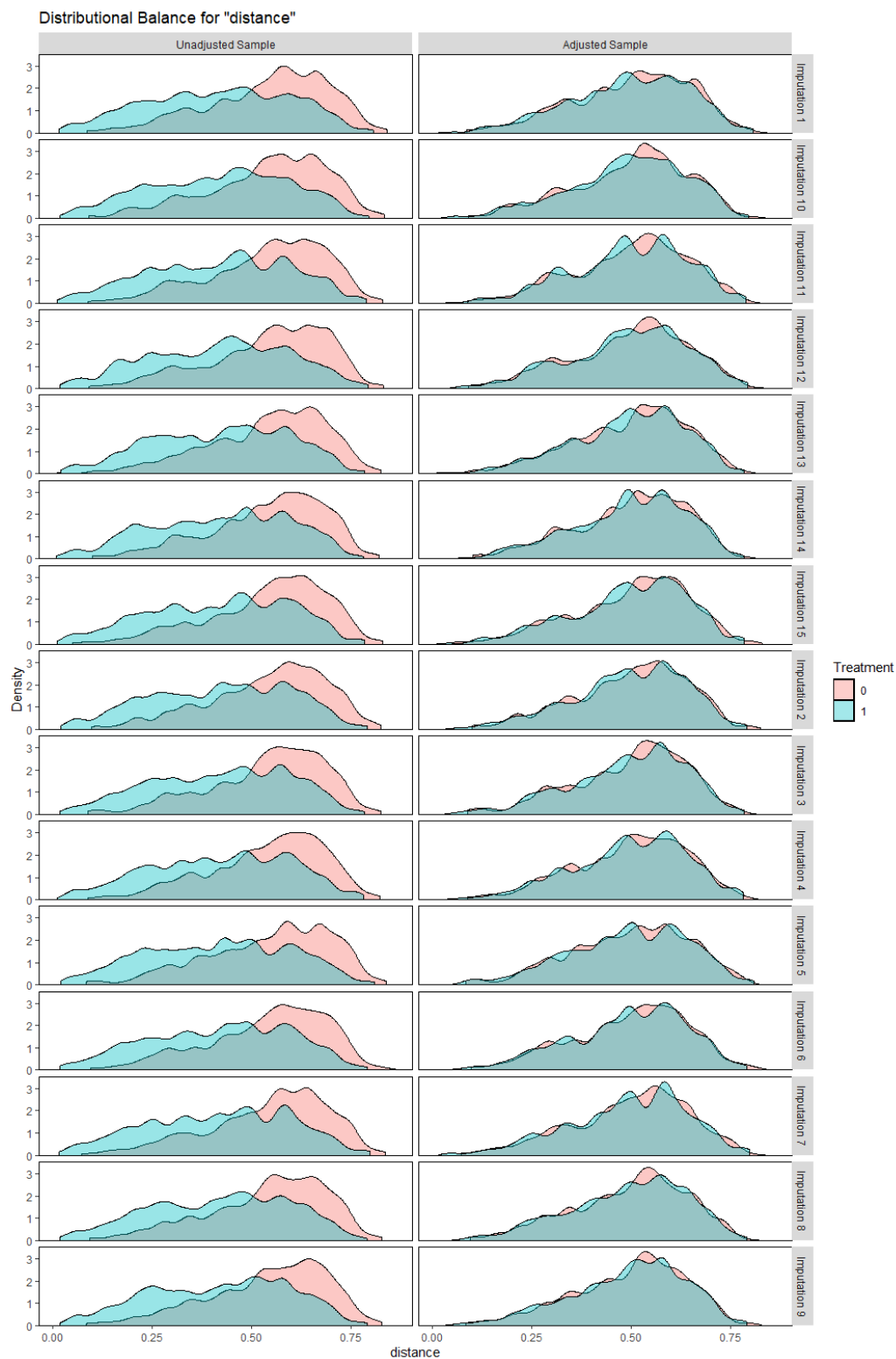
Supplementary table 2- Balance summary across imputations

	XR- Negative	XR- Positive	Total
All	573	625	1,198
Matched	430	430	860
Unmatched	143	195	338
Discarded	0	0	0

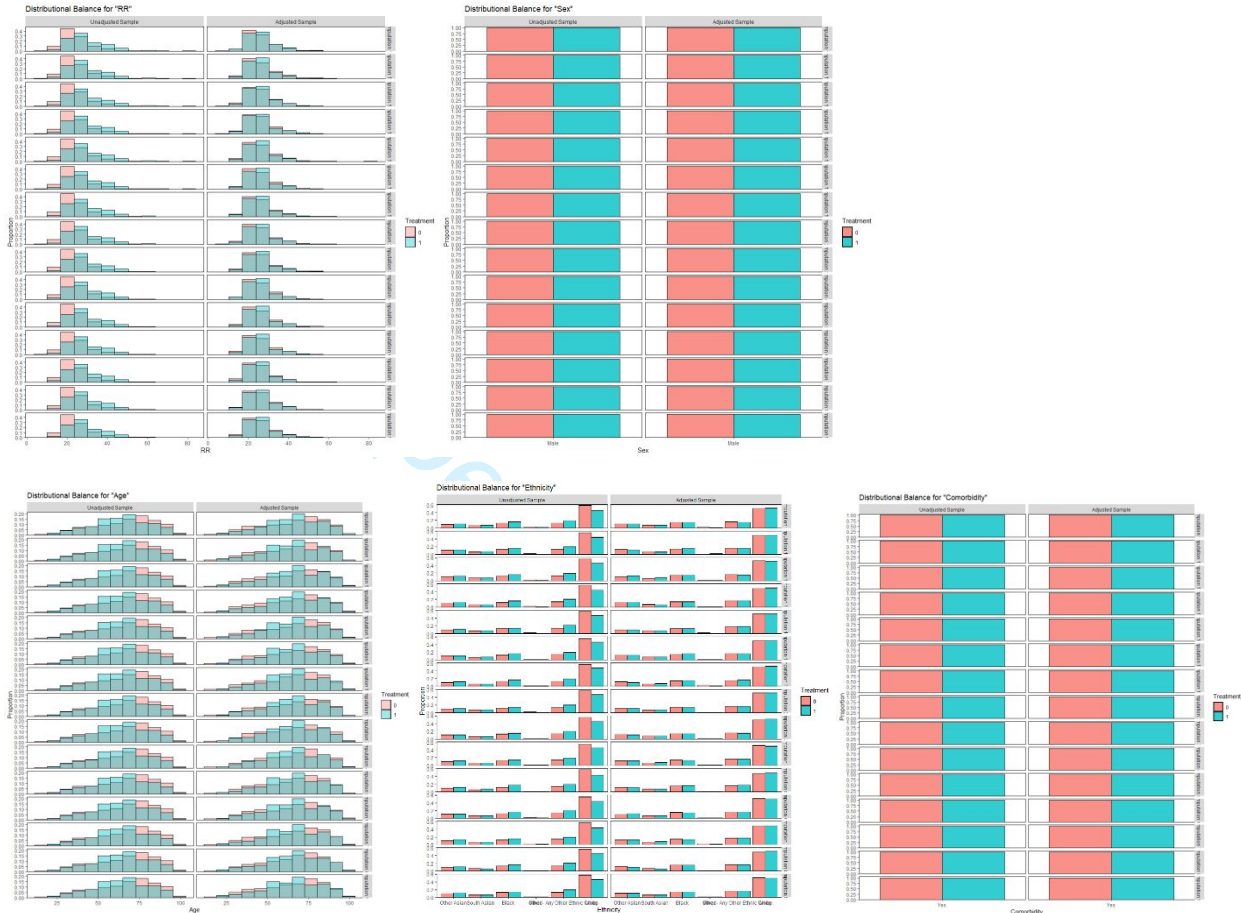
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplementary table 3- Average Sample sizes pre- and post- matching across imputed data sets

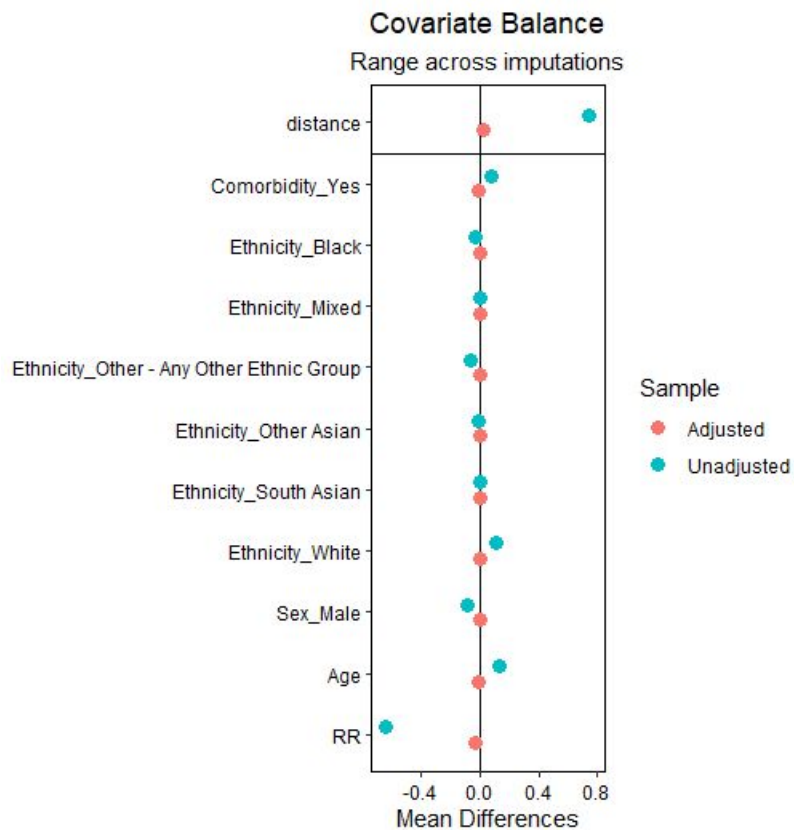
For peer review only



1
2
3
4 **Supplementary figure 2-** Density plot of propensity scores pre- and post- matching in each imputed dataset; treatment units represent a positive X-ray for COVID-19, whereas a control unit represents a
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23



24
25
26
27
28
29
30
31
32
33
34
35
36
37
38 **Supplementary figure 3-** Histogram of distributions for each matching covariate pre- and post- matching in
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Supplementary figure 4- Love plot of pooled balances across imputed datasets in matching covariates after matching

CXR in COVID Analysis

Dr Aditya Borakati

Royal Free Hospital, London, UK

a.borakati@doctors.org.uk

10/06/2020

Software Environment and Packages

R version 4.0.0 (2020-04-24)

Platform: x86_64-w64-mingw32/x64 (64-bit)

Running under: Windows 10 x64 (build 19041)

Matrix products: default

locale:

LC_COLLATE=English_United Kingdom.1252 LC_CTYPE=English_United Kingdom.1252

LC_MONETARY=English_United Kingdom.1252 LC_NUMERIC=C

LC_TIME=English_United Kingdom.1252

attached base packages:

stats graphics grDevices utils datasets methods base

other attached packages:

corrplot 0.84

Taiyun Wei and Viliam **Simko** (2017). R package "**corrplot**": Visualization of a Correlation **Matrix** (Version 0.84). Available from

<https://github.com/taiyun/corrplot>

MKmisc 1.6

Kohl **M** (2019). MKmisc: Miscellaneous functions from M. Kohl. R package version 1.6,

<http://www.stamats.de>

epiR 1.0-14

Mark Stevenson with contributions from Telmo Nunes, Cord Heuer, Jonathon Marshall, Javier Sanchez, Ron Thornton, Jeno Reiczigel, Jim Robison-Cox, Paola Sebastiani, Peter Solymos, Kazuki Yoshida, Geoff Jones, Sarah

Pirikahu, Simon Firestone, Ryan Kyle, Johann Popp, Mathew Jay and Charles

Reynard. (2020). epiR: Tools **for** the Analysis of Epidemiological Data. R

package version 1.0-14. <https://CRAN.R-project.org/package=epiR>

Matching 4.9-7

Jasjeet S. **Sekhon** (2011). Multivariate and Propensity Score Matching

Software with Automated Balance Optimization: The Matching Package **for** R.

Journal of Statistical Software, 42(7), 1-52. URL <http://www.jstatsoft.org/v42/i07/>.

MASS 7.3-51.5

1
2
3 Venables, W. N. & Ripley, B. **D.** (2002) Modern Applied Statistics with S.
4 Fourth Edition. Springer, New York. ISBN 0-387-95457-0
5 Ordinal 2019.12-10
6 Christensen, R. H. **B.** (2019). ordinal - Regression Models **for** Ordinal Data. R package
7 version 2019.12-10. <https://CRAN.R-project.org/package=ordinal>.
8 Hmisc 4.4-0
9 Frank E Harrell Jr, with contributions from Charles Dupont and many
10 **others.** (2020). Hmisc: Harrell Miscellaneous. R package version 4.4-0.
11 <https://CRAN.R-project.org/package=Hmisc>
12 Formula 1.2-3
13 Achim Zeileis, Yves **Croissant** (2010). Extended Model Formulas **in** R:
14 Multiple Parts and Multiple Responses. Journal of Statistical Software
15 34(1), 1-13. doi:10.18637/jss.v034.i01
16 lattice 0.20-41
17 Sarkar, **Deepayan** (2008) Lattice: Multivariate Data Visualization with R.
18 Springer, New York. ISBN 978-0-387-75968-5
19 mice 3.8.0
20 Stef van Buuren, Karin Groothuis-**Oudshoorn** (2011). mice: Multivariate
21 Imputation by Chained Equations **in** R. Journal of Statistical Software,
22 45(3), 1-67. URL <https://www.jstatsoft.org/v45/i03/>.
23 readxl 1.3.1
24 Hadley Wickham and Jennifer **Bryan** (2019). readxl: Read Excel Files. R
25 package version 1.3.1. <https://CRAN.R-project.org/package=readxl>
26 finalfit 1.0.1
27 Ewen Harrison, Tom Drake and Riinu **Ots** (2020). finalfit: Quickly Create
28 Elegant Regression Results Tables and Plots when Modelling. R package
29 version 1.0.1. <https://CRAN.R-project.org/package=finalfit>
30 MatchIt 3.0.2
31 Daniel E. Ho, Kosuke Imai, Gary King, Elizabeth A. **Stuart** (2011). MatchIt:
32 Nonparametric Preprocessing **for** Parametric Causal Inference. Journal of
33 Statistical Software, Vol. 42, No. 8, pp. 1-28. URL
34 <http://www.jstatsoft.org/v42/i08/>
35 tableone 0.11.1
36 Kazuki **Yoshida** (2020). tableone: Create 'Table 1' to Describe Baseline
37 Characteristics. R package version 0.11.1.
38 <https://CRAN.R-project.org/package=tableone>
39 forcats 0.5.0
40 Hadley **Wickham** (2020). forcats: Tools **for** Working with Categorical
41 **Variables** (Factors). R package version 0.5.0.
42 <https://CRAN.R-project.org/package=forcats>
43 stringr 1.4.0
44 Hadley **Wickham** (2019). stringr: Simple, Consistent Wrappers **for** Common
45 String Operations. R package version 1.4.0.
46 <https://CRAN.R-project.org/package=stringr>
47 dplyr 0.8.5
48 Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2020).

1
 2
 3 dplyr: A Grammar of Data Manipulation. R package version **0.8.5**.
 4 <https://CRAN.R-project.org/package=dplyr>
 5 purrr **0.3.4**
 6 Lionel Henry and Hadley **Wickham** (2020). purrr: Functional Programming
 7 Tools. R package version **0.3.4**. <https://CRAN.R-project.org/package=purrr>
 8 readr **1.3.1**
 9 Hadley Wickham, Jim Hester and Romain **Francois** (2018). readr: Read
 10 Rectangular Text Data. R package version **1.3.1**.
 11 <https://CRAN.R-project.org/package=readr>
 12 tidyrr **1.0.2**
 13 Hadley Wickham and Lionel **Henry** (2020). tidyrr: Tidy Messy Data. R package
 14 version **1.0.2**. <https://CRAN.R-project.org/package=tidyrr>
 15 tibble **3.0.0**
 16 Hadley Wickham and Lionel **Henry** (2020). tidyrr: Tidy Messy Data. R package
 17 version **1.0.2**. <https://CRAN.R-project.org/package=tidyrr>
 18 ggplot2 **3.3.0**
 19 H. Wickham. ggplot2: Elegant Graphics **for** Data Analysis. Springer-Verlag
 20 New York, **2016**.
 21 tidyverse **1.3.0**
 22 Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source
 23 Software, **4**(43), **1686**, <https://doi.org/10.21105/joss.01686>
 24 forestplot **1.9**
 25 Max Gordon and Thomas **Lumley** (2019). forestplot: Advanced Forest Plot Using '**grid**'
 26 Graphics. R package version **1.9**. <https://CRAN.R-project.org/package=forestplot>
 27 MatchThem **0.9.3**
 28 Farhad Pishgar and Noah **Greifer** (2020). MatchThem: Matching and Weighting Multiply
 29 Imputed Datasets. R package version **0.9.3**. [https://CRAN.R-](https://CRAN.R-project.org/package=MatchThem)
 30 [project.org/package=MatchThem](https://CRAN.R-project.org/package=MatchThem)
 31 miceadds **3.9-14**
 32 Robitzsch, A., & Grund, S. (2020). miceadds: Some Additional Multiple Imputation
 33 Functions, Especially **for 'mice'**. R package version **3.9-14**. [https://CRAN.R-](https://CRAN.R-project.org/package=miceadds)
 34 [project.org/package=miceadds](https://CRAN.R-project.org/package=miceadds)
 35 cobalt **4.2.2**
 36 Noah **Greifer** (2020). cobalt: Covariate Balance Tables and Plots. R package version **4.2.2**.
 37 <https://CRAN.R-project.org/package=cobalt>
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

Load Packages and Data

Load Packages:

```
library(MKmisc)
```

```
library(tidyverse)
```

```
library(tableone)
```

```
library(MatchIt)
```

```

1
2
3 library(finalfit)
4 library(readxl)
5 library(cobalt)
6 library(mice)
7 library(miceadds)
8 library(Hmisc)
9 library(epiR)
10 library(MatchThem)
11 library(ordinal)
12 library(forestplot)
13
14
15

```

Power Calculation

This code calculates the sample size (positive and negative by gold standard test) needed to evaluate a diagnostic test with 56% sensitivity at 80% power with alpha 0.05. The 56% value is the lower confidence reported by Wong et al. and lower sensitivities typically require higher sample sizes, the result is the same whether specificity or sensitivities are passed as arguments, the previously published specificities are higher than sensitivities so for a generous estimate, the sensitivity was used.

```

24 power.diagnostic.test(sens = 0.56,
25
26     sig.level = 0.05,
27     delta = 0.1,
28     power = 0.8) %>% print()->power
29
30

```

Diagnostic test exact power calculation

```

34     sens = 0.56
35     n = 165
36     n1 = 165
37     delta = 0.1
38     sig.level = 0.05
39     power = 0.8
40     prev = NULL
41
42
43

```

NOTE: n is number of cases, n1 is number of controls

Load Data:

```

49
50 data <- read_csv(
51
52     "FullDataWithCT.csv",
53     col_types = cols(
54         Age = col_integer(),
55         Albumin = col_number(),
56
57
58
59
60

```



```

1
2
3
4 CK = col_number(),
5 CT = col_character(),
6 CRP = col_number(),
7 DDimer = col_number(),
8 DateOfDeath = col_date(format = "%d/%m/%Y"),
9 DateOfDischarge = col_date(format = "%d/%m/%Y"),
10 DateOfVisit = col_date(format = "%d/%m/%Y"),
11 DateOfSymptomOnset = col_date(format = "%d/%m/%Y"),
12 DiastolicBP = col_number(),
13 FiO2 = col_skip(),
14 GCS = col_number(),
15 HR = col_number(),
16 MRN = col_skip(),
17 NEWS = col_number(),
18 'NEWS2(noFiO2)' = col_skip(),
19 Neutrophils = col_number(),
20 RR = col_number(),
21 Sats = col_number(),
22 'Supplemental Oxygen' = col_skip(),
23 SystolicBP = col_number(),
24 Temperature = col_number(),
25 Troponin = col_number(),
26 CTBSTI = col_integer()
27 )
28 )
29
30
31
32
33
34

```

Data Cleaning

Format data into factors/ differences between dates:

```

41 data <- mutate_if(data, is.character, as.factor)
42
43 data$DayOfSymptoms <-
44   difftime(data$DateOfVisit, data$DateOfSymptomOnset, units = "days")
45 data$TimeToDeath <-
46   abs(difftime(data$DateOfDeath, data$DateOfVisit, units = "days"))
47 data$DayOfSymptoms <- as.numeric(data$DayOfSymptoms)
48 data$TimeToDeath <- as.numeric(data$TimeToDeath)
49
50

```

Recode ethnicities as too many options:

This code collapses the ethnicity categories into 'White', 'Black', 'South Asian', 'Other Asian', 'Mixed' or 'Other';

```

1
2
3
4 fct_collapse(
5
6   data$Ethnicity,
7   White = c(
8     "White - British",
9     "White - Irish",
10    "White - Any Other White Background"
11  )
12 ) -> data$Ethnicity
13
14 fct_collapse(
15   data$Ethnicity,
16   Black = c(
17     "Black - Any Other Black Background",
18     "Black or Black British - African",
19     "Black or Black British - African",
20     "Black or Black British - Caribbean"
21  )
22 ) -> data$Ethnicity
23
24 fct_collapse(
25   data$Ethnicity,
26   'South Asian' = c(
27     "Asian or Asian British - Bangladeshi",
28     "Asian or Asian British - Indian",
29     "Asian or Asian British - Pakistani"
30  )
31 ) -> data$Ethnicity
32
33 fct_collapse(data$Ethnicity,
34   'Other Asian' = c("Asian - Any Other Asian Background",
35     "Other - Chinese")) -> data$Ethnicity
36
37 fct_collapse(
38   data$Ethnicity,
39   'Mixed' = c(
40     "mixed - Any Other mixed Background",
41     "Mixed - Any Other Mixed Background",
42     "Mixed - White and Asian",
43     "Mixed - White and Black African",
44     "mixed - White and Black Caribbean",
45     "Mixed - White and Black Caribbean"
46  )
47 ) -> data$Ethnicity
48
49

```

New XR positive column for "Classic Covid" or not:

```

50
51
52
53 data$XRPositive <-
54
55 ifelse(data$XR Chest == "Classic COVID", "Positive", "Negative")
56
57
58
59
60

```

```
data$XRPositive <- as.factor(data$XRPositive)
```

Follow Up Swabs + Initial Swabs Positive:

Creates new column 'OverallPos' which includes initial RT-PCR swab and follow-up swabs in 30 days of attendance, if any are positive the value will be positive in this column

```
data$OverallPos<-case_when(data$RTPCR == "Positive" | data$FollowUpPos ==
"Positive"~"Positive")
replace_na(data$OverallPos,"Negative")->data$OverallPos
```

Create new vector with all variable names (i.e. the column headers)

```
explanatory <- names(data)
```

Paired XR and RT-PCR data

Creates new variable 'completedata' which contains only patients who had both CXR and RT-PCR in ED

```
completedata <- filter(data, !is.na(data$XRPositive) & !is.na(data$RTPCR))
```

Remove missing data variable

```
completedata <- completedata[-c(31)]
```

Format complete data variables

```
completedata$OverallPos <- as.factor(completedata$OverallPos)
```

```
completedata$ThirtyDayFU<-as.factor(completedata$ThirtyDayFU)
```

```
completedata$TimeToDeath <-
```

```
  abs(difftime(completedata$DateOfDeath,
    completedata$DateOfVisit, units = "days"))
```

```
completedata$TimeToDeath <- as.numeric(completedata$TimeToDeath)
```

Set 'XRchest' as ordinal variable on scale of 'Alternative pathology' as lowest value and 'Classical COVID' as highest

```
completedata$XRchest <- ordered(
```

```
  completedata$XRchest,
  levels = c(
```

```

"Alternative pathology",
"No abnormalities",
"Indeterminate",
"Classic COVID"
)
)

```

Convert CT BSTI grade column into factor:

```
completedata$CTBSTI<-as.factor(completedata$CTBSTI)
```

Demographic table of raw data

This code creates an unformatted demographic table (table 2 in manuscript), for the raw data, stratified by RT-PCR status, significance testing between RT-PCR +ve and -ve groups is carried out automatically using chi squared, t-tests, ANOVA etc.; there is also a column for the proportion of missing data

```

CreateTableOne(vars = explanatory,
               strata = 'OverallPos',
               data = completedata) -> demogtable

##### List nonnormal factors for summarisation as median / IQR and non parametric statistical
test

explanatorynonnormal<-c("Sats","RR","GCS","SystolicBP","Temperature","HR",
"Neutrophils",
+ "DDimer","Albumin","CRP","CK","Troponin")
as.data.frame(print(demogtable, nonnormal = explanatorynonnormal, missing = TRUE))-
>demogtable

write.csv(demogtable, file = "Demogtable.csv")

```

Age (mean (SD))	62.74 (17.72)	66.18 (17.58)	0.001
Ethnicity (%)		0.097	
Other Asian	29 (8.0)	72 (11.8)	
South Asian	27 (7.5)	38 (6.2)	
Black	41 (11.4)	91 (14.9)	
Mixed	6 (1.7)	6 (1.0)	
Other - Any Other Ethnic Group	56 (15.5)	105 (17.2)	
White	202 (56.0)	297 (48.8)	
Sex = Male (%)	233 (53.6)	480 (62.9)	0.002
Sats (median [IQR])	95.00 [92.00, 98.00]	93.00 [88.00, 96.00]	<0.001
nonnorm			
RR (median [IQR])	22.00 [20.00, 28.00]	26.00 [20.00, 32.00]	<0.001
nonnorm			

1				
2				
3	GCS (median [IQR])	15.00 [15.00, 15.00]	15.00 [15.00, 15.00]	0.043
4	nonnorm			
5	SystolicBP (median [IQR])	134.00 [119.00, 151.50]	130.00 [115.00, 145.00]	0.009
6	nonnorm			
7	DiastolicBP (mean (SD))	79.54 (16.40)	75.61 (14.51)	<0.001
8	HR (median [IQR])	96.00 [83.00, 110.00]	94.00 [81.00, 108.00]	0.092
9	nonnorm			
10	Temperature (median [IQR])	37.10 [36.60, 38.00]	37.70 [37.00, 38.40]	<0.001
11	nonnorm			
12	XR Chest (%)		<0.001	
13	Alternative pathology	4 (0.9)	3 (0.4)	
14	No abnormalities	178 (40.9)	136 (17.8)	
15	Indeterminate	83 (19.1)	169 (22.1)	
16	Classic COVID	170 (39.1)	455 (59.6)	
17	CTPA = PE (%)	16 (30.2)	28 (45.9)	0.127
18	Comorbidity = Yes (%)	297 (79.0)	482 (80.3)	0.669
19	Dyspnoea = Yes (%)	274 (69.4)	497 (75.5)	0.034
20	Neutrophils (median [IQR])	6.42 [4.55, 9.11]	5.25 [3.69, 7.61]	<0.001
21	nonnorm			
22	DDimer (median [IQR])	1250.00 [619.00, 3059.00]	1105.00 [626.00, 2428.50]	
23	0.204 nonnorm			
24	Albumin (median [IQR])	39.00 [35.00, 42.00]	37.00 [34.00, 40.00]	<0.001
25	nonnorm			
26	CRP (median [IQR])	51.00 [13.00, 117.00]	83.00 [42.00, 158.00]	<0.001
27	nonnorm			
28	CK (median [IQR])	91.00 [54.00, 169.00]	146.50 [78.00, 342.75]	<0.001
29	nonnorm			
30	Troponin (median [IQR])	19.00 [7.00, 53.00]	20.00 [9.00, 53.00]	0.278
31	nonnorm			
32	Admitted = Discharged (%)	104 (24.0)	128 (16.8)	0.003
33	AdmittedToITU = Yes (%)	5 (1.3)	32 (4.8)	0.005
34	RTPCR = Positive (%)	0 (0.0)	738 (96.7)	<0.001
35	CT = 1 (%)	37 (57.8)	26 (86.7)	0.011
36	NEWS (mean (SD))	4.36 (3.06)	5.48 (2.71)	0.032
37	ThirtyDayFU (%)		<0.001	
38	1	219 (78.2)	367 (58.3)	
39	2	14 (5.0)	49 (7.8)	
40	3	18 (6.4)	60 (9.5)	
41	4	29 (10.4)	154 (24.4)	
42	CTBSTI (%)		<0.001	
43	0	23 (22.1)	6 (3.3)	
44	1	52 (50.0)	157 (85.8)	
45	2	14 (13.5)	14 (7.7)	
46	3	15 (14.4)	6 (3.3)	
47	DayOfSymptoms (mean (SD))	9.84 (9.63)	8.56 (15.80)	0.368
48	TimeToDeath (mean (SD))	50.33 (77.93)	57.76 (70.02)	0.618
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				

XRPositive = Positive (%)	170 (39.1)	455 (59.6)	<0.001
OverallPos = Positive (%)	0 (0.0)	763 (100.0)	

Limited dataset comprising relevant data and those without significant missingness:

```
limcompletedata <- dplyr::select(completedata,
```

```
  c("Age",
    "XR Chest",
    "Ethnicity",
    "Sex",
    "RR",
    "Sats",
    "GCS",
    "Temperature",
    "HR",
    "SystolicBP",
    "DiastolicBP",
    "Neutrophils",
    "DDimer",
    "CRP",
    "Troponin",
    "Albumin",
    "CK",
    "OverallPos",
    "Admitted",
    "AdmittedToITU",
    "ThirtyDayFU",
    "Dyspnoea",
    "Comorbidity",
    "XRPositive"))
```

Imputation

This code generates 15 imputed datasets using the permuted mean matching method, based on the 'limcompletedata' dataset which has filtered the most relevant fields, with minimal missing data initially

```
imputed <- mice(limcompletedata, m = 15, method = 'pmm')
```

Imputation Diagnostics Density plot, this corresponds to supplementary figure 1:

```
densityplot(imputed)
```



Propensity Score Matching

This code matches data in the imputed datasets on whether the XR was reported classical COVID or not, the matching is done based on the covariates Sex, Age, Comorbidity, Ethnicity and Respiratory Rate

```
library(MatchThem)
```

```
##### MatchThem package requires dependent variable to be coded as 0 or 1
imputed[["data"]][["XRPositive"]] %>% recode_factor("Positive" = "1", "Negative" = "0") -
->imputed[["data"]][["XRPositive"]]
matchthem(
  XRPositive ~ Sex + Age + Comorbidity + Ethnicity + RR,
  data = imputed,
  method = 'nearest',
  verbose = FALSE,
  replace = FALSE,
  ratio = 1,
  caliper = 0.2,
  m.order = "random",) -> matchedtest
### Set XRchest to unordered for binomial analyses
matchedtest[["datasets"]]c(1:15)[["XRchest"]] %>% factor(ordered = FALSE) ->
matched2[["datasets"]]c(1:15)[["XRchest"]]
```

Match Balance Diagnostics

Creates plots and table with mean difference and distribution of values in covariates between XR +ve and -ve groups after matching across all imputed datasets:

```
##### Supplementary tables 1,2 and 3:
```

```
bal.tab(matchedtest)
##### Supplementary figure 2
bal.plot(matchedtest)
##### Supplementary figure 3:
bal.plot(matchedtest, var.name = "Age", type = "histogram", which = "both")
```

```

1
2
3
4 bal.plot(matchedtest, var.name = "Sex", type = "histogram", which = "both")
5 bal.plot(matchedtest, var.name = "Ethnicity", type = "histogram", which = "both")
6 bal.plot(matchedtest, var.name = "RR", type = "histogram", which = "both")
7 bal.plot(matchedtest, var.name = "Comorbidity", type = "histogram", which = "both")
8 ##### Supplementary figure 4:
9 love.plot(matchedtest)
10

```

Matched Demographics Table:

Stack matched imputed datasets into one large dataset and split into COVID +ve and -ve groups:

```

11
12
13
14
15 ##### 'all=FALSE' gets matched data only
16
17 stacked<-MatchThem::complete(matchedtest, n = c(1:15), all = FALSE)
18 stacked<-stacked %>% filter(.imp>0)
19

```

Creates demographics table as above, but on propensity matched imputed datasets, corresponds to Table 4:

```

20
21
22
23 CreateTableOne(strata = "OverallPos", data = stacked)-> table4
24
25 ##### Means and SD kept as is, mean counts calculated after dividing by 15 (as 15 imputed
26 datasets)
27

```

Creates demographic table stratified by XR Positive or Negative on matched imputed datasets, corresponds to Table 5:

```

28
29
30
31 CreateTableOne(strata = "XRPositive", data = stacked)-> table5
32
33 ##### Means and SD kept as is, mean counts calculated after dividing by 15 (as 15 imputed
34 datasets)
35

```

Summary statistics for pooled data:

```

36
37
38 ##### Normal means sd
39
40 explanatorynorm<-c("Age","Temperature","HR","SystolicBP")
41 stacked %>% group_by(OverallPos) %>%
42 summarise_at(vars(explanatorynorm),list(mean.default, sd))->summarynormalOverallPos
43 stacked %>% group_by(XRPositive) %>%
44 summarise_at(vars(explanatorynorm),list(mean.default, sd))->summarynormalXRPositive
45
46
47 ##### Non normal medians and IQR
48 stacked %>% group_by(OverallPos) %>%
49 summarise_at(vars(explanatorynorm),list(median, IQR))->summarynormalOverallPos
50 stacked %>% group_by(XRPositive) %>%
51 summarise_at(vars(explanatorynorm),list(median, IQR))->summarynormalXRPositive
52
53
54
55
56
57
58
59
60

```


Diagnostic Accuracy

This section generates the diagnostic accuracy statistics (e.g. sensitivity, specificity) for CXR and CT with RT-PCR as the reference standard using the matched imputed datasets

This code creates a contingency table of False/ True Positives and Negatives for Chest X-ray taken from the demographic tables above:

```
contingxr <-
  matrix(c(305,243,125,187),
        nrow = 2,
        ncol = 2)

colnames(contingxr) <- c("PCR+", "PCR-")

rownames(contingxr) <- c("XR+", "XR-")
```

This function calculates diagnostic accuracy test statistics:

```
epi.tests(contingxr, conf.level = 0.95) -> xraccuracy
```

Giving the diagnostic accuracy output for CXR in table 3:

```
xraccuracy
```

	Outcome +	Outcome -	Total
Test +	305	125	430
Test -	243	187	430
Total	548	312	860

Point estimates and 95 % CIs:

Apparent prevalence	0.50 (0.47, 0.53)
True prevalence	0.64 (0.60, 0.67)
Sensitivity	0.56 (0.51, 0.60)
Specificity	0.60 (0.54, 0.65)
Positive predictive value	0.71 (0.66, 0.75)
Negative predictive value	0.43 (0.39, 0.48)
Positive likelihood ratio	1.39 (1.19, 1.62)
Negative likelihood ratio	0.74 (0.65, 0.84)

NB diagnostic accuracy values in table available in list view of xraccuracy variable

CT Data and Accuracy

Only those with CT and RT PCR:

```
CTdata <-  
  filter(data, is.na(data$CTBSTI) == FALSE &  
           is.na(data$RTPCR) == FALSE)
```

Select relevant variables

```
CTdata <-  
  dplyr::select(CTdata, c("Age",  
                           "XRChest",  
                           "Ethnicity",  
                           "Sex",  
                           "RR",  
                           "Sats",  
                           "GCS",  
                           "Temperature",  
                           "HR",  
                           "SystolicBP",  
                           "DiastolicBP",  
                           "Neutrophils",  
                           "DDimer",  
                           "CRP",  
                           "Troponin",  
                           "OverallPos",  
                           "Admitted",  
                           "AdmittedToITU",  
                           "ThirtyDayFU",  
                           "Dyspnoea",  
                           "Comorbidity",  
                           "XRPositive",  
                           "OverallPos",  
                           "CTBSTI"))
```

Set RT-PCR as factor:

```
CTdata$OverallPos<-as.factor(CTdata$OverallPos)
```

Rename 1 and 0 to Positive and Negative:

```

1
2
3
4 CTdata$CTPositive <-
5
6   ifelse(CTdata$CTBSTI == "1", "Positive", "Negative")
7 CTdata$CTPositive <- as.factor(CTdata$CTPositive)
8

```

Regression with CT as outcome variable:

```

10
11 CT <- finalfit(
12
13   CTdata,
14   "OverallPos",
15   c(
16     "Age",
17     "Sex",
18     "RR",
19     "GCS",
20     "CTPositive",
21     "Temperature",
22     "HR",
23     "SystolicBP",
24     "DiastolicBP",
25     "Sats",
26     "Dyspnoea",
27     "Comorbidity"
28   ),
29   confint_level = 0.95
30 )
31
32
33
34
35

```

Contingency table of True/False Positives and Negatives for CT taken from Regression table:

```

36
37
38 contingct <-
39
40   matrix(c(CT[7,4], CT[6,4], CT[7,3], CT[6,3]),
41     nrow = 2,
42     ncol = 2)
43
44   colnames(contingct) <- c("PCR+", "PCR-")
45   rownames(contingct) <- c("CT+", "CT-")
46   substr(contingct, start = 1, stop = 3)->contingct
47   sapply(contingct,as.numeric)->contingct
48   matrix(contingct, nrow = 2, ncol = 2)->contingct
49   colnames(contingct) <- c("PCR+", "PCR-")
50
51   rownames(contingct) <- c("CT+", "CT-")
52
53

```

Diagnostic accuracy statistics for CT

```
epi.tests(contingct, conf.level = 0.95) -> ctaccuracy
```

	Outcome +	Outcome -	Total
Test +	162	55	217
Test -	29	56	85
Total	191	111	302

Point estimates and 95 % CIs:

Apparent prevalence	0.72 (0.66, 0.77)
True prevalence	0.63 (0.58, 0.69)
Sensitivity	0.85 (0.79, 0.90)
Specificity	0.50 (0.41, 0.60)
Positive predictive value	0.75 (0.68, 0.80)
Negative predictive value	0.66 (0.55, 0.76)
Positive likelihood ratio	1.71 (1.41, 2.08)
Negative likelihood ratio	0.30 (0.21, 0.44)

NB Diagnostic accuracy values found in list view rather than output

CT and XR accuracy comparison

In this section mean differences of diagnostic accuracy statistics between CT and Chest X-ray with confidence intervals and p-values are calculated

Sensitivity

Upper confidence limit for difference in sensitivity

```
ubsens<-(ctaccuracy[["elements"]][["se.up"]]-xraccuracy[["elements"]][["se.low"]])
```

Lower confidence limit for difference in sensitivity

```
lbsens<-(ctaccuracy[["elements"]][["se.low"]]-xraccuracy[["elements"]][["se.up"]])
```

Mean difference in sensitivity

```
meansens<-ctaccuracy[["elements"]][["se"]]-xraccuracy[["elements"]][["se"]]
```

Standard error for sensitivity

```
sesens<-(ubsens-lbsens)/(2*1.96)
```

value for difference in sensitivity

```
meansens/sesens->zsens
```

P-value for difference in sensitivity

```
psens <- exp(-0.717*zsens - 0.416*zsens^2)
```

Format values into 'mean difference (95% CI) p-value' rounded to 2 d.p.

```
sprintf("%s (%s-%s)",
        round(meansens, digits = 2), round(lbsens, digits = 2),
        round(ubsens, digits = 2))->diffsens
diffsensp<-c(diffsens,psens)
```

Subsequent analyses in this section follow the code above

```
##Specificity
```

```
ubspec<-(ctaccuracy[["elements"]][["sp.up"]]-xraccuracy[["elements"]][["sp.low"]])
lbspec<-(ctaccuracy[["elements"]][["sp.low"]]-xraccuracy[["elements"]][["sp.up"]])
meanspec<-ctaccuracy[["elements"]][["sp"]]-xraccuracy[["elements"]][["sp"]]
sespec<-(ubspec-lbspec)/(2*1.96)
meanspec/sespec->zspec
pspec <- exp(-0.717*zspec - 0.416*zspec^2)
sprintf("%s (%s-%s)",
        round(meanspec, digits = 2), round(lbspec, digits = 2),
        round(ubspec, digits = 2))->diffspec
diffspecp<-c(diffspec,pspec)

ubda<-(ctaccuracy[["elements"]][["da.up"]]-xraccuracy[["elements"]][["da.low"]])
lbda<-(ctaccuracy[["elements"]][["da.low"]]-xraccuracy[["elements"]][["da.up"]])
meanda<-ctaccuracy[["elements"]][["da"]]-xraccuracy[["elements"]][["da"]]
seda<-(ubda-lbda)/(2*1.96)
meanda/seda->zda
pda <- exp(-0.717*zda - 0.416*zda^2)
sprintf("%s (%s-%s)",
        round(meanda, digits = 2), round(lbda, digits = 2),
        round(ubda, digits = 2))->diffda
diffdap<-c(diffda,pda)
##Positive Likelihood Ratio
```

```

1
2
3
4 ublrpos<- (ctaccuracy[["elements"]][["lrpos.up"]]-xraccuracy[["elements"]][["lrpos.low"]])
5 lblrpos<- (ctaccuracy[["elements"]][["lrpos.low"]]-xraccuracy[["elements"]][["lrpos.up"]])
6 meanlrpos<- (ctaccuracy[["elements"]][["lrpos"]]-xraccuracy[["elements"]][["lrpos"]])
7 selrpos<- (ublrpos-lblrpos)/(2*1.96)
8 meanlrpos/selrpos->zlrpos
9 plrpos <- exp(-0.717*zlrpos - 0.416*zlrpos^2)
10
11 sprintf("%s (%s-%s)",
12         round(meanlrpos, digits = 2), round(lblrpos, digits = 2),
13         round(ublrpos, digits = 2))->difflrpos
14 difflrposp<-c(difflrpos,plrpos)
15 ##Negative Likelihood Ratios
16 ublrneg<- (ctaccuracy[["elements"]][["lrneg.up"]]-xraccuracy[["elements"]][["lrneg.low"]])
17 lblrneg<- (ctaccuracy[["elements"]][["lrneg.low"]]-xraccuracy[["elements"]][["lrneg.up"]])
18 meanlrneg<- (ctaccuracy[["elements"]][["lrneg"]]-xraccuracy[["elements"]][["lrneg"]])
19 selrneg<- (ublrneg-lblrneg)/(2*1.96)
20 meanlrneg/selrneg->zlrneg
21 plrneg <- exp(-0.717*zlrneg - 0.416*zlrneg^2)
22
23 sprintf("%s (%s-%s)",
24         round(meanlrneg, digits = 2), round(lblrneg, digits = 2),
25         round(ublrneg, digits = 2))->difflrneg
26 difflrnegp<-c(difflrneg,plrneg)
27
28
29
30 ##Positive Predictive Value
31 ppv<- (ctaccuracy[["elements"]][["ppv.low"]]-xraccuracy[["elements"]][["ppv.up"]])
32 meanppv<- (ctaccuracy[["elements"]][["ppv"]]-xraccuracy[["elements"]][["ppv"]])
33 seppv<- (ubppv-lbppv)/(2*1.96)
34 meanppv/seppv->zppv
35 pppv <- exp(-0.717*zppv - 0.416*zppv^2)
36
37 sprintf("%s (%s-%s)",
38         round(meanppv, digits = 2), round(lbppv, digits = 2),
39         round(ubppv, digits = 2))->diffppv
40 diffppvvp<-c(diffppv,pppv)
41
42
43 ##Negative Predictive Value
44 npv<- (ctaccuracy[["elements"]][["npv.low"]]-xraccuracy[["elements"]][["npv.up"]])
45 meannpv<- (ctaccuracy[["elements"]][["npv"]]-xraccuracy[["elements"]][["npv"]])
46 senpv<- (ubnpv-lbnpv)/(2*1.96)
47 meannpv/senpv->znpv
48 pnpv <- exp(-0.717*znpv - 0.416*znpv^2)
49
50 sprintf("%s (%s-%s)",
51         round(meannpv, digits = 2), round(lbnpv, digits = 2),
52         round(ubnpv, digits = 2))->diffnpv
53 diffnpvp<-c(diffnpv,pnpv)
54
55
56
57
58
59
60

```

```

1
2
3
4 ##Apparent Prevalence
5 meantp<-ctaccuracy[["elements"]][["tp"]]-xraccuracy[["elements"]][["tp"]]
6 setp<-(ubtp-lbtp)/(2*1.96)
7 meantp/setp->ztp
8 ptp <- exp(-0.717*ztp - 0.416*ztp^2)
9 sprintf("%s (%s-%s)",
10         round(meantp, digits = 2), round(lbtp, digits = 2),
11         round(ubtp, digits = 2))->difftp
12 diffftp<-c(difftp,ptp)
13
14
15 ##True Prevalence
16 meanap<-ctaccuracy[["elements"]][["ap"]]-xraccuracy[["elements"]][["ap"]]
17 seap<-(ubap-lbap)/(2*1.96)
18 meanap/seap->zap
19 pap <- exp(-0.717*zap - 0.416*zap^2)
20 sprintf("%s (%s-%s)",
21         round(meanap, digits = 2), round(lbap, digits = 2),
22         round(ubap, digits = 2))->diffap
23 diffapp<-c(diffap,pap)
24
25
26
27
28

```

Intermodality Agreement

This section contains code to analyse the level of agreement in the unmatched CT dataset which contains only data with CT, XR and RT-PCR

First- comparing CT and XR agreement

```

33
34 library(irr)
35
36 kappa2(c(CTdata$XRPositive,CTdata$CTPositive), weight = "squared")
37 d<-CTdata %>% select(c("CTPositive","XRPositive"))
38 View(d)
39 kappa2(d, weight = "squared")
40

```

Output:

Cohen's Kappa for 2 Raters (Weights: squared)

```

41
42
43
44
45
46 Subjects = 287
47 Raters = 2
48 Kappa = 0.406
49
50
51 z = 7.14
52 p-value = 9.37e-13
53

```

The following code compares RT-PCR, CT and XR

```
d2<-CTdata %>% select(c("CTPositive", "XRPositive", "OverallPos"))
```

```
View(d2)
```

```
kappam.fleiss(d2)
```

Output:

Fleiss' Kappa for m Raters

Subjects = 287

Raters = 3

Kappa = 0.361

z = 10.6

p-value = 0

Diagnostic Accuracy Analysis when Indeterminate Reports of CXR and CT are taken as positive

XR Indeterminates

New column for positive if indeterminate

```
stacked$XRIndPositive<-ifelse(stacked$XR Chest=="Classic COVID" | stacked$XR Chest ==  
"Indeterminate",
```

```
"Positive", "Negative")
```

```
stacked$XRIndPositive<-as.factor(stacked$XRIndPositive)
```

```
stacked %>% filter(OverallPos == "Positive")->stackedpos
```

```
stacked %>% filter(OverallPos == "Negative")->stackedneg
```

```
summary(stackedpos$XRIndPositive)
```

```
summary(stackedneg$XRIndPositive)
```

```
contingxrind<-matrix(c(441,107,186,126),nrow = 2,ncol = 2)
```

```
colnames(contingxrind) <- c("PCR+", "PCR-")
```

```
rownames(contingxrind) <- c("XR+", "XR-")
```

```
epi.tests(contingxrind)->xrindaccuracy
```

In this section mean differences of diagnostic accuracy statistics between CT (when CT indeterminates are not counted as positive) and Chest X-ray with confidence intervals and p-values are calculated, follows the same pattern as code previously

```
##### Sensitivity
```



```

##### Upper confidence limit for difference in sensitivity

ubsens<-(ctaccuracy[["elements"]][["se.up"]]-xrindaccuracy[["elements"]][["se.low"]])
##Lower confidence limit for difference in sensitivity
lbsens<-(ctaccuracy[["elements"]][["se.low"]]-xrindaccuracy[["elements"]][["se.up"]])
##Mean difference in sensitivity
meansens<-ctaccuracy[["elements"]][["se"]]-xrindaccuracy[["elements"]][["se"]]
##Standard error for sensitivity
sesens<-((ubsens-lbsens)/(2*1.96))
##Z value for difference in sensitivity
meansens/sesens->zsens
##P-value for difference in sensitivity
psens <- exp(-0.717*zsens - 0.416*zsens^2)
###Format values into 'mean difference (95% CI) p-value' rounded to 2 d.p.
sprintf("%s (%s-%s)",
        round(meansens, digits = 2), round(lbsens, digits = 2),
        round(ubsens, digits = 2))>diffsens
diffsensp<-c(diffsens,psens)

###Subsequent analyses in this section follow the code above
##Specificity
ubspec<-(ctaccuracy[["elements"]][["sp.up"]]-xrindaccuracy[["elements"]][["sp.low"]])
lbspec<-(ctaccuracy[["elements"]][["sp.low"]]-xrindaccuracy[["elements"]][["sp.up"]])
meanspec<-ctaccuracy[["elements"]][["sp"]]-xrindaccuracy[["elements"]][["sp"]]
sespec<-((ubspec-lbspec)/(2*1.96))
meanspec/sespec->zspec
pspec <- exp(-0.717*zspec - 0.416*zspec^2)
sprintf("%s (%s-%s)",
        round(meanspec, digits = 2), round(lbspec, digits = 2),
        round(ubspec, digits = 2))>diffspec
diffspecp<-c(diffspec,pspec)

ubda<-(ctaccuracy[["elements"]][["da.up"]]-xrindaccuracy[["elements"]][["da.low"]])
lbda<-(ctaccuracy[["elements"]][["da.low"]]-xrindaccuracy[["elements"]][["da.up"]])
meanda<-ctaccuracy[["elements"]][["da"]]-xrindaccuracy[["elements"]][["da"]]
seda<-((ubda-lbda)/(2*1.96))
meanda/seda->zda
pda <- exp(-0.717*zda - 0.416*zda^2)
sprintf("%s (%s-%s)",
        round(meanda, digits = 2), round(lbda, digits = 2),
        round(ubda, digits = 2))>diffda
diffdap<-c(diffda,pda)
##Positive Likelihood Ratio
ublrpos<-(ctaccuracy[["elements"]][["lrpos.up"]]-xrindaccuracy[["elements"]][["lrpos.low"]])
lblrpos<-(ctaccuracy[["elements"]][["lrpos.low"]]-xrindaccuracy[["elements"]][["lrpos.up"]])

```

```

1
2
3
4 meanlrpos<-ctaccuracy[["elements"]][["lrpos"]]-xrindaccuracy[["elements"]][["lrpos"]]
5 selrpos<-(ublrpos-lblrpos)/(2*1.96)
6 meanlrpos/selrpos->zlrpos
7 plrpos <- exp(-0.717*zlrpos - 0.416*zlrpos^2)
8 sprintf("%s (%s-%s)",
9         round(meanlrpos, digits = 2), round(lblrpos, digits = 2),
10        round(ublrpos, digits = 2))->difflrpos
11 difflrposp<-c(difflrpos,plrpos)
12 ##Negative Likelihood Ratios
13 ublrneg<-(ctaccuracy[["elements"]][["lrneg.up"]]-xrindaccuracy[["elements"]][["lrneg.low"]])
14 lblrneg<-(ctaccuracy[["elements"]][["lrneg.low"]]-xrindaccuracy[["elements"]][["lrneg.up"]])
15 meanlrneg<-ctaccuracy[["elements"]][["lrneg"]]-xrindaccuracy[["elements"]][["lrneg"]]
16 selrneg<-(ublrneg-lblrneg)/(2*1.96)
17 meanlrneg/selrneg->zlrneg
18 plrneg <- exp(-0.717*zlrneg - 0.416*zlrneg^2)
19 sprintf("%s (%s-%s)",
20        round(meanlrneg, digits = 2), round(lblrneg, digits = 2),
21        round(ublrneg, digits = 2))->difflrneg
22 difflrnegp<-c(difflrneg,plrneg)
23
24
25
26
27
28 ##Positive Predictive Value
29 ppv<-(ctaccuracy[["elements"]][["ppv.low"]]-xrindaccuracy[["elements"]][["ppv.up"]])
30 meanppv<-ctaccuracy[["elements"]][["ppv"]]-xrindaccuracy[["elements"]][["ppv"]]
31 seppv<-(ubppv-lbppv)/(2*1.96)
32 meanppv/seppv->zppv
33 pppv <- exp(-0.717*zppv - 0.416*zppv^2)
34 sprintf("%s (%s-%s)",
35        round(meanppv, digits = 2), round(lbppv, digits = 2),
36        round(ubppv, digits = 2))->diffppv
37 diffppvp<-c(diffppv,pppv)
38
39
40
41 ##Negative Predictive Value
42 npv<-(ctaccuracy[["elements"]][["npv.low"]]-xrindaccuracy[["elements"]][["npv.up"]])
43 meannpv<-ctaccuracy[["elements"]][["npv"]]-xrindaccuracy[["elements"]][["npv"]]
44 senpv<-(ubnpv-lbnpv)/(2*1.96)
45 meannpv/senpv->znpv
46 pnpv <- exp(-0.717*znpv - 0.416*znpv^2)
47 sprintf("%s (%s-%s)",
48        round(meannpv, digits = 2), round(lbnpv, digits = 2),
49        round(ubnpv, digits = 2))->diffnpv
50 diffnpvp<-c(diffnpv,pnpv)
51
52
53
54
55 ##True Prevalence
56 meantp<-ctaccuracy[["elements"]][["tp"]]-xrindaccuracy[["elements"]][["tp"]]
57
58
59
60

```

```

1
2
3
4 setp<-(ubtp-lbtp)/(2*1.96)
5 meantp/setp->ztp
6 ptp <- exp(-0.717*ztp - 0.416*ztp^2)
7 sprintf("%s (%s-%s)",
8     round(meantp, digits = 2), round(lbtp, digits = 2),
9     round(ubtp, digits = 2))->difftp
10 diffpp<-c(difftp,ptp)
11
12
13 ##Apparent Prevalence
14 meanap<-ctaccuracy[["elements"]][["ap"]]-xrindaccuracy[["elements"]][["ap"]]
15 seap<-(ubap-lbap)/(2*1.96)
16 meanap/seap->zap
17 pap <- exp(-0.717*zap - 0.416*zap^2)
18 sprintf("%s (%s-%s)",
19     round(meanap, digits = 2), round(lbap, digits = 2),
20     round(ubap, digits = 2))->diffap
21 diffapp<-c(diffap,pap)
22
23

```

CT Indeterminates

New column for positive if indeterminate

```

24
25
26
27
28 CTdata$CTIndPositive<-ifelse(CTdata$CTBSTI=="1" | CTdata$CTBSTI == "2",
29
30     "Positive", "Negative")
31
32 CTdata$CTIndPositive<-as.factor(CTdata$CTIndPositive)
33 CTdata %>% group_by(OverallPos, CTIndPositive) %>% summarise(n=n())->valuesctind
34 ctcontingind<-matrix(data = c(178,13,70,41),
35     nrow = 2, ncol = 2)
36
37 colnames(ctcontingind)<-c("PCR+ve", "PCR-ve")
38 rownames(ctcontingind)<-c("CT+ve", "CT-ve")
39 epi.tests(ctcontingind)->ctindaccuracy
40
41
42
43

```

Pooled Regression after Multiple Imputation and Propensity Score Matching

Binomnal Logistic regression with RT-PCR as dependent variable

```

44
45
46
47
48
49
50
51 matchedtest %>% with(glm(formula(ff_formula(dependent = "OverallPos",
52
53     explanatory = c("Age",
54         "Ethnicity",
55         "Sex",
56
57
58
59
60

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

```

```

"RR",
"GCS",
"Temperature",
"HR",
"SystolicBP",
"Neutrophils",
"DDimer",
"CRP",
"Troponin",
"Albumin",
"CK",
"Sats",
"Admitted",
"AdmittedToITU",
"ThirtyDayFUTwo",
"Dyspnoea",
"Comorbidity",
"XRchest")),
family = "binomial"), all = FALSE)->overallposmatchimp
overallposmatchimp %>% pool()->P
multivarpooledeoverallpos = P %>%
fit2df(estimate_name = "OR (multiple imputation)", exp = TRUE)

```

'multivarpooledeoverallpos' produces multivariate odds ratios for each explanatory variable, corresponding to Table 4

Pooled Univariate Odds Ratios for OverallPos as dependent variable

This code is run with each of the explanatory variables in table 4 as arguments to produce their respective odds Ratios in table 4

```

40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

```

```

matchedtest %>% with(glm(formula(ff_formula(dependent = "OverallPos",
explanatory = "XRchest"
)),
family = "binomial"))->overallposmatchimpunivar
overallposmatchimpunivar %>% pool()->P
univarpooledeoverallpos = P %>%
fit2df(estimate_name = "OR (univariate)", exp = TRUE)->univaroverallpos
univaroverallpos

```

Binomial Logistic Regression with Positive Chest X-ray Report as Dependent Variable

This code follows the format above to produce univariate and multivariate odds ratios for each explanatory variable for having a positive XR report

Univariate XRPositive as dependent

(different explanatory variables passed into function to produce Odds ratios for each)

```

matchedtest %>% with(glm(formula(ff_formula(dependent = "XRPositive",
                                explanatory = "Comorbidity"
                                )),
family = "binomial"))->XRchestmatchimp
XRchestmatchimp %>% pool()->P
multivarpooledXRchest = P %>%
  fit2df(estimate_name = "OR (univariate)", exp = TRUE)->univarXRchest
univarXRchest

```

Multivariate XRPositive as dependent

```

matchedtest %>% with(glm(formula(ff_formula(dependent = "XRPositive",
                                explanatory = c("Age",
                                "OverallPos",
                                "Ethnicity",
                                "Sex",
                                "RR",
                                "GCS",
                                "Temperature",
                                "HR",
                                "SystolicBP",
                                "Neutrophils",
                                "DDimer",
                                "CRP",
                                "Troponin",
                                "Albumin",
                                "CK",
                                "Sats",
                                "Admitted",
                                "AdmittedToITU",
                                "ThirtyDayFUTwo",
                                "Dyspnoea",
                                "Comorbidity"
                                )),
family = "binomial"))->XRchestmatchimp
XRchestmatchimp %>% pool()->P
multivarpooledXRchest = P %>%
  fit2df(estimate_name = "OR (multivariate)", exp = TRUE)->multivarXRchest
multivarXRchest

```

Pooled Ordinal Logistic Regression with XRPositive as dependent

This code also produces multivariate odds ratios for table 5, however, uses ordinal linear regression after the CXR report variable is converted to an ordered categorical variable, with alternative pathology as the lowest and classic covid as the highest value (see table 3)

```

matchedtest %>% with(clm(formula = XRchest ~ Age +
                        OverallPos+
                        Ethnicity+
                        Sex+
                        RR+
                        GCS+
                        Temperature+
                        HR+
                        SystolicBP+
                        Neutrophils+
                        DDimer+
                        CRP+
                        Troponin+
                        Sats+
                        Admitted+
                        AdmittedToITU+
                        ThirtyDayFUTwo+
                        Dyspnoea+
                        Comorbidity))->XRchestmatchimpord
pool(object = XRchestmatchimpord[["analyses"]])->P
multivarpooledXRchestord = P %>%
  fit2df(estimate_name = "OR (multivariate)", exp = TRUE)->multivarXRchestord
multivarXRchestord

```

Forest Plots

Creates forest plots for post matched regression tables above:

```

Figure1Forest <- read_excel("Figure1Forest.xlsx",
                           col_types = c("text", "numeric", "numeric",
                                           "numeric", "text", "text"))

tabletext1<-cbind(Figure1Forest$explanatory, Figure1Forest$summary)
forestplot (tabletext1, Figure1Forest$Mean,
            Figure1Forest$Lower, Figure1Forest$Upper, is.summary = FALSE,

```

```

clip = c(0, 2),
  xlab="\u2190 Decreased Odds SARS-CoV 2    Increased Odds SARS-CoV 2
\u2192",
  zero=1, cex=0.9, lineheight = unit(6,"mm"), boxsize=0.4, colgap=unit(6,"mm"),
  lwd.ci=2, ci.vertices=TRUE, ci.vertices.height = 0.4,
title="Odds Ratio of Positivity for SARS-CoV 2 by RT-PCR",
txt_gp=fpTxtGp(label=gpar(cex=1.25),
  ticks=gpar(cex=1.1),
  xlab=gpar(cex = 1.2),
  title=gpar(cex = 1.2)),
graphwidth = unit(200,"mm")
)

```

Figure 2:



Figure 3 (XR dependent):

```

Figure2Forest <- read_excel("Figure2Forest.xlsx",
  col_types = c("text", "numeric", "numeric",
    "numeric", "text", "text"))

tabletext2<-cbind(Figure2Forest$explanatory,Figure2Forest$summary)
forestplot (tabletext2, Figure2Forest$Mean,
  Figure2Forest$Lower, Figure2Forest$Upper, is.summary = FALSE,
  clip = c(0, 2),
  xlab="\u2190 Decreased Odds of Classical X-Ray    Increased Odds of Classical X-
Ray \u2192",
  zero=1, cex=0.9, lineheight = unit(6,"mm"), boxsize=0.5, colgap=unit(6,"mm"),
  lwd.ci=2, ci.vertices=TRUE, ci.vertices.height = 0.4,
title="Odds Ratio of Classical COVID-19 Findings on Chest X-Ray",
txt_gp=fpTxtGp(label=gpar(cex=1.25),
  ticks=gpar(cex=1.1),
  xlab=gpar(cex = 1.2),
  title=gpar(cex = 1.2)),
graphwidth = unit(200,"mm")
)

```



Correlation Matrix

This section creates a plot of correlation between all the variables in the raw data

```
library(corrplot);library(Hmisc)
```

Relevel factors so relevant value is first

```
relevel(data$XRPositive, "Negative")->data$XRPositive
```

```
relevel(data$Admitted, "Discharged")->data$Admitted
relevel(data$AdmittedToITU, "No")->data$AdmittedToITU
```

New variable for correlation matrix

```
cor<-data
```

Remove variables with high missings/ data which won't work e.g. date, RT-PCR removed as it only represents initial ED swab, OverallPos used instead as this includes subsequent swabs in 30 days

```
cor<-subset(data, select = -c(CT,DateOfDeath,DateOfDischarge,RTPCR,
DateOfVisit,DateOfSymptomOnset,FollowUpPos,TimeToDeath,NEWS))'
```

Format and re-name values

```
cor$CTPositive <-
  ifelse(cor$CTBSTI == "1", "Positive", "Negative")
cor$CTPositive<-as.factor(cor$CTPositive)
cor$CTPositive<-relevel(cor$CTPositive,"Negative")
cor$Death<-as.factor(ifelse(cor$ThirtyDayFU == "4", "Dead", "Alive"))
relevel(cor$Death, "Alive")->cor$Death
cor$OverallPos<-as.factor(cor$OverallPos)
cor<-sapply(cor, as.numeric)
```

Create new numerical correlation matrix

```
cor(cor, method = "spearman", use = "pairwise.complete.obs")->cormatrixall
```

This variable also contains p-values so identification of only significant correlations is possible:

```
cormatrixall2 <- rcorr(as.matrix(cor), type = "spearman")
```

Function to create and format correlation matrix plot

```
corrplot(cormatrixall2$r, method = "color", type = "full", order = "hclust",
p.mat = cormatrixall2$p, sig.level = 0.05, insig = "blank",
tl.col = "black", outline = "white",
```



```
title = "Correlation Matrix of Explanatory and Outcome Variables",
line = -1, cex.main = 2, adj.main = 0.5)
```



STARD Flow Diagram

See instructions from <https://www.r-bloggers.com/flow-charts-in-r/>

Produces flow charts in Figure 1, (images need to be stretched out, output as svgs)

```
library(grid)
```

```
library(Gmisc)
```

```
## Warning: package 'Gmisc' was built under R version 4.0.2
```

```
## Loading required package: Rcpp
```

```
## Loading required package: htmlTable
```

```
## Warning: package 'htmlTable' was built under R version 4.0.2
```

```
grid.newpage()
```

```
# set some parameters to use repeatedly
```

```
leftx <- .25
```

```
midx <- .5
```

```
rightx <- .75
```

```
width <- .4
```

```
gp <- gpar(fill = "white")
```

```
# create boxes
```

```
(totalattendance <- boxGrob("Number of Patients Attending Emergency Department (ED) in
Study Period\n n = 1862",
```

```
  x=midx, y=.9, box_gp = gp, width = 0.7))
```

```
(numberwithxr <- boxGrob("Total Number of Patients with Chest X-ray\n n = 1772",
```

```
  x=midx, y=.75, box_gp = gp, width = width))
```

```
# connect boxes like this
```

```
connectGrob(totalattendance, numberwithxr, "v")
```

```
(numberwithoutxr <- boxGrob("No Chest X-ray\n n = 90",
```

```
  x=rightx, y=.825, box_gp = gp, width = unit(2, "inch"), height = .05))
```

```
connectGrob(totalattendance, numberwithoutxr, "-")
```

```
(XRPos <- boxGrob("Chest X-ray Positive for COVID-19 \n n = 750",
```

```
  x=leftx, y=.6, box_gp = gp, width = width))
```

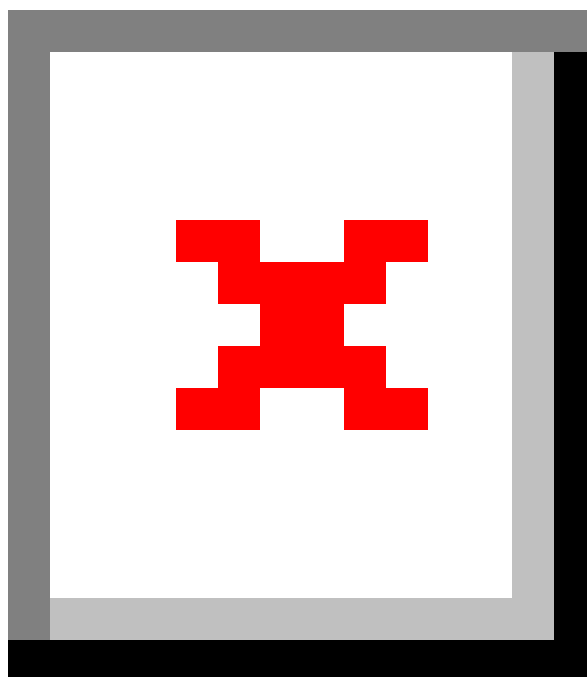
```

1
2
3 (XRNeg <- boxGrob("Chest X-ray Negative for COVID-19\n n = 1022",
4     x=rightx, y=.6, box_gp = gp, width = width))
5
6
7 connectGrob(numberwithxr, XRPos, "N")
8 connectGrob(numberwithxr, XRNeg, "N")
9
10 (RTPCRXRPos <- boxGrob("Chest X-Ray Positive with RT-PCR swab\n n = 625",
11     x=leftx, y=.4, box_gp = gp, width = width))
12 (RTPCRXRNeg <- boxGrob("Chest X-Ray Negative with RT-PCR swab \n n = 573",
13     x=rightx, y=.4, box_gp = gp, width = width))
14
15
16 connectGrob(XRPos, RTPCRXRPos, "N")
17 connectGrob(XRNeg, RTPCRXRNeg, "N")
18
19 (NoRTPCRXRPos <- boxGrob("No RT-PCR Swab\n n = 125",
20     x=0.4, y=.5, box_gp = gp, width = unit(1.5,"inch")))
21 (NoRTPCRXRNeg <- boxGrob("No RT-PCR Swab\n n = 449",
22     x=0.9, y=.5, box_gp = gp, width = unit(1.5,"inch")))
23
24
25 connectGrob(XRPos, NoRTPCRXRPos, "-")
26 connectGrob(XRNeg, NoRTPCRXRNeg, "-")
27
28
29 (MatchedXRPos <- boxGrob("Chest X-Ray Positive \nafter Propensity Score Matching\n n =
30 430",
31     x=leftx, y=.225, box_gp = gp, width = width))
32 (MatchedXRNeg <- boxGrob("Chest X-Ray Negative \nafter Propensity Score Matching \n n
33 = 430",
34     x=0.65, y=.25, box_gp = gp, width = unit(4.2,"inch")))
35
36
37 connectGrob(RTPCRXRPos, MatchedXRPos, "N")
38 connectGrob(RTPCRXRNeg, MatchedXRNeg, "N")
39
40
41
42 (UnmatchedXRPos <- boxGrob("Unmatched\n n = 195",
43     x=0.4, y=.325, box_gp = gp, width = unit(1.5,"inch")))
44 (UnmatchedXRNeg <- boxGrob("Unmatched\n n = 143",
45     x=0.9, y=.325, box_gp = gp, width = unit(1.5,"inch")))
46
47
48 connectGrob(RTPCRXRPos, UnmatchedXRPos, "-")
49 connectGrob(RTPCRXRNeg, UnmatchedXRNeg, "L")
50
51 (DiagXRPositive <- boxGrob("COVID-19 Positive n=305\n COVID-19 Negative n=125",
52     x=leftx, y=0.1, box_gp = gp, width = width))
53 (DiagXRNegative <- boxGrob("COVID-19 Positive n=243 \n COVID-19 Negative n=187",
54     x=rightx, y=0.1, box_gp = gp, width = width))
55
56
57
58
59
60

```

```
1
2
3 connectGrob(MatchedXRPos, DiagXRPositive, "N")
4 connectGrob(MatchedXRNeg, DiagXRNegative, "vertical")
5
6
7
8 (XRInd <- boxGrob("Chest X-Ray Indeterminate \n n = 197",
9                 x=0.88, y=.25, box_gp = gp, width = unit(2.5, "inch")))
10
11 connectGrob(MatchedXRNeg, XRInd, "horizontal")
12
13
14 (DiagXRInd <- boxGrob("COVID-19 Positive n=136\n COVID-19 Negative n=63",
15                     x=0.88, y=0.170, box_gp = gp, width = unit(2, "inch")))
16 connectGrob(XRInd, DiagXRInd, "vertical")
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



```
#####CT Flow Chart#####
```

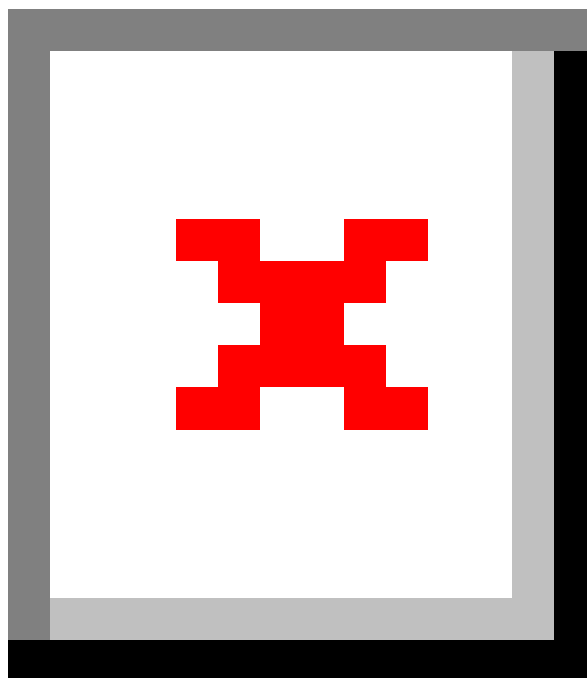
```
grid.newpage()  
(totalattendance <- boxGrob("Number of Patients Attending Emergency Department (ED) in  
Study Period\n n = 1862",  
x=midx, y=.9, box_gp = gp, width = 0.7))
```

```

1
2
3 (numberwithCT <- boxGrob("Total Number with Chest Computed Tomography (CT)\n n =
4 319",
5     x=midx, y=.75, box_gp = gp, width = width))
6 connectGrob(totalattendance, numberwithCT, "vertical")
7
8
9 (numberwithoutCT <- boxGrob("No Chest CT\n n = 1543",
10     x=rightx, y=.825, box_gp = gp, width = unit(2, "inch"), height = .05))
11
12 connectGrob(totalattendance, numberwithoutCT, "-")
13
14
15
16 (CTPos <- boxGrob("CT Positive for COVID-19 \n n = 232",
17     x=leftx, y=.6, box_gp = gp, width = width))
18 (CTNeg <- boxGrob("CT Negative for COVID-19\n n = 87",
19     x=rightx, y=.6, box_gp = gp, width = width))
20
21
22 connectGrob(numberwithCT, CTPos, "N")
23 connectGrob(numberwithCT, CTNeg, "N")
24
25
26 (RTPCRCTPos <- boxGrob("CT Positive with RT-PCR swab\n n = 217",
27     x=leftx, y=.4, box_gp = gp, width = width))
28 (RTPCRCTNeg <- boxGrob("CT Negative with RT-PCR swab \n n = 85",
29     x=rightx, y=.4, box_gp = gp, width = width))
30
31 connectGrob(CTPos, RTPCRCTPos, "N")
32 connectGrob(CTNeg, RTPCRCTNeg, "N")
33
34
35 (NoRTPCRCTPos <- boxGrob("No RT-PCR Swab\n n = 15",
36     x=0.4, y=.5, box_gp = gp, width = unit(1.5, "inch")))
37 (NoRTPCRCTNeg <- boxGrob("No RT-PCR Swab\n n = 2",
38     x=0.9, y=.5, box_gp = gp, width = unit(1.5, "inch")))
39
40 connectGrob(CTPos, NoRTPCRCTPos, "-")
41 connectGrob(CTNeg, NoRTPCRCTNeg, "-")
42
43
44 (DiagCTPositive <- boxGrob("COVID-19 Positive n=162\n COVID-19 Negative n=55",
45     x=leftx, y=0.1, box_gp = gp, width = width))
46 (DiagCTNegative <- boxGrob("COVID-19 Positive n=29\n COVID-19 Negative n=56",
47     x=rightx, y=0.1, box_gp = gp, width = width))
48
49 connectGrob(RTPCRCTPos, DiagCTPositive, "N")
50 connectGrob(RTPCRCTNeg, DiagCTNegative, "N")
51
52
53
54 (CTInd <- boxGrob("CT Reported Indeterminate \n n = 31",
55     x=0.9, y=.275, box_gp = gp, width = unit(3, "inch")))
56
57
58
59
60

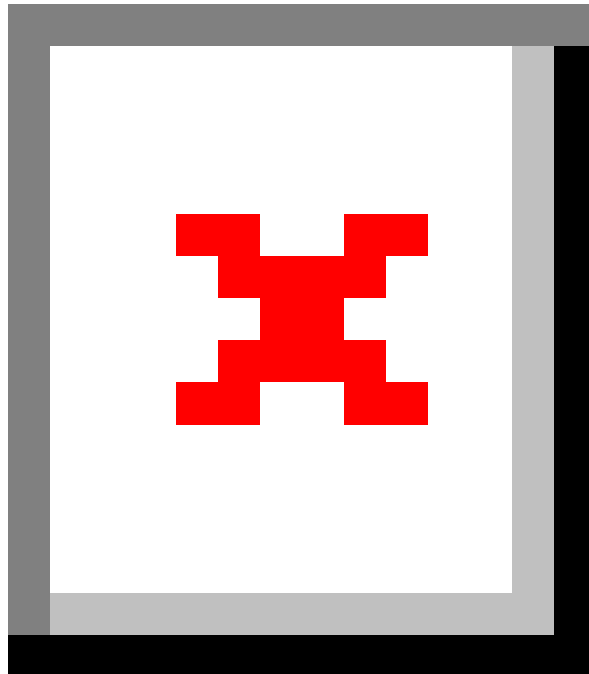
```

```
1
2
3
4 connectGrob(RTPCRCTNeg, CTInd, "N")
5
6
7 (DiagCTInd <- boxGrob("COVID-19 Positive n=16\n COVID-19 Negative n=15",
8   x=0.9, y=0.170, box_gp = gp, width = unit(2,"inch")))
9 connectGrob(CTInd, DiagCTInd, "vertical")
10
11
12
13
14
15
16
17
18
19
```



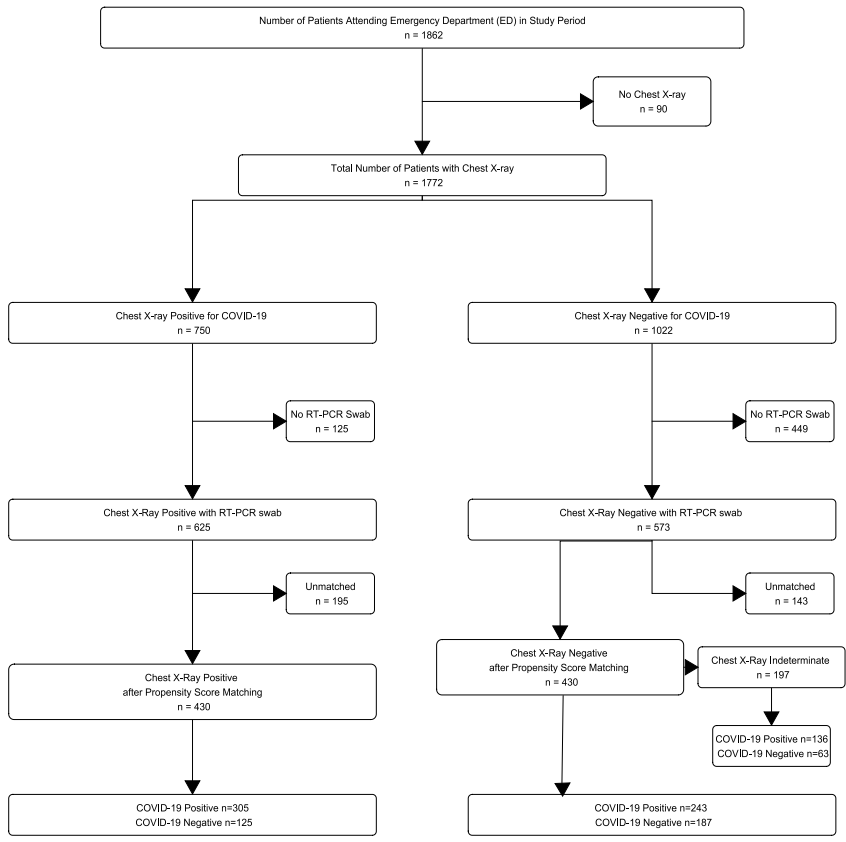
```
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55 ###Labels####
56
57
58
59
60
```

```
1
2
3 grid.newpage()
4 (indextest <- boxGrob("Index Tests",
5   x=midx, y=.9, box_gp = gpar(fill="light blue"), width = 0.7))
6
7
8 (reftest <- boxGrob("Index Tests and Reference Standards",
9   x=midx, y=.4, box_gp = gpar(fill="light blue"), width = 0.7))
10
11
12 (finaldiag <- boxGrob("Final Diagnoses",
13   x=midx, y=0.1, box_gp = gpar(fill="light blue"), width = 0.7))
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
```



Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	5
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
<i>Participants</i>	6	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	5,20
	11	Rationale for choosing the reference standard (if alternatives exist)	N/A
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	5
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	20
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	12
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	6,7
	15	How indeterminate index test or reference standard results were handled	5
	16	How missing data on the index test and reference standard were handled	N/A, excluded
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	N/A
	18	Intended sample size and how it was determined	7
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	22, diagram below
	20	Baseline demographic and clinical characteristics of participants	21
	21a	Distribution of severity of disease in those with the target condition	21
	21b	Distribution of alternative diagnoses in those without the target condition	N/A
	22	Time interval and any clinical interventions between index test and reference standard	N/A
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	22
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	22
	25	Any adverse events from performing the index test or the reference standard	N/A
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12
	27	Implications for practice, including the intended use and clinical role of the index test	14
OTHER INFORMATION			
	28	Registration number and name of registry	N/A
	29	Where the full study protocol can be accessed	N/A
	30	Sources of funding and other support; role of funders	N/A

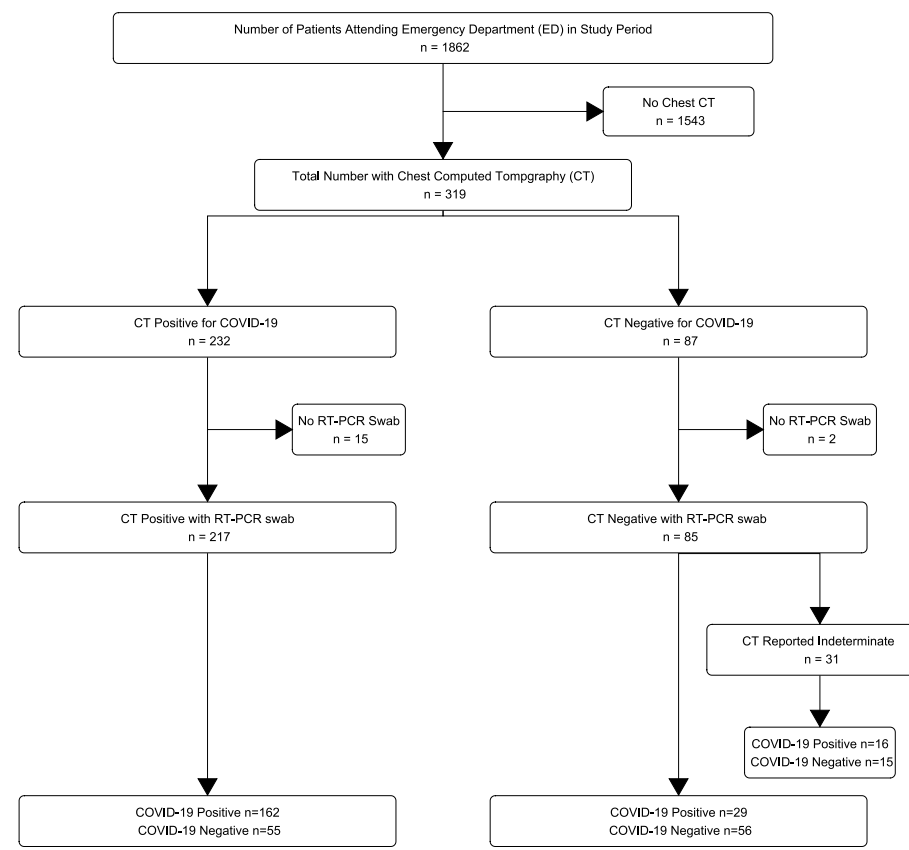
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47



Index Tests

Index Tests and Reference Standards

Final Diagnoses



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Supplementary Figure- STARD Flow Diagram

For peer review only

BMJ Open

Diagnostic Accuracy of X-ray versus CT in COVID-19: A Propensity Matched Database Study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-042946.R1
Article Type:	Original research
Date Submitted by the Author:	13-Sep-2020
Complete List of Authors:	Borakati, Aditya; Royal Free Hospital, Emergency Department; University College London, Division of Surgery and Interventional Science Perera, Adrian; Royal Free Hospital, Emergency Department Johnson, James; Royal Free Hospital, Emergency Department Sood, Tara; Royal Free Hospital, Emergency Department
Primary Subject Heading:	Emergency medicine
Secondary Subject Heading:	Radiology and imaging, Medical management, Infectious diseases, Respiratory medicine, Diagnostics
Keywords:	COVID-19, Chest imaging < RADIOLOGY & IMAGING, ACCIDENT & EMERGENCY MEDICINE, GENERAL MEDICINE (see Internal Medicine), Diagnostic radiology < RADIOLOGY & IMAGING

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Chest X-Ray Has Poor Sensitivity and Prognostic Significance in COVID-19: A Propensity Matched Database Study

Authors: Aditya Borakati (0000-0003-0457-4944), Adrian Perera (0000-0002-1753-2586), James Johnson (0000-0001-7796-659X), Tara Sood (0000-0003-1620-916X)

Emergency Department, Royal Free Hospital, Pond Street, London, UK NW3 2QG

Aditya Borakati, Academic Foundation Doctor, Division of Surgery and Interventional Science, Royal Free Campus, University College London

Adrian Perera, Emergency Medicine and Intensive Care Medicine Specialist Registrar, Emergency Department, Royal Free Hospital

James Johnson, Foundation Doctor, Emergency Department, Royal Free Hospital

Tara Sood, Consultant in Emergency Medicine, Emergency Department, Royal Free Hospital

Correspondence:

Dr Aditya Borakati

Academic Foundation Doctor

Emergency Department, Royal Free Hospital, London, UK NW3 2QG

a.borakati@doctors.org.uk

Author contribution (CRediT) statement:

Aditya Borakati: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project Administration

Adrian Perera: Conceptualization, Methodology, Investigation, Writing- Review & Editing, Supervision, Project Administration

James Johnson: Investigation

Tara Sood: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project Administration

Aditya Borakati is the overall guarantor of this work.

Word count: 4236

Abstract

Objectives: To identify the diagnostic accuracy of common imaging modalities, chest X-ray (CXR) and computed tomography (CT) for diagnosis of COVID-19 in the general emergency population in the UK and to find the association between imaging features and outcomes in these patients.

Design: Retrospective analysis of electronic patient records

Setting: Tertiary academic health science centre and designated centre for high consequence infectious diseases in London, UK.

Participants: 1,198 patients who attended the emergency department with paired RT-PCR swabs for SARS-CoV 2 and CXR between 16th March and 16th April 2020

Main outcome measures: Sensitivity and specificity of CXR and CT for diagnosis of COVID-19 using the British Society of Thoracic Imaging reporting templates. Reference standard was any reverse transcriptase polymerase chain reaction (RT-PCR) positive naso-oropharyngeal swab within 30 days of attendance. Odds ratios of CXR in association with vital signs, laboratory values and 30-day outcomes were calculated.

Results: Sensitivity and specificity of CXR for COVID-19 diagnosis were 0.56 (95% CI 0.51-0.60) and 0.60 (95% CI 0.54-0.65), respectively. For CT scans these were 0.85 (95% CI 0.79-0.90) and 0.50 (95% CI 0.41-0.60), respectively. This gave a statistically significant mean increase in sensitivity with CT compared with CXR, of 29% (95% CI 19%-38%, $p < 0.0001$). Specificity was not significantly different between the two modalities.

Chest X-ray findings were not statistically significantly or clinically meaningfully associated with vital signs, laboratory parameters or 30-day outcomes.

Conclusions: Computed tomography has substantially improved diagnostic performance over CXR in COVID-19. CT should be strongly considered in the initial assessment for suspected COVID-19. This gives potential for increased sensitivity and considerably faster turnaround time, where capacity allows and balanced against excess radiation exposure risk.

Key words: X-Rays, Computed Tomography, COVID-19, severe acute respiratory syndrome coronavirus 2, Emergency Medicine, Diagnostic Imaging

Statistical review: The statistical methods in this manuscript and associated code have been reviewed by Dr Federico Ricciardi of the Department of Statistical Science at University College London and confirmed as robust and accurate.

Ethical approval: This study was registered with the local institutional review board as a service evaluation using anonymised data only. No formal ethics committee review was required.

1
2
3 **Declarations of Interests:** The authors have no relevant conflicts of interest to declare.
4 All authors have completed the [Unified Competing Interest form](#) (available on request
5 from the corresponding author) and declare: no support from any organisation for the
6 submitted work; no financial relationships with any organisations that might have an
7 interest in the submitted work in the previous three years, no other relationships or
8 activities that could appear to have influenced the submitted work.
9
10

11 **Transparency declaration:** The lead author (AB) affirms that this manuscript is an
12 honest, accurate, and transparent account of the study being reported; that no important
13 aspects of the study have been omitted; and that any discrepancies from the study as
14 planned (and, if relevant, registered) have been explained.
15
16

17 **Funding:** No funding was received for this study.
18

19 **Copyright:** The Corresponding Author has the right to grant on behalf of all authors and
20 does grant on behalf of all authors, a worldwide licence to the Publishers and its
21 licensees in perpetuity, in all forms, formats and media (whether known now or created
22 in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii)
23 translate the Contribution into other languages, create adaptations, reprints, include
24 within collections and create summaries, extracts and/or, abstracts of the Contribution,
25 iii) create any other derivative work(s) based on the Contribution, iv) to exploit all
26 subsidiary rights in the Contribution, v) the inclusion of electronic links from the
27 Contribution to third party material where-ever it may be located; and, vi) licence any
28 third party to do any or all of the above
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

SARS-CoV 2 and its resulting disease, COVID-19, have propagated exponentially worldwide, with over 10 million cases in 188 countries at the time of writing [1,2].

The gold standard for diagnosis of the virus is the detection of viral RNA through reverse transcriptase polymerase chain reaction (RT-PCR) of respiratory tract samples. However, this method has several limitations including: (1) low sensitivity at 59-71% [3,4], (2) relatively slow turnaround times ranging from a few hours to several days [5], (3) high expense and (4) limited capacity for testing in many countries.

Computed tomography (CT) has been shown to be more sensitive than RT-PCR for diagnosis of COVID-19 [3,4], while being significantly faster and cheaper. This comes with a large radiation dose and capacity is still lacking in many countries.

Plain film chest X-ray (CXR) is ubiquitous worldwide, with a 30-70x lower dose of radiation[6] and is commonly performed as an initial investigation in COVID-19.

Studies have so far only evaluated imaging in those with confirmed infection, it is therefore, not possible to calculate the specificity of these modalities. In the context of the global pandemic, infection may be widespread in the community, often with subclinical infection [7,8]. A reliable and rapid method to detect infection in the general population, who may present to medical personnel with other complaints, is needed.

Despite its extensive use, the specificity and sensitivity of CXR in the general emergency population for diagnosis of COVID-19 is unknown, nor how imaging features correlate with severity.

This study evaluated the performance of CXR in diagnosing COVID-19 in the emergency department (ED) of a tertiary care hospital.

Methods

This study was conducted at the Royal Free Hospital, London, UK, an academic health science centre and nationally designated centre for High Consequence Infectious Diseases [9].

All individuals attending the emergency department who had paired posterior-anterior chest radiographs and RT-PCR nasopharyngeal swabs for COVID-19 at the time of initial attendance between 16th March 2020 and 16th April 2020 were included.

All chest radiographs were reported by a Consultant Radiologist and rated on an ordinal scale for probability of COVID-19: Alternative pathology identified, not COVID-19; Clear chest, unlikely COVID; Indeterminate findings for COVID-19; Classical findings of COVID-19, based on the British Society of Thoracic Imaging's (BSTI) reporting templates (table 1) [10]. These were reported prior to RT-PCR results being available.

RT-PCR of swabs were performed in laboratories either at our centre or at a public health laboratory (PHE Collindale, UK), according to published national standard operating procedures [11]. Subsequent RT-PCR swabs taken within 30 days of initial ED attendance were also included.

CT scans performed within 30 days of attendance were retrieved. These were also reported according to the BSTI template. CT pulmonary angiogram was performed in the ED if the D-dimer was >5000 to exclude pulmonary emboli as per the locally agreed protocol. Subsequent CT chest imaging (whether pulmonary angiogram, contrast or non-contrast) was performed on the basis of clinical suspicion.

Prospectively recorded data was extracted from the Cerner Millennium electronic patient record system (Cerner Corp., Kansas City, MO).

Primary Outcome

The primary outcome is sensitivity and specificity of initial CXR, where it is reported as having classic COVID-19 features in the ED. This is compared with RT-PCR swab as the reference standard for diagnosis of COVID-19.

In the event of multiple RT-PCR swabs during one attendance, a single positive swab was taken as an overall positive test during one admission.

Secondary Outcomes

In those patients who also had CT scans of the thorax, the diagnostic accuracy was compared with CXR, with RT-PCR again as the reference standard. Sensitivity and specificity of CXR when X-rays reported as indeterminate or atypical for COVID-19 were classed as positive was also calculated.

Chest x-ray findings were correlated with vital signs at attendance and blood results, including: neutrophil counts, D-dimer and C-reactive protein, which have been associated with poor prognosis in COVID-19 [12]. Hazard ratios for clinical outcomes including direct admission to the intensive treatment unit (ITU) from ED and 30-day mortality rates were also calculated for CXR reporting categories.

Statistical Analysis

In the event of missing data, multiple imputation was conducted using a Predictive Mean Matching algorithm, via the MICE R package, as described previously [13]. Briefly, this uses a linear regression model (or logistic regression model for categorical data), to find a random value based on already observed data, to replace missing fields [14]. Variables without missing data fields were not modified. The number of imputed datasets was similar in number to the percentage of missing data as suggested by White and colleagues [15]. Balance diagnostics with density plots are available in supplementary file 1, adequate balance was assessed via visual inspection of imputed distributions with respect to the original dataset.

The propensity for a CXR being reported as positive or negative for COVID-19 was calculated for several plausible covariates that may influence image characteristics such as Age, Gender, Ethnicity, pre-existing morbidities and the respiratory rate of the patient using a generalised linear model [16]. X-ray positive and negative groups were then matched in each imputed dataset using the nearest neighbour algorithm, with a calliper of 0.2 of the propensity score standard deviation, without replacement and in random sequential order to obtain a 1:1 match as described elsewhere [17].

The balance of the match data was assessed quantitatively with mean differences of covariates in each of the X-ray groups pre- and post-matching, with a difference of less than 0.1% considered a good match (supplementary tables 1-3). Visual inspection of matches was also conducted to ensure balance (supplementary figures 1-4).

After matching, outcome data were adjusted for covariates including age, gender, ethnicity and presence of co-morbidities as well as C-reactive protein, D-dimer, troponin and vital signs. This was achieved by generalised linear regression for continuous outcome data, binomial logistic regression for binary categorical outcomes, or ordinal logistic regression in the case of CXR where it is the outcome variable.

These regression models were run on each imputed dataset and outcomes were pooled together across each imputed data set according to Rubin's rules [18] to give an overall estimate.

Diagnostic Accuracy Statistics

Chest X-rays reported as classical for COVID-19 as per the BSTI guidelines were considered a positive test in the primary analysis. In a secondary analysis X-rays reported as 'Indeterminate' or 'Atypical' for COVID-19 were also considered positive. All other reports were classified as a negative test. These were compared to nasopharyngeal aspirate RT-PCR results, which were taken as the gold standard for diagnosis of COVID-19. Where more than one swab was taken during the study period (up to 30 days after initial attendance), a single positive result was taken as a positive result for calculation of diagnostic accuracy statistics.

Sensitivity, specificity, predictive values and diagnostic accuracy were calculated using the propensity matched data after imputation and pooled across imputed datasets with 95% confidence intervals. Apparent and true prevalence based on this dataset are also given for interpretation of the predictive values.

Chest CTs were also reported according to the BSTI guidelines as with X-ray. Diagnostic statistics were calculated on raw, unmatched and non-imputed data (due to a low volume of

1
2
3 data for imputation and matching) in the same manner as X-ray. Mean differences and 95%
4 confidence intervals between CT and X-ray for each of the diagnostic statistics are given, with a
5 p-value calculated from the confidence intervals.
6

7 Agreement between the modalities was assessed on the unmatched dataset, in the sample
8 where CT, CXR and RT-PCR were all available using Cohen's (for two group agreement) and Fleiss'
9 Kappa (when all 3 are compared).
10

11 **Data Presentation**

12 Descriptive statistics are given as means and standard deviations for normally distributed data
13 and as medians and interquartile ranges for non-normally distributed data, before and after
14 matching and multiple imputation (for the latter these statistics are pooled across imputations).
15

16 Association of explanatory variables with SARS-CoV 2 and Chest X-ray findings are given as
17 odds ratios in uni- and multi-variate configurations.
18

19 Data was considered statistically significant if $p < 0.05$. Given the large number of analyses in
20 this paper, data is separately highlighted if $p < 0.001$ as a secondary threshold to address the
21 potential for false positives with multiple testing.
22

23 Analyses were conducted using R 4.0.0 (R Foundation for Statistical Computing, Vienna,
24 Austria) and code for the analyses is given in supplementary file 2.
25

26 **Sample size calculation**

27 In this study, the lower confidence interval for sensitivity of CXR as reported by Wong et al.[19]
28 (56%) was used as an estimate of likely sensitivity for COVID-19 . A power of 80% at an alpha
29 of 0.05 was used to calculate the sample size for sensitivities and specificities of 56%. This
30 gave an estimated sample size of 165 in each of the COVID-19 negative and positive groups by
31 RT-PCR (total 330).
32

33 **Ethical approval**

34 This study was registered with the local institutional review board as a service evaluation using
35 anonymised data only. No formal ethics committee review was required.
36

37 **Reporting Guidelines**

38 This study is reported according to the STARD guidelines [20] for diagnostic accuracy studies.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Results

1,198 eligible patients with both CXR and RT-PCR were identified in the study period (figure 1). Their characteristics, stratified by positivity for SARS-CoV 2 infection by RT-PCR is summarized in table 2. This showed that those with confirmed SARS-CoV 2 infection were more likely to be male, older (mean age 66.2 vs 62.7), have lower saturations, higher respiratory rates, whilst being more likely to be admitted and die within 30 days. There was a significant association with X-ray images and SARS-CoV 2 at baseline, with 59.6% having classic imaging features of COVID-19 in those with positive swabs versus 39.1% in those with negative swabs. There was 8.6% missing data overall in the dataset when variables with >50% missing data were removed and 15 imputations were performed on these remaining variables only.

After multiple imputation for missing data and pooled propensity score matching for plausible covariates that may affect CXR reporting, there were 430 patients in each of the X-ray positive and X-ray negative groups, for a total of 860 patients. Adequate balance was achieved for relevant covariates with a mean difference of <0.1 between groups (supplementary table 2).

Computed tomography (CT) was performed in 302 patients with paired RT-PCR during the same time period, with a median serial interval of 4.5 days (inter quartile range 0-17) after the initial attendance in ED and of these 30.1% were within one day of attendance.

Diagnostic Accuracy

The pooled sensitivity and specificity of CXR was 0.56 (95% CI 0.51-0.60) and 0.60 (95% CI 0.54-0.65), respectively (table 3). This gave an overall diagnostic accuracy of 0.57 (95% CI 0.54-0.61) for CXR.

In comparison, sensitivity and specificity for CT was 0.85 (95% CI 0.79-0.90) and 0.50 (95% CI 0.41-0.60), respectively. This gave a statistically significant mean increase in sensitivity with CT compared with CXR by 29% (95% CI 19%-38%, $p<0.0001$). Specificity was not significantly different between the two modalities. Diagnostic accuracy and negative predictive values were also significantly increased with CT at 0.15 and 0.22, respectively, while the negative likelihood ratio was significantly decreased at -0.44. This shows that the post-test odds of being negative for SARS-CoV 2 by RT-PCR with a negative CT is significantly lower.

Taking X-rays reported as indeterminate as positive increased the sensitivity of CXR to 0.80 (95% CI 0.77-0.84), however reduced specificity to 0.40 (95% CI 0.35-0.46). When CT scans reported as indeterminate are also considered positive the sensitivity of CT increased to 0.93 (95% CI 0.89-0.96), whilst mean specificity reduced to 0.37 (95% CI 0.28-0.47), although this was not statistically different from when indeterminate CTs are considered negative. Sensitivity of CT remained significantly higher than CXR (when indeterminates are considered positive for both) by 0.13 (95% CI 0.05-0.19, $p<0.001$), specificity was not significantly different between the two.

When comparing only the unimputed, unmatched subset of data where CT, RT-PCR and CXR were all performed ($n=287$), the agreement between CT and CXR was poor (Cohen's kappa 0.406). Agreement between all three modalities was also poor (Fleiss' kappa 0.361).

Association of CXR with Markers of Severity and Outcomes

Association of covariates with RT-PCR results is shown in table 4 and figure 2. Those who tested positive for SARS-CoV 2 by RT-PCR were significantly more likely to have a classical X-ray (OR 1.79 95% CI 1.25-2.56, $p < 0.002$) as would be expected by the diagnostic accuracy statistics (table 4). When the CXR report is considered as an ordered scale, worsening grades of report were associated more strongly with RT-PCR positivity, with a 1.94 x increase in odds for each grade.

Positive chest X-rays for COVID-19 were significantly associated with lower oxygen saturations (OR 0.94 95% CI 0.92-0.97, $p < 0.001$) and temperatures (2.30 95% CI 1.46-3.63, $p < 0.001$) in the ED following propensity score matching and multivariate regression (table 5 and figure 3).

They also had higher rates of admission to a general ward from the ED (OR 2.30 95% CI 1.46-3.63, $p < 0.001$) but no significant association with 30 day outcomes. There was a statistically significant increase in C-reactive protein with a positive X-ray, however, this is unlikely to be clinically meaningful due to the minimal association (OR 1.00 95% CI 1.00-1.01).

Discussion

This study is the first to report the diagnostic accuracy of CXR and CT in the general emergency population during the COVID-19 pandemic.

We show that CXR has poor sensitivity and specificity for diagnosis of COVID-19, whilst CT has 29% higher sensitivity. Many international radiological guidelines advise against CT scanning for the initial assessment of COVID-19 [21–23] or where there are equivocal CXRs, whilst in other countries CT scanning is performed as a routine first line investigation. Our results suggest that CT should be considered in the initial assessment of COVID-19 and that CXR findings poorly correlate with CT findings in this setting. We also show that indeterminate and non-classical features of COVID-19 significantly increase the sensitivity of these imaging modalities, without a significant decrease in specificity. Further, we demonstrate the limited prognostic value of CXR in COVID-19.

These findings mirror what has previously been reported in the literature on individuals with confirmed COVID-19. Wong et al. [19] showed a sensitivity of 59% for initial X-ray in confirmed COVID-19 infection, similarly initial case series in China also reported a sensitivity of 59.1% [12].

A recent in press article from Italy reported a much higher sensitivity of 89% for CXR in a smaller general emergency population (n=535) without confirmed COVID-19 at attendance [24]. However, this used telephone follow up for clinical symptoms of COVID-19 as a reference standard in individuals with an initial negative RT-PCR swab and appeared to classify any abnormal X-ray as positive, which may inflate this figure. When indeterminate CXRs are counted as positive in this study, the sensitivity would be in line with this Italian data. In the US, a study of patients attending an urgent care centre with confirmed COVID-19, showed a much lower sensitivity at 41.7% for CXR where any abnormality was found on the images [25]. In this study 97/636 reports were re-classified from 'possible pneumonia' to 'normal' on second reading from a radiologist, highlighting the importance of inter-rater agreement and possibly explaining this low estimate.

Computed tomography has been reported in previous studies as being up to 98% sensitive for the diagnosis of COVID-19 in confirmed patients, when RT-PCR is used as the reference standard in confirmed patients [3,4]. These studies used any potential features of COVID-19 (e.g. ground glass opacification, crazy paving) as a positive scan, regardless of spatial distribution or features more characteristic of alternate pathology, unlike the BSTI guidelines used in this study. When we classified indeterminate CTs as positive like these latter studies, our estimates match their sensitivity values.

Consequently, a much lower specificity of 25% was found with initial RT-PCR in previous literature; however, it is reported that 10 out of 15 (67%) of these negatives subsequently tested positive. This would give an adjusted specificity of 75%, considering subsequent swabs as a reference standard, which combined with the wider CIs in these smaller studies, would bring estimates in line with the specificity in this paper. More recent meta-analyses have placed the pooled sensitivity of CT in populations with confirmed COVID-19 only, at 89.76% (95% CI 84.42%-93.84%) [26], in line with the estimates identified here.

There is limited coverage in the literature on association of X-ray findings with clinical and laboratory parameters and outcomes in the COVID-19 pandemic. This study demonstrates that classic appearances of COVID-19 were associated with initial lower saturations and lower

1
2
3 temperature. Volume opacification of the lung fields were not quantified as a surrogate of
4 severity; however, the use of the BSTI grading templates does this somewhat. When the X-ray
5 report is considered as a graded scale from low likelihood of COVID-19 and severity to high
6 likelihood and severity of disease there was no significant difference in association with vital
7 signs or laboratory parameters compared with when the X-ray report is merely considered as a
8 binary positive and negative outcome for COVID-19.
9

10 Borghesi and colleagues have devised a X-ray grading system, the Brixia score, for severity in
11 admitted patients with confirmed SARS-CoV 2 infection [27]. They further found a significant
12 increase in the severity of CXR by this scoring system in those who were discharged versus
13 those who died [28,29].
14

15
16 Here, there were no relevant associations between CXR and laboratory values. This analysis
17 also found no association with positive X-rays and 30 day outcomes after multivariate analyses,
18 unlike Borghese et al. This is also in contrast to Guan et al. who found higher rates of ITU
19 admission and death in those with positive imaging findings. However, these studies analysed
20 only those with confirmed SARS-CoV 2 infection. The divergence observed in this study may be
21 due to classifying those with 'Alternate pathology/ Indeterminate' or 'CVXC3/ CVXC2' as per the
22 BSTI templates, negative for COVID-19 in these analyses. Other studies classified X-rays with
23 any abnormality as a positive for COVID-19. These alternate distributions may still be reflective
24 of underlying COVID-19 and we show significantly higher sensitivity for both CT and CXR when
25 these are classed as positive. It may be that correlating indeterminate X-rays (in addition to
26 classical images) with vitals, laboratory markers and 30 day outcomes would yield significant
27 associations. However this may be unlikely, Xu and Zhang et al. found that those with classical
28 bilateral and diffuse involvement in upper and lower lobes had more severe disease than those
29 without [30,31].
30
31

32 There were a total of 70 confirmed pulmonary emboli (PEs) in our dataset out of 114 CT
33 pulmonary angiograms (61.0%, 5.84% of all patients attending) performed in the emergency
34 department. The incidence of venous thromboembolism is reported as ranging from 20-30% in
35 admitted confirmed SARS-CoV 2 positive patients [32]. Although we have not focused on this
36 cohort of patients in this paper for the sake of brevity and simplicity, this high incidence
37 represents a further advantage for CT over CXR.
38
39

40 CT, even with the absence of contrast has been shown to have strong accuracy in the diagnosis
41 of pulmonary emboli and many imaging features correlate with the presence of pulmonary
42 emboli. Sensitivities of non-contrast CT for diagnosis of PE have been reported at 96.9% and
43 specificity at 71.9% [33,34].
44

45 We therefore see the advantages of CT scanning in COVID-19 as threefold over other
46 diagnostic techniques: 1) The rapid turnaround; 2) Increased sensitivity and 3) The possibility to
47 identify pulmonary emboli in COVID-19, which are a significant burden in this group.
48

49 This must be balanced against the excess radiation exposure with CT. Radiation from CT and
50 its association with carcinogenesis is difficult to quantify and no definitive epidemiological
51 studies have confirmed excess risk of cancer[35]. Modern CT scanners and software
52 reconstruction techniques continue to minimise radiation exposure and many ways of shielding
53 parts of the body from radiation also exist. Nevertheless, the excess risk of lifetime cancer is
54 estimated at 1 per 5,000 CT examinations[36].
55
56
57

Strengths and Limitations

This study is the largest conducted on imaging in the COVID-19 pandemic and one of the only studies conducted in the general population during the pandemic rather than only in confirmed patients. This enables greater applicability to the clinical setting where the diagnosis is uncertain, in addition to being able to calculate specificity, which is not possible in most studies. This study was planned to be powered to detect a sensitivity and specificity of 56% for CXR and greatly exceeded the sample size necessary for this.

Comprehensive statistical analyses were conducted to account for confounders in both factors influencing reporting of CXR and in factors affecting outcomes. The data was collected from prospectively maintained electronic records; however, the retrieval took place retrospectively with its inherent disadvantages. We were not able to collect data on several relevant covariates such as specific comorbidities or markers of severity such as lymphocytes. Furthermore, there was a significant amount of missing data that required multiple imputation to replace, although the fit of this imputed data was good, actual, observed data would be ideal.

Inter-rater reliability of imaging reports was not analysed in this paper and there was the potential for individual radiologists to have greater or lesser accuracy in the diagnosis of COVID-19. The literature has so far suggested a strong degree of agreement between radiologists in reporting of COVID-19 images [28].

The single centre nature of this study further limits generalisability and the potential for inter-hospital disagreement in imaging, in addition to inter-rater disagreement.

Finally, the median time for patients to receive a CT scan was 4.5 days following initial attendance to ED. Thus, the scans may not have been directly comparable to the initial CXR, both because of the progression of disease and because the SARS-CoV 2 status may have been confirmed at this point, biasing the reporting of these scans.

Future Research

Although this study used RT-PCR of nasopharyngeal swabs as a reference standard, newer methods exist for diagnosis of the disease. Serological assays for antibodies against SARS-CoV 2 are increasingly available and may represent a better gold standard in diagnosis for future research [37]. RT-PCR is limited by swabbing technique for nasopharyngeal samples and the fact that the virus is more avid in the lower respiratory tract [38]. However, many patients may not seroconvert prior to death limiting this test to survivors only.

Point of care lung ultrasound is a new technique for diagnosis of COVID-19 which may mitigate many of the issues noted with the modalities discussed so far. It has no radiation, is fast, cheap and may be able to detect lower respiratory tract disease unlike nasopharyngeal swab. However, there is limited evidence beyond small case series on its diagnostic accuracy [39–41]. Further, like other ultrasound techniques accuracy will likely be operator dependent [42] and experience will need to be built up for robust results in evaluating suspected COVID-19.

Finally, much research has been conducted in the use of artificial intelligence techniques to correctly diagnose COVID-19 based on imaging [43–45]. These techniques would obviate capacity limitations in reporting imaging as well as eliminate inter-reporter variability. However, as with any supervised machine learning technique, large, generalisable datasets, with correctly

1
2
3 pre-classified positive and negative cases (which in turn will depend on a truly accurate
4 reference standard) are needed [46].
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Conclusion

Chest X-ray has poor sensitivity and specificity in diagnosing COVID-19 in the general population during the pandemic. CT scanning has demonstrated excellent sensitivity and should strongly be considered during the pandemic in the initial assessment of COVID-19. This needs to be balanced against the risk of excess radiation with CT, where capacity allows.

Summary box

What is already known on this topic

-Small observational studies, predominantly in China, have reported on imaging features in COVID-19 after a confirmed RT-PCR swab test

-These studies have shown limited sensitivity for chest X-ray, but excellent sensitivity for CT scans, it is not possible to calculate the specificity of these modalities as they only included patients with confirmed COVID-19, therefore it is not possible to assess their utility in the general population who may or may not have COVID-19

-Literature on this general population attending emergency departments and the accuracy of these imaging techniques is limited

-International guidelines including from the British Society of Thoracic Imaging and American College of Radiology do not recommend the use of CT in initial evaluation of suspected COVID-19, largely due to capacity concerns

What this study adds

-This study shows that Chest x-ray has poor sensitivity and specificity in patients with suspected COVID-19 attending the emergency department, whilst CT has excellent sensitivity and is 29% more sensitive than CXR in our study cohort; there was also poor agreement between CT and CXR findings in COVID-19

-Patients with indeterminate imaging without classical distribution of COVID-19 should still be considered at high risk of having the disease

-Our data suggest that CT should be employed more widely as an initial investigation, where capacity allows and balanced against the risk of excess radiation exposure

Acknowledgements

We would like to thank Scott Wilson from the Royal Free Hospital's clinical practice group analytics department for retrieving the data from the hospital's data warehouse.

We would like to thank Dr Federico Ricciardi of the Department of Statistical Science and PRIMENT Clinical Trials Unit at University College London for reviewing the statistical methods in this study.

1
2
3 **Data availability**

4 Anonymised data is available on reasonable request from the corresponding author. Analysis
5 scripts are attached as a supplementary file.
6

7
8 **Declarations of Interest**

9 The authors declare no conflicts of interest.
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

References

- 1 COVID-19 Map. Johns Hopkins Coronavirus Resour. Cent. <https://coronavirus.jhu.edu/map.html> (accessed 30 Jun 2020).
- 2 Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020;**0**. doi:10/ggnsjk
- 3 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**. doi:10/ggmw6p
- 4 Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* 2020;**200432**. doi:10/ggnnkj
- 5 Konrad R, Eberle U, Dangel A, et al. Rapid establishment of laboratory diagnostics for the novel coronavirus SARS-CoV-2 in Bavaria, Germany, February 2020. *Eurosurveillance* 2020;**25**. doi:10/ggp6bw
- 6 Lin EC. Radiation Risk From Medical Imaging. *Mayo Clin Proc* 2010;**85**:1142–6. doi:10/c445mk
- 7 Mizumoto K, Kagaya K, Zarebski A, et al. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance* 2020;**25**:2000180. doi:10/ggn4bd
- 8 Arons MM, Hatfield KM, Reddy SC, et al. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. *N Engl J Med* 2020;**382**:2081–90. doi:10/ggszfg
- 9 High consequence infectious diseases (HCID). GOV.UK. <https://www.gov.uk/guidance/high-consequence-infectious-diseases-hcid> (accessed 24 May 2020).
- 10 Desai S. COVID-19 BSTI Reporting templates | The British Society of Thoracic Imaging. *Br. Soc. Thorac. Imaging*. 2020. <https://www.bsti.org.uk/covid-19-resources/covid-19-bsti-reporting-templates/> (accessed 29 Apr 2020).
- 11 NHS England. Guidance and Standard Operating Procedure: COVID-19 virus testing in NHS Laboratories. 2020. <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 24 May 2020).
- 12 Guan W, Ni Z, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020;**0**:null. doi:10/ggm6dh
- 13 Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. *J Stat Softw* 2011;**45**. doi:10/gdq9c
- 14 Ginkel JR van, Linting M, Rippe RCA, et al. Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data. *J Pers Assess* 2020;**102**:297–308. doi:10/gftj5w

- 1
2
3 15 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and
4 guidance for practice. *Stat Med* 2011;**30**:377–99. doi:10.1002/sim.4067
5
6 16 He H, McDermott MP. A robust method using propensity score stratification for correcting
7 verification bias for binary tests. *Biostat Oxf Engl* 2012;**13**:32–47. doi:10/c4jzn6
8
9 17 Ho DE, Imai K, King G, et al. MatchIt : Nonparametric Preprocessing for Parametric Causal
10 Inference. *J Stat Softw* 2011;**42**. doi:10/gdwtnq
11
12 18 Marshall A, Altman DG, Holder RL, et al. Combining estimates of interest in prognostic
13 modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res*
14 *Methodol* 2009;**9**:57. doi:10.1186/1471-2288-9-57
15
16 19 Wong HYF, Lam HYS, Fong AH-T, et al. Frequency and Distribution of Chest Radiographic
17 Findings in COVID-19 Positive Patients. *Radiology* 2020;:201160.
18 doi:10/ggqbp4
19
20 20 Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic
21 accuracy studies: explanation and elaboration. *BMJ Open* 2016;**6**:e012799.
22 doi:10.1136/bmjopen-2016-012799
23
24 21 Rubin GD, Ryerson CJ, Haramati LB, et al. The Role of Chest Imaging in Patient
25 Management during the COVID-19 Pandemic: A Multinational Consensus
26 Statement from the Fleischner Society. *Radiology* 2020;:201365.
27 doi:10/ggrmg4
28
29 22 ACR Recommendations for the use of Chest Radiography and Computed Tomography (CT)
30 for Suspected COVID-19 Infection. [https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-](https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection)
31 [COVID19-Infection](https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection) (accessed 5 Jun 2020).
32
33 23 British Society of Thoracic Imaging. COVID-19: BSTI STATEMENT AND GUIDANCE.
34 2020;:1. https://www.bsti.org.uk/media/resources/files/COVID11.3.20_2.pdf (accessed 5 Jun
35 2020).
36
37 24 Schiaffino S, Tritella S, Cozzi A, et al. Diagnostic Performance of Chest X-Ray for COVID-
38 19 Pneumonia During the SARS-CoV-2 Pandemic in Lombardy, Italy. *J Thorac Imaging*
39 2020;**Publish Ahead of Print**. doi:10/ggx268
40
41 25 Weinstock MB, Echenique A, Russell JW, et al. Chest X-Ray Findings in 636 Ambulatory
42 Patients with COVID-19 Presenting to an Urgent Care Center: A Normal Chest X-Ray Is no
43 Guarantee. ;:10.
44
45 26 Bao C, Liu X, Zhang H, et al. Coronavirus Disease 2019 (COVID-19) CT Findings: A
46 Systematic Review and Meta-analysis. *J Am Coll Radiol* 2020;**17**:701–9. doi:10/ggr28p
47
48 27 Borghesi A, Zigliani A, Masciullo R, et al. Radiographic severity index in COVID-19
49 pneumonia: relationship to age and sex in 783 Italian patients. *Radiol Med (Torino)*
50 2020;**125**:461–4. doi:10/ggtvwp
51
52
53
54
55
56
57
58
59
60

- 1
2
3 28 Borghesi A, Maroldi R. COVID-19 outbreak in Italy: experimental chest X-ray scoring system
4 for quantifying and monitoring disease progression. *Radiol Med (Torino)* 2020;**125**:509–13.
5 doi:10/ggtvwn
6
- 7 29 Borghesi A, Zigliani A, Golemi S, et al. Chest X-ray severity index as a predictor of in-
8 hospital mortality in coronavirus disease 2019: A study of 302 patients from Italy. *Int J Infect*
9 *Dis* 2020;**96**:291–3. doi:10.1016/j.ijid.2020.05.021
10
- 11 30 Xu Y-H, Dong J-H, An W-M, et al. Clinical and computed tomographic imaging features of
12 novel coronavirus pneumonia caused by SARS-CoV-2. *J Infect* 2020;**80**:394–400.
13 doi:10/ggqwf3
14
- 15 31 Zhang J-J, Dong X, Cao Y-Y, et al. Clinical characteristics of 140 patients infected with
16 SARS-CoV-2 in Wuhan, China. *Allergy Published Online First*: 19 February 2020.
17 doi:10/ggpx6g
18
- 19 32 Lodigiani C, Iapichino G, Carenzo L, et al. Venous and arterial thromboembolic
20 complications in COVID-19 patients admitted to an academic hospital in Milan, Italy. *Thromb*
21 *Res* 2020;**191**:9–14. doi:10/ggvcft
22
- 23 33 Chien C-H, Shih F-C, Chen C-Y, et al. Unenhanced multidetector computed tomography
24 findings in acute central pulmonary embolism. *BMC Med Imaging* 2019;**19**:65.
25 doi:10/ggzg85
26
- 27 34 Mohamed N, Othman MoustafaHM, Hassan L, et al. The accuracy of non-contrast chest
28 computed tomographic Scan in the detection of pulmonary thromboembolism. *J Curr Med*
29 *Res Pract* 2019;**4**:61. doi:10/ggzg83
30
- 31 35 McCollough CH, Bushberg JT, Fletcher JG, et al. Answers to Common Questions About the
32 Use and Safety of CT Scans. *Mayo Clin Proc* 2015;**90**:1380–92. doi:10/f3jgqx
33
- 34 36 Moser JB, Sheard SL, Edyvean S, et al. Radiation dose-reduction strategies in thoracic CT.
35 *Clin Radiol* 2017;**72**:407–20. doi:10/f95q7p
36
- 37 37 Long Q-X, Liu B-Z, Deng H-J, et al. Antibody responses to SARS-CoV-2 in patients with
38 COVID-19. *Nat Med* 2020;:1–4. doi:10.1038/s41591-020-0897-1
39
- 40 38 Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in Different Types of Clinical
41 Specimens. *JAMA* 2020;**323**:1843–4. doi:10/ggpp6h
42
- 43 39 Smith MJ, Hayward SA, Innes SM, et al. Point-of-care lung ultrasound in patients with
44 COVID-19 – a narrative review. *Anaesthesia*;n/a. doi:10/ggr2p7
45
- 46 40 Haaksma ME, Heldeweg MLA, Matta JEL, et al. Lung ultrasound findings in patients with
47 novel SARS-CoV2. *medRxiv* 2020;:2020.05.18.20105775.
48 doi:10.1101/2020.05.18.20105775
49
- 50 41 Benchoufi M, Bokobza J, Chauvin AA, et al. Lung injury in patients with or suspected
51 COVID-19 : a comparison between lung ultrasound and chest CT-scanner severity
52 assessments, an observational study. *medRxiv* 2020;:2020.04.24.20069633.
53 doi:10.1101/2020.04.24.20069633
54
55
56
57
58
59
60

- 1
2
3 42 Fine D, Perring S, Herbetko J, et al. Three-dimensional (3D) ultrasound imaging of the
4 gallbladder and dilated biliary tree: reconstruction from real-time B-scans. *Br J Radiol*
5 1991;**64**:1056–7. doi:10/fqr9mh
6
7 43 Shi F, Wang J, Shi J, et al. Review of Artificial Intelligence Techniques in Imaging Data
8 Acquisition, Segmentation and Diagnosis for COVID-19. *IEEE Rev Biomed Eng* 2020;;1–1.
9 doi:10/ggs2km
10
11 44 Li L, Qin L, Xu Z, et al. Artificial Intelligence Distinguishes COVID-19 from Community
12 Acquired Pneumonia on Chest CT. *Radiology* 2020;;200905. doi:10/ggpdgp
13
14 45 Wang L, Wong A. COVID-Net: A Tailored Deep Convolutional Neural Network Design for
15 Detection of COVID-19 Cases from Chest X-Ray Images. *ArXiv200309871 Cs Eess*
16 Published Online First: 11 May 2020.<http://arxiv.org/abs/2003.09871> (accessed 13 Jun
17 2020).
18
19 46 Kotsiantis SB. Use of machine learning techniques for educational proposes: a decision
20 support system for forecasting students' grades. *Artif Intell Rev* 2012;**37**:331–44.
21 doi:10/fmbng4
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Tables

Ordinal scale for study	BSTI grade	Features on X-ray
0	CVCX3- Non-COVID-19	Alternative pathology such as pneumothorax with no features of COVID-19 identified
1	CVCX0- Normal	No pathology seen
2	CVCX2- Indeterminate for COVID-19 or atypical features	Poor quality film or central/ basal consolidation
3	CVCX1- Classic findings of COVID-19	Peripheral ground glass opacities

Table 1- Ordinal scale used in this study based on the British Society of Thoracic Imaging (BSTI) Reporting Template [12]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	SARS-CoV 2 RT-PCR		p-value	Missing (%)
	Negative	Positive		
n (%)	435 (36.3)	763 (63.7)		
Number of Swabs (%)	810 (48.3)	868 (51.7)		
Age (mean (SD))	62.74 (17.72)	66.18 (17.58)	0.001*	0
Ethnicity			0.097	19
Other- Asian (%)	29 (8.0)	72 (11.8)		
South- Asian (%)	27 (7.5)	38 (6.2)		
Black (%)	41 (11.4)	91 (14.9)		
Mixed (%)	6 (1.7)	6 (1.0)		
Other (%)	56 (15.5)	105 (17.2)		
White (%)	202 (56.0)	297 (48.8)		
Sex – Male (%)	233 (53.6)	480 (62.9)	0.002*	0
Oxygen Saturation (median (IQR))	95 (6)	93 (8)	<0.001**	6.3
Respiratory Rate (median (IQR))	22 (8)	26 (12)	<0.001**	6.3
Glasgow Coma Scale (median (IQR))	15 (0)	15 (0)	0.043*	6.6
Systolic BP (median (IQR))	134 (32)	130 (30)	0.009*	15.8
Heart Rate (median (IQR))	96 (27)	94 (27)	0.092	6.4
Temperature (median (IQR))	37.1 (1.4)	37.7 (1.4)	<0.001**	6.7
Chest X-ray report			<0.001**	0
Alternative pathology (%)	4 (0.9)	3 (0.4)		
No abnormalities (%)	178 (40.9)	136 (17.8)		
Indeterminate (%)	83 (19.1)	169 (22.1)		
Classic COVID-19 (%)	170 (39.1)	455 (59.6)		
Presence of comorbidities (%)	297 (79.0)	482 (80.3)	0.669	18.5
Dyspnoea (%)	274 (69.4)	497 (75.5)	0.034	12.1
Neutrophils (median (IQR))	6.42 (4.56)	5.25 (3.92)	<0.001**	2.3
D-Dimer (median (IQR))	1250 (2440)	1105 (1803)	0.204	23.2
Albumin (median (IQR))	39 (7)	37 (6)	<0.001**	10
C-Reactive Protein (median (IQR))	91.0 (115)	146.5 (264.8)	<0.001**	3
Creatine Kinase (median (IQR))	51 (104)	145 (260)	<0.001**	23.3
Troponin (median (IQR))	19 (46)	20 (44)	0.278	19.1
Admitted (%)	331 (76.0)	635 (83.2)	0.003*	0.1
Admitted to ITU (%)	5 (1.3)	32 (4.8)	0.005*	12.4
Thirty Day Follow Up Status			<0.001**	24
Discharged (%)	219 (78.2)	367 (58.3)		
On Ambulatory Follow Up (%)	14 (5.0)	49 (7.8)		
Admitted (%)	18 (6.4)	60 (9.5)		
Died (%)	29 (10.4)	154 (24.4)		
CT report			<0.001**	0
No pathology identified (%)	23 (22.1)	6 (3.3)		
Classic COVID-19 findings (%)	52 (50.0)	157 (85.8)		
Indeterminate for COVID-19 (%)	14 (13.5)	14 (7.7)		
Alternative pathology identified (%)	15 (14.4)	6 (3.3)		
Day of Symptoms (mean (SD))	9.84 (9.63)	8.56 (15.80)	0.368	69.2

1
2
3 **Table 2-** Baseline characteristics of dataset stratified by overall SARS-CoV 2 RT-PCR status, including
4 subsequent swabs during the study period- NB there were 480 additional swabs on 399 unique patients
5 with a median of 2 and mean of 3.5 per patient; *significant at $p < 0.05$; **significant at $p < 0.001$
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

	Chest X-ray	CT Chest	Mean Difference	p-value
Total (n)	860	302		
True Positives (n)	305	162	-	-
False Positives (n)	125	55	-	-
True Negatives (n)	187	56	-	-
False Negatives (n)	243	29	-	-
Apparent prevalence (95% CI)	0.50 (0.47-0.53)	0.72 (0.66-0.77)	0.22 (0.04-0.21)	<0.0001**
True prevalence (95% CI)	0.64 (0.60-0.67)	0.63 (0.58-0.69)	-0.00 (-0.09-0.03)	0.111
Sensitivity (95% CI)	0.56 (0.51-0.60)	0.85 (0.79-0.90)	0.29 (0.19-0.38)	<0.0001**
Specificity (95% CI)	0.60 (0.54-0.65)	0.50 (0.41-0.60)	-0.10 (-0.25-0.04)	0.119
Positive Predictive Value (95% CI)	0.71 (0.66-0.75)	0.75 (0.68-0.80)	0.04 (-0.06-0.14)	0.492
Negative Predictive Value (95% CI)	0.43 (0.39-0.48)	0.66 (0.55-0.76)	0.22 (0.06-0.37)	0.005*
Positive Likelihood Ratio (95% CI)	1.39 (1.19-1.62)	1.71 (1.41- 2.08)	0.32 (-0.22-0.89)	0.258
Negative Likelihood Ratio (95% CI)	0.74 (0.64-0.84)	0.30 (0.21-0.44)	-0.44 (-0.64 - -0.21)	0.022*
Diagnostic Accuracy (95% CI)	0.57 (0.54-0.61)	0.72 (0.66-0.77)	0.15 (0.06-0.23)	<0.0001**

Table 3- Diagnostic Accuracy Metrics for CXR and CT Chest with RT-PCR for SARS-CoV 2, as the reference standard; *significant difference at the <0.05 level; **significant difference at the <0.0001 level

		SARS-CoV 2 RT-PCR		OR (univariable)	OR (multivariable)	
		Negative	Positive			
1						
2						
3						
4	n	312	548			
5						
6	Chest X-ray report	Alternative pathology (%)	3 (0.8)	3 (0.5)	-	-
7		No abnormalities (%)	123 (39.6)	104 (19.1)	0.76 (0.08-6.82, p=0.801)	0.48 (0.03-8.82, p=0.620)
8		Indeterminate/ atypical findings (%)	61 (19.5)	136 (4.8)	1.99 (0.22-17.81, p=0.535)	0.92 (0.05-16.88, p=0.952)
9						
10		Classic COVID (%)	125 (40.1)	305 (55.6)	2.17 (0.24-19.19, p=0.484)	1.14 (0.06-20.98, p=0.927)
11	Age	Mean (SD)	61.8 (17.9)	67.0 (17.7)	1.02 (1.01-1.02, p<0.001)**	1.02 (1.00-1.03, p=0.028)*
12	Sex	Female (%)	138 (44.3)	212 (38.7)	-	-
13		Male (%)	174 (55.7)	336 (61.3)	1.26 (0.93-1.70, p=0.137)	1.19 (0.83-1.71, p=0.340)
14						
15	Ethnicity	Other Asian (%)	31 (9.9)	66 (12.0)	-	-
16		White (%)	164 (52.7)	270 (49.2)	0.76 (0.44-1.31, p=0.326)	0.73 (0.38-1.40, p=0.339)
17						
18		Black (%)	39 (12.4)	84 (15.3)	1.01 (0.52-1.98, p=0.974)	0.92 (0.43-1.97, p=0.827)
19						
20		Mixed (%)	6 (1.8)	4 (0.8)	0.36 (0.08-1.62, p=0.184)	0.74 (0.11-4.94, p=0.754)
21						
22		South Asian (%)	22 (7.0)	36 (6.6)	0.77 (0.34-1.76, p=0.531)	0.68 (0.28-1.65, p=0.390)
23		Other (%)	51 (16.2)	89 (16.2)	0.82 (0.43-1.55, p=0.535)	0.88 (0.45-1.74, p=0.716)
24						
25	Comorbidity	No (%)	65 (20.8)	95 (17.4)	-	-
26		Yes (%)	247 (79.2)	453 (82.6)	1.25 (0.82-1.89, p=0.296)	1.00 (0.53-1.88, p=0.993)
27						
28	Dyspnoea on attendance	No (%)	90 (28.8)	139 (25.4)	-	-
29		Yes (%)	222 (71.2)	409 (74.6)	1.19 (0.82-1.73, p=0.356)	0.84 (0.53-1.32, p=0.447)
30						
31	Oxygen Saturation	Median (IQR)	96 (6)	93 (8)	0.94 (0.91-0.97, p<0.001)**	0.97 (0.93-1.00, p=0.072)
32	Respiratory rate	Median (IQR)	23 (8)	25 (8)	1.04 (1.01-1.07, p=0.002)*	1.01 (0.98-1.05, p=0.462)
33	Glasgow Coma Scale	Median (IQR)	15 (0)	15 (0)	1.02 (0.89-1.17, p=0.819)	1.21 (0.98-1.48, p=0.073)
34	Temperature	Mean (SD)	37.2 (1.4)	37.7 (1.1)	1.48 (1.26-1.73, p<0.001)**	1.44 (1.20-1.74, p<0.001)**
35	Heart Rate	Mean (SD)	96.7 (20.5)	94.9 (21.5)	1.00 (0.99-1.00, p=0.305)	1.00 (0.99-1.01, p=0.702)
36						
37	Systolic Blood Pressure	Mean (SD)	136.2 (25.8)	132.6 (24.5)	0.99 (0.99-1.00, p=0.086)	0.99 (0.98-1.00, p=0.097)
38						
39						
40	Neutrophils	Median (IQR)	6.26 (4.52)	5.05 (3.93)	0.92 (0.89-0.96, p<0.001)**	0.87 (0.82-0.91, p<0.001)**
41	D-Dimer	Median (IQR)	1220 (2343)	1061 (1814)	1.00 (1.00-1.00, p=0.403)	1.00 (1.00-1.00, p=0.419)
42	C-Reactive Protein	Median (IQR)	45 (100)	77 (107)	1.00 (1.00-1.01, p<0.001)**	1.00 (1.00-1.01, p=0.021)*
43	Troponin	Median (IQR)	20 (55)	21 (46)	1.00 (1.00-1.00, p=0.890)	1.00 (1.00-1.00, p=0.667)
44	Albumin	Median (IQR)	39 (7)	37 (6)	0.97 (0.94-1.00, p=0.071)	1.02 (0.98-1.06, p=0.432)
45	Creatine Kinase	Median (IQR)	94 (131)	145 (263)	1.00 (1.00-1.00, p=0.119)	1.00 (1.00-1.00, p=0.152)
46	Admitted from ED	Admitted (%)	235 (75.2)	453 (82.7)	-	-
47		Discharged (%)	77 (24.8)	95 (17.3)	1.56 (1.06 -2.33, p=0.022)**	1.35 (0.79-2.30, p=0.272)
48						
49	Admitted To ITU from ED	No (%)	307 (98.5)	532 (97.1)	-	-
50		Yes (%)	5 (1.5)	16 (2.9)	1.92 (0.60-6.18, p=0.274)	1.06 (0.25-4.40, p=0.940)
51						
52	Thirty Day Follow up Status	Discharged (%)	259 (83.0)	368 (67.1)	-	-
53		Admitted (%)	22 (6.9)	47 (8.5)	1.53 (0.82-2.87, p=0.181)	1.64 (0.77-3.51, p=0.198)
54						
55						
56						
57						
58						
59						
60						

Dead (%) 31 (10.1) 133 (24.4) **3.00 (1.86-4.84, p<0.001)**** **2.81 (1.22-6.50, p=0.017)***

Table 4- Association of covariates with RT-PCR status for SARS-CoV 2, following propensity score matching and binomial logistic regression; SD- Standard deviation; IQR- Interquartile Range; *p<0.05; **p<0.001

		X-ray report		OR (univariable)	OR with XR as binary outcome (multivariable)	OR with XR as ordinal variable (multivariable)
		Other X-ray Findings	Classical COVID-19			
15n		430	430			
16RT-PCR for SARS-CoV 2	Negative (%)	187 (43.4)	125 (29.1)	-	-	-
	Positive (%)	243 (56.6)	305 (70.9)	1.85 (1.36-2.56, p<0.001)**	1.79 (1.25-2.56, p<0.002)*	1.94 (1.37-2.76, p<0.001)**
21Age	Mean (SD)	65.0 (18.9)	65.3 (16.9)	1.00 (0.99-1.01, p=0.849)	0.99 (0.98-1.00, p=0.164)	1.00 (0.99-1.01, p=0.542)
22Sex	Female (%)	176 (40.9)	175 (40.6)	-	-	-
	Male (%)	254 (59.1)	255 (59.3)	1.01 (0.75-1.37, p=0.940)	0.87 (0.63-1.20, p=0.400)	1.02 (0.49-2.09, p=0.967)
24Ethnicity	Other Asian (%)	49 (11.4)	48 (11.2)	-	-	-
	South Asian (%)	29 (6.7)	29 (6.7)	1.04 (0.52-2.04, p=0.912)	1.02 (0.47-2.17, p=0.965)	1.02 (0.49-2.09, p=0.967)
	Black (%)	61 (14.2)	61 (14.2)	1.02 (0.55-1.85, p=0.957)	0.88 (0.46-1.69, p=0.719)	0.92 (0.52-1.65, p=0.789)
	Mixed (%)	5 (1.2)	5 (1.2)	0.92 (0.21-4.00, p=0.911)	0.86 (0.18-4.17, p=0.853)	0.85 (0.17-4.30, p=0.838)
	Other (%)	70 (16.3)	70 (16.3)	1.02 (0.58-1.79, p=0.943)	0.98 (0.52-1.82, p=0.942)	0.93 (0.53-1.64, p=0.810)
	White (%)	216 (50.2)	217 (50.5)	1.03 (0.63-1.67, p=0.913)	0.97 (0.57-1.67, p=0.926)	0.90 (0.55-1.47, p=0.666)
31Comorbidity	No (%)	82 (19.1)	78 (18.1)	-	-	-
	Yes (%)	348 (80.9)	352 (81.9)	0.95 (0.66-1.36, p=0.777)	0.93 (0.59-1.49, p=0.782)	0.88 (0.57-1.37, p=0.592)
33Dyspnoea	No (%)	191 (29.3)	103 (24.0)	-	-	-
	Yes (%)	304 (70.7)	327 (76.0)	1.31 (0.92-1.88, p=0.123)	1.20 (0.80-1.82, p=0.380)	1.22 (0.83-1.80, p=0.301)
36Oxygen Saturation	Median (IQR)	95 (7)	93 (7)	0.94 (0.91-0.96, p<0.001)**	0.94 (0.92-0.97, p<0.001)**	0.94 (0.91-0.97, p<0.001)**
38Respiratory rate	Median (IQR)	24 (10)	24 (10)	1.01 (0.99-1.02, p=0.570)	0.97 (0.94-1.00, p=0.063)	0.98 (0.96-1.01, p=0.157)
39Glasgow Coma Scale	Median (IQR)	15 (0)	15 (0)	1.04 (0.92-1.19, p=0.524)	1.05 (0.90-1.23, p=0.503)	1.05 (0.92-1.21, p=0.464)
42Temperature	Mean (SD)	37.6 (1.1)	37.5 (1.3)	0.93 (0.83-1.06, p=0.297)	0.79 (0.67-0.93, p=0.006)*	0.85 (0.73-0.99, p=0.031)*
43Heart Rate	Mean (SD)	95.7 (21.4)	95.5 (21.0)	1.00 (0.99-1.01, p=0.888)	1.00 (0.99-1.01, p=0.864)	1.00 (0.99-1.01, p=0.872)
44Systolic Blood Pressure	Mean (SD)	133.8 (25.0)	134.0 (25.6)	1.00 (0.99-1.01, p=0.907)	1.00 (0.99-1.01, p=0.335)	1.00 (1.00-1.01, p=0.478)
46Neutrophils	Median (IQR)	5.44 (4.54)	5.67 (4.03)	1.00 (0.97-1.04, p=0.892)	0.96 (0.92-1.01, p=0.143)	0.96 (0.92-1.01, p=0.115)
47D-Dimer	Median (IQR)	1119 (2221)	1119 (1850)	1.00 (1.00-1.00, p=0.513)	1.00 (1.00-1.00, p=0.568)	1.00 (1.00-1.00, p=0.385)
49C-Reactive Protein	Median (IQR)	46 (93)	88 (110)	1.00 (0.99-1.00, p<0.001)**	1.00 (1.00-1.01, p<0.001)**	1.00 (1.00-1.01, p<0.001)**
51Troponin	Median (IQR)	23 (54)	20 (46)	1.00 (1.00-1.00, p=0.231)	1.00 (1.00-1.00, p=0.277)	1.00 (1.00-1.00, p=0.059)
52Albumin	Median (IQR)	39 (7)	37 (6)	0.93 (0.90-0.96, p<0.001)**	0.93 (0.90-0.97, p=0.001)*	0.94 (0.91-0.97, p=0.001)*
54Creatine Kinase	Median (IQR)	110 (183)	134 (239)	1.00 (1.00-1.00, p=0.535)	1.00 (1.00-1.00, p=0.242)	1.00 (1.00-1.00, p=0.186)

1							
2							
3	Admitted from	Admitted (%)	315 (73.3)	373 (86.7)	2.37 (1.63-3.46,	2.30 (1.46-3.63,	2.22 (1.47-3.33,
4	ED				p<0.001)**	p<0.001)**	p<0.001)**
5		Discharged (%)	115 (26.7)	57 (13.3)	-	-	-
6	Admitted to ITU	No (%)	423 (98.4)	416 (96.7)	-	-	-
7	from ED						
8		Yes (%)	7 (1.6)	14 (3.3)	2.17 (0.69-6.67, p=0.181)	1.27 (0.32-5.00, p=0.732)	1.34 (0.36-5.00, p=0.653)
9	30 Day Follow	Discharged (%)	316 (73.5)	311 (72.3)	-	-	-
10	Up Status						
11		Admitted (%)	34 (7.9)	34 (7.9)	1.31 (0.81-2.13, p=0.282)	1.32 (0.69-2.53, p=0.392)	1.43 (0.78-2.63, p=0.653)
12		Dead (%)	80 (18.6)	85 (19.8)	1.03 (0.73-1.45, p=0.886)	1.38 (0.80-2.37, p=0.247)	1.41 (0.87-2.27, p=0.157)
13							
14							
15							

Table 5- Association of covariates with CXR report following propensity score matching and either binomial or ordinal logistic regression; SD- Standard deviation; IQR- Interquartile Range; *p<0.05; **p<0.001

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

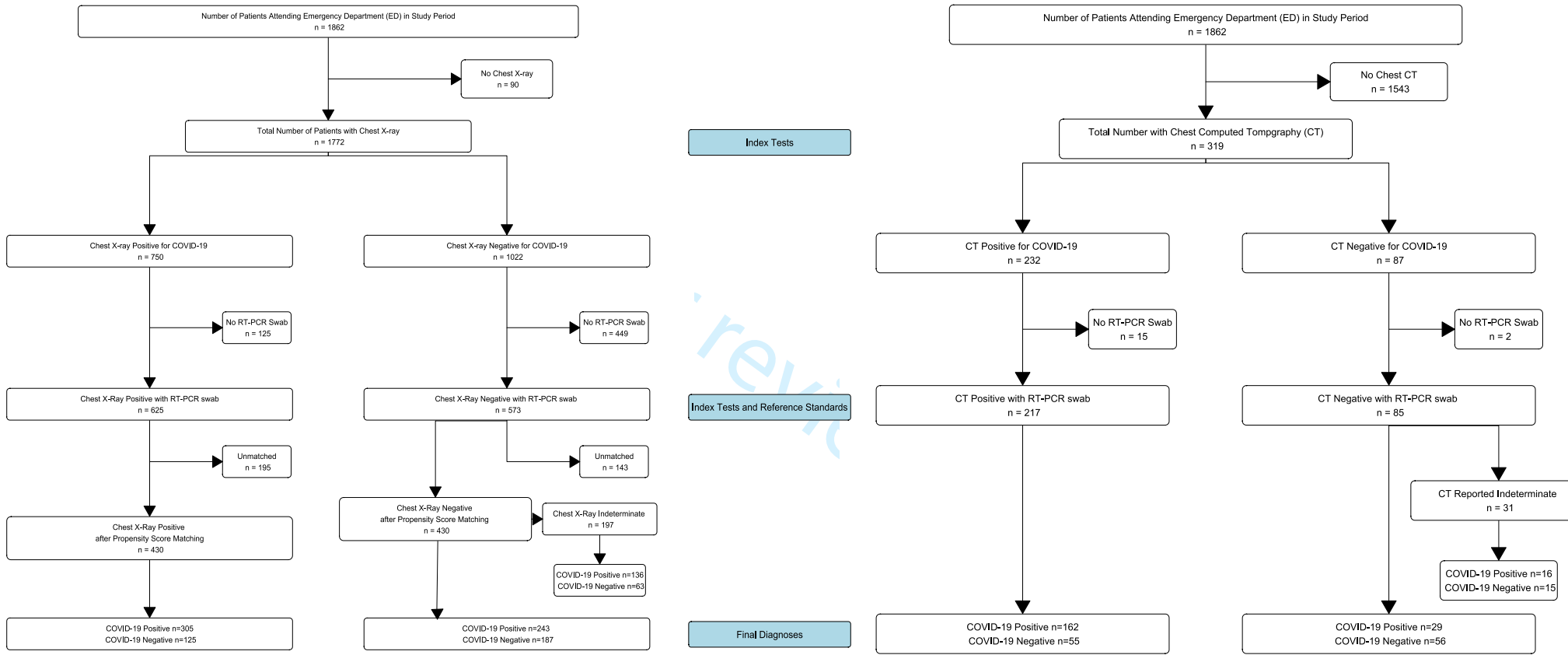


Figure 1- Inclusion and exclusion of patients during study period with test results

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Odds Ratio of Positivity for SARS-CoV 2 by RT-PCR

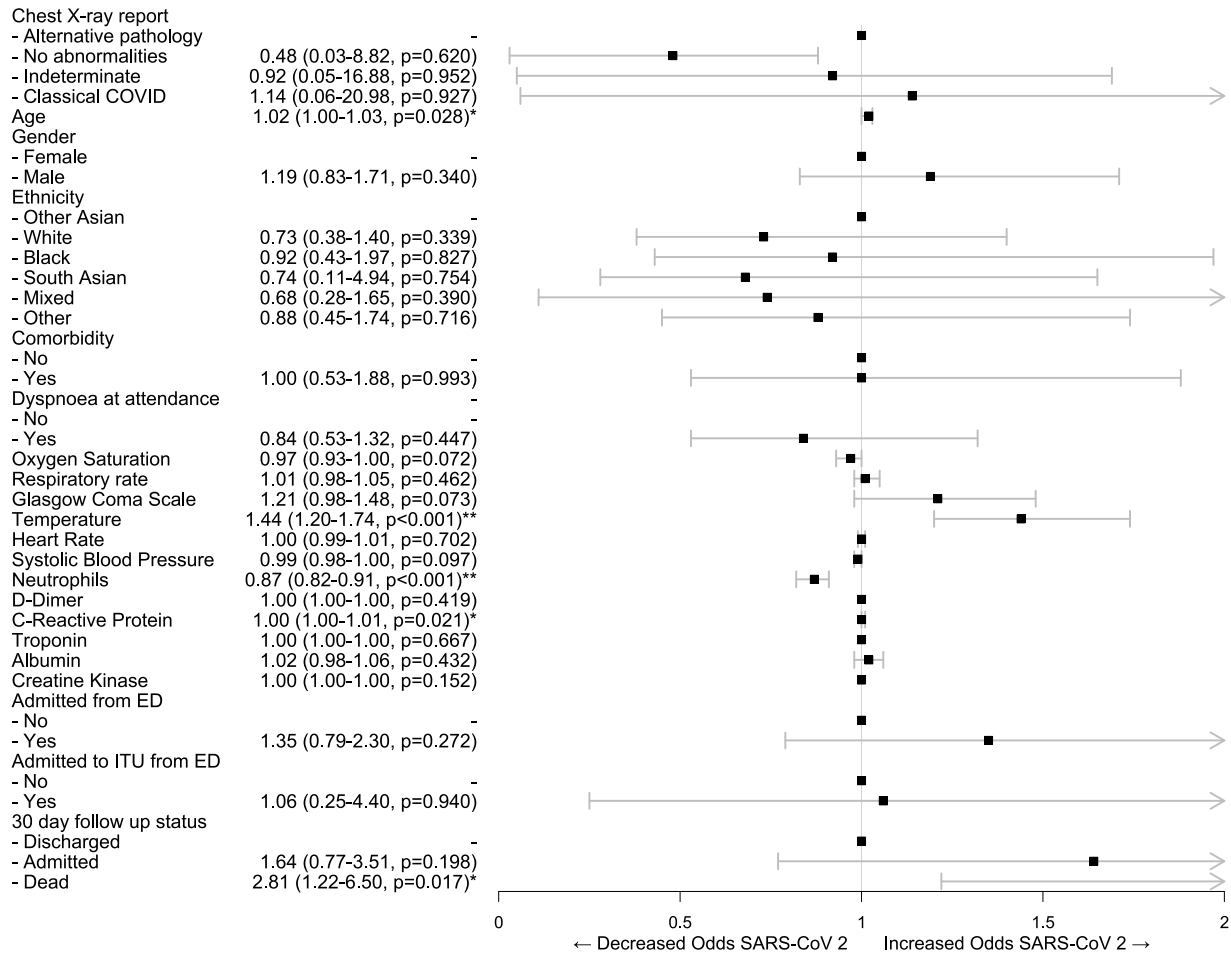


Figure 2- Forest plot of odds ratios of variables associated with RT-PCR positivity for SARS-CoV 2, following multiple imputation, propensity score matching and binomial logistic regression; *significant difference at the <0.05 level; **significant difference at the <0.001 level

Odds Ratio of Classical COVID-19 Findings on Chest X-Ray

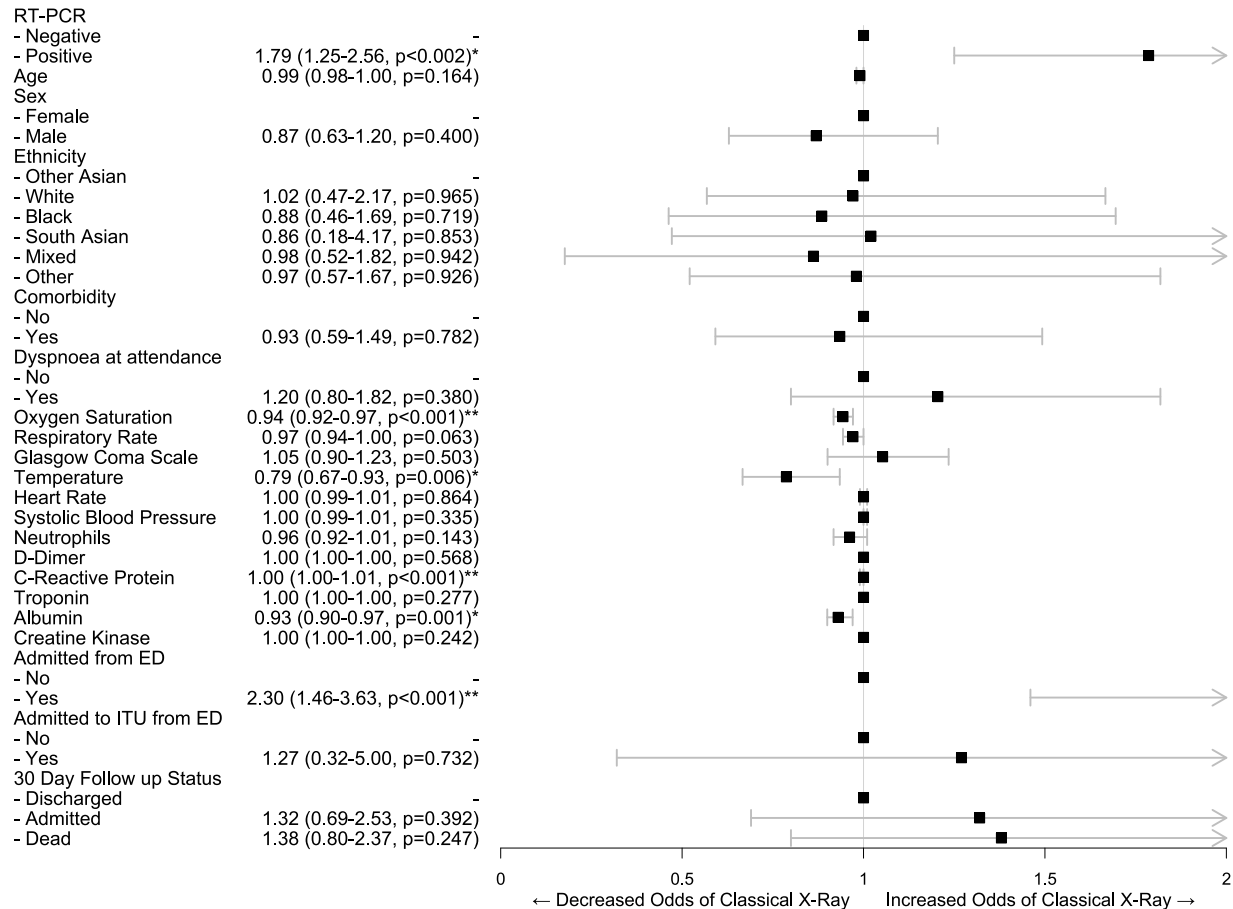
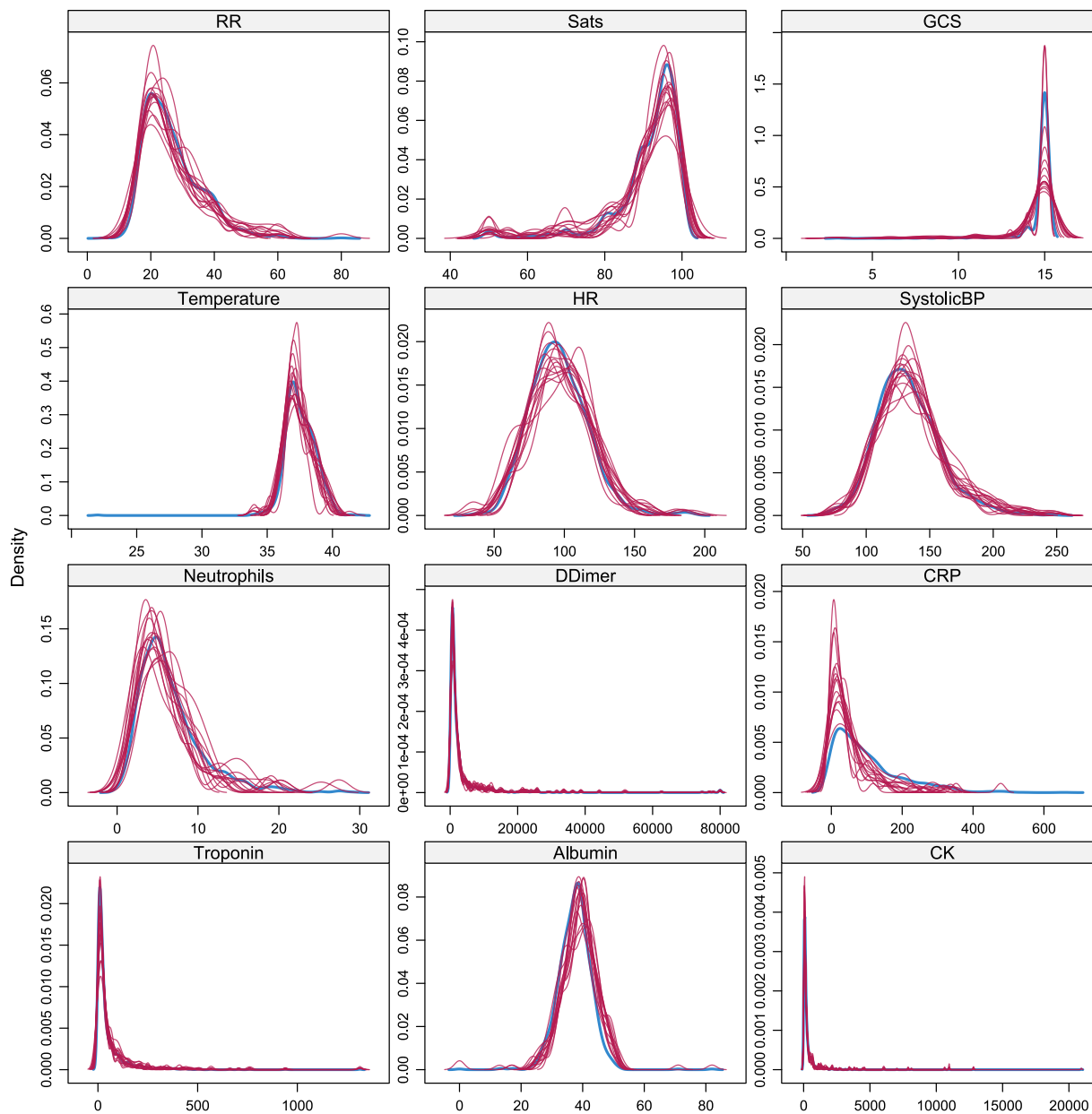


Figure 3- Forest plot of odds ratios of variables associated with classical Chest X-ray features COVID-19 following propensity score matching and binomial logistic regression; *significant difference at the <0.05 level; **significant difference at the <0.001 level

Supplementary file 1



Supplementary figure 1- Density plots of imputed datasets; Blue represents original dataset; other colours are individual imputed datasets (n=15)

Covariate:	Means Treated	Means Control	Standard Deviation Control	Mean Difference
Overall Propensity Score	0.422997940	0.53935303	0.1449627	-0.1163550897
Female	36.3782051	45.026178	0.4979547	-8.64797288
Male	63.6217949	54.973822	0.4979547	8.64797288
Age	63.796474359	66.19022688	18.5893357	-23.937525171
Comorbidity- Yes	76.1217949	84.467714	0.3625287	-8.34591892
Ethnicity- South Asian	6.5705128	6.631763	0.2490539	-0.06124983
Ethnicity- Black	16.1858974	11.518325	0.3195219	4.66757283
Ethnicity- Mixed	0.9615385	1.396161	0.1174340	-0.43462210
Ethnicity- Other	18.9102564	13.263525	0.3394765	5.64673110
Ethnicity- White	46.6346154	57.766143	0.4943635	-11.13152772
Respiratory Rate	29.214743590	24.01745201	7.2639816	5.1972915828

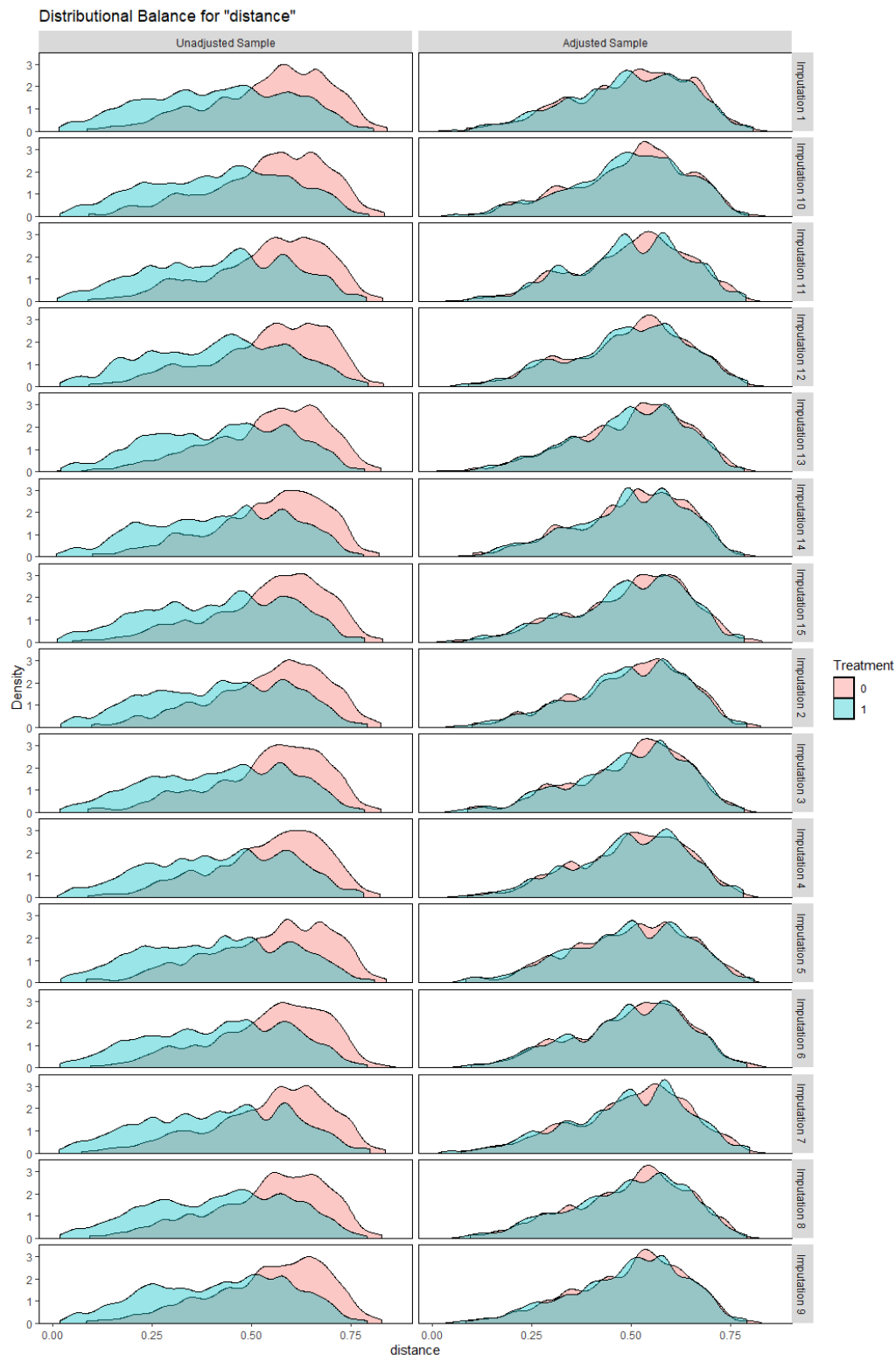
Supplementary table 1- Means of data before multiple imputation and propensity score matching

	Type	Minimum Difference Adjusted	Mean Difference Adjusted	Maximum Difference Adjusted
Distance	Distance	0.016988	0.027107	0.040963
Sex = Male	Binary	-0.03917	-0.0028	0.015982
Age	Contin.	-0.04586	-0.01371	0.027589
Comorbidity = Yes	Binary	-0.02331	-0.00778	0.004598
Ethnicity = Other Asian	Binary	-0.01392	0.002362	0.016471
Ethnicity = South Asian	Binary	-0.01399	-0.00136	0.011905
Ethnicity = Black	Binary	-0.01852	0.000443	0.015982
Ethnicity = Mixed	Binary	-0.00464	0.001403	0.007042
Ethnicity = Other	Binary	-0.01152	4.30E-06	0.00939
Ethnicity = White	Binary	-0.02353	-0.00285	0.018433
Respiratory Rate	Contin.	-0.06157	-0.03478	-0.00442

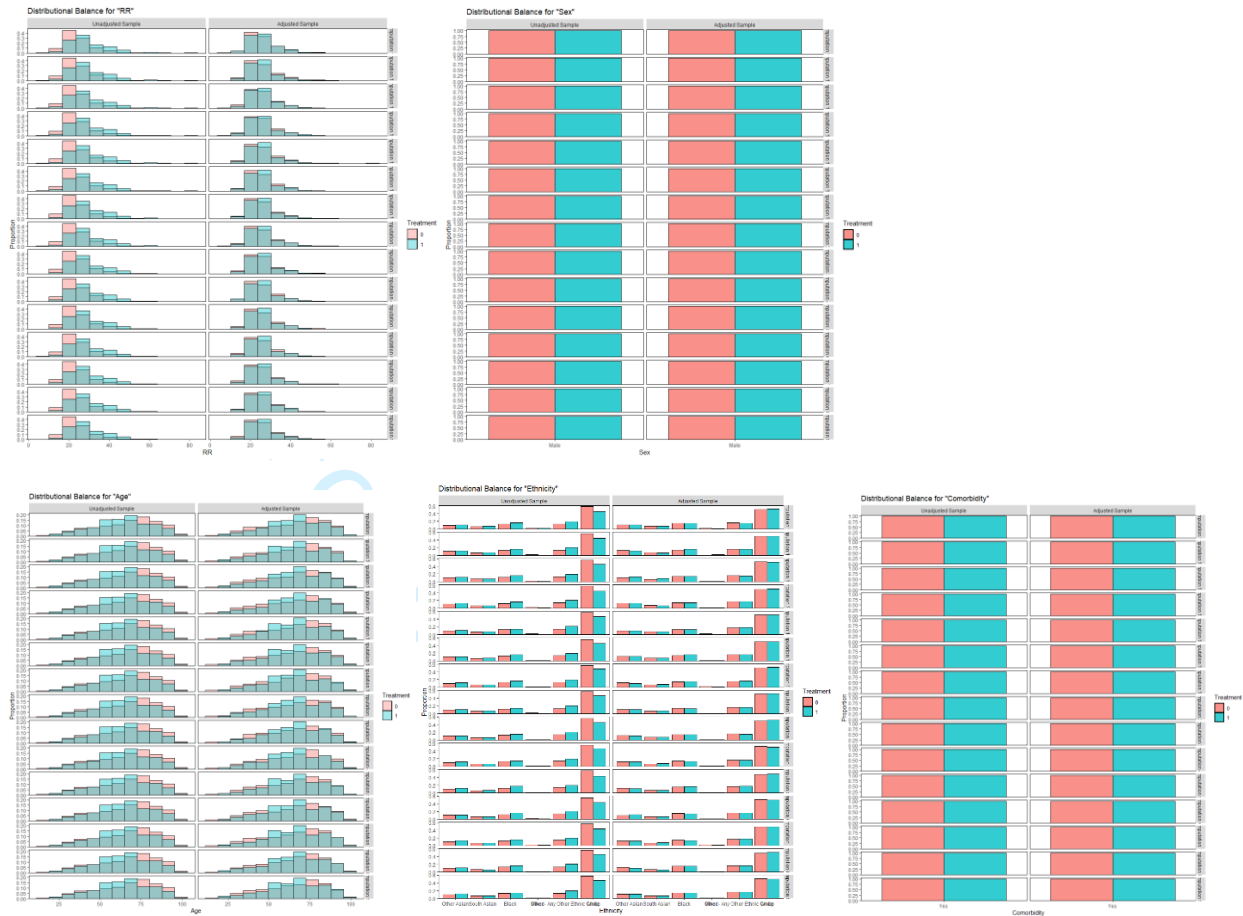
Supplementary table 2- Balance summary across imputations

	XR- Negative	XR- Positive	Total
All	573	625	1,198
Matched	430	430	860
Unmatched	143	195	338
Discarded	0	0	0

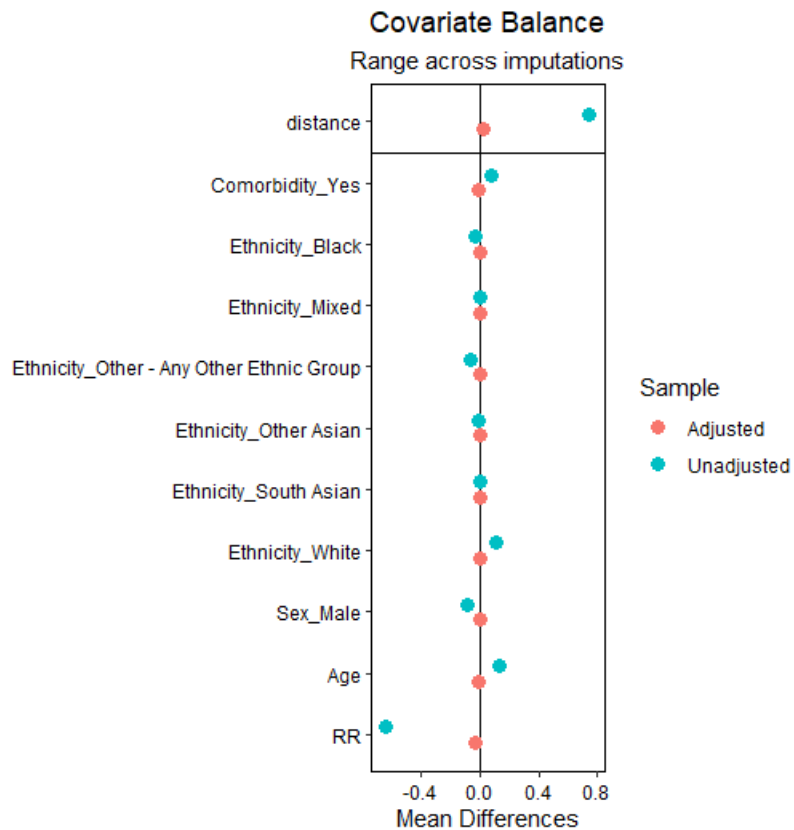
Supplementary table 3- Average Sample sizes pre- and post- matching across imputed data sets



Supplementary figure 2- Density plot of propensity scores pre- and post- matching in each imputed dataset; treatment units represent a positive X-ray for COVID-19, whereas a control unit represents a negative X-ray



Supplementary figure 3- Histogram of distributions for each matching covariate pre- and post- matching in each imputed dataset; treatment units represent a positive X-ray for COVID-19, whereas a control unit represents a negative X-ray



Supplementary figure 4- Love plot of pooled balances across imputed datasets in matching covariates after matching

CXR in COVID Analysis

Dr Aditya Borakati

Royal Free Hospital, Pond Street, London, NW3 2QG
a.borakati@doctors.org.uk

2020-10-06

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1 Software Environment and Packages	7
1.1 Load Packages and Data	9
1.1.1 Load Packages:	9
1.2 Power Calculation	9
2 Load Data:	11
3 Data Cleaning	13
3.0.1 Follow Up Swabs + Initial Swabs Positive: .	14
3.0.2 Paired XR and RT-PCR data	14
4 Demographic table of raw data	17
5 Imputation	21
6 Propensity Score Matching	23
6.1 Match Balance Diagnostics	23
7 Matched Demographics Table:	25
8 Diagnostic Accuracy	27
8.1 CT Data and Accuracy	28
8.2 CT and XR accuracy comparison	30
8.2.1 Sensitivity	30
8.3 Intermodality Agreement	34
8.3.1 Diagnostic Accuracy Analysis when Indeterminate Reports of CXR and CT are taken as positive	35
9 Pooled Regression after Multiple Imputation and Propensity Score Matching	39
9.0.1 Pooled Univariate Odds Ratios for OverallPos as dependent variable	39
9.0.2 Binomial Logistic Regression with Positive Chest X-ray Report as Dependent Variable	40
9.0.3 Univariate XRPositive as dependent	40
9.0.4 Multivariate XRPositive as dependent	40
9.0.5 Pooled Ordinal Logistic Regression with XRPositive as dependent	41

iv

9.1 Forest Plots	41
9.2 Correlation Matrix	43
9.3 STARD Flow Diagram	46

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1 Software Environment and Packages

```

R version 4.0.0 (2020-04-24)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 19041)
Matrix products: default
locale:
 LC_COLLATE=English_United Kingdom.1252 LC_CTYPE=English_United Kingdom.1252
 LC_MONETARY=English_United Kingdom.1252 LC_NUMERIC=C
 LC_TIME=English_United Kingdom.1252
attached base packages:
 stats    graphics  grDevices  utils      datasets  methods   base
other attached packages:
 corrplot 0.84
   Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of
   a Correlation Matrix (Version 0.84). Available from
   https://github.com/taiyun/corrplot
MKmisc 1.6
   Kohl M (2019). MKmisc: Miscellaneous functions from M. Kohl_. R package
   version 1.6, http://www.stamats.de
epiR 1.0-14
   Mark Stevenson with contributions from Telmo Nunes, Cord Heuer, Jonathon
   Marshall, Javier Sanchez, Ron Thornton, Jenő Reiczigel, Jim Robison-Cox,
   Paola Sebastiani, Peter Solymos, Kazuki Yoshida, Geoff Jones, Sarah
   Pirikahu, Simon Firestone, Ryan Kyle, Johann Popp, Mathew Jay and Charles
   Reynard. (2020). epiR: Tools for the Analysis of Epidemiological Data. R
   package version 1.0-14. https://CRAN.R-project.org/package=epiR
Matching 4.9-7
   Jasjeet S. Sekhon (2011). Multivariate and Propensity Score Matching
   Software with Automated Balance Optimization: The Matching Package for R.
   Journal of Statistical Software, 42(7), 1-52. URL
   http://www.jstatsoft.org/v42/i07/.
MASS 7.3-51.5
   Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S.
   Fourth Edition. Springer, New York. ISBN 0-387-95457-0
Ordinal 2019.12-10
   Christensen, R. H. B. (2019). ordinal - Regression Models for Ordinal Data. R
   package version 2019.12-10. https://CRAN.R-
   project.org/package=ordinal.
Hmisc 4.4-0
   Frank E Harrell Jr, with contributions from Charles Dupont and many
   others. (2020). Hmisc: Harrell Miscellaneous. R package version 4.4-0.
   https://CRAN.R-project.org/package=Hmisc
Formula 1.2-3
   Achim Zeileis, Yves Croissant (2010). Extended Model Formulas in R:
   Multiple Parts and Multiple Responses. Journal of Statistical Software
   34(1), 1-13. doi:10.18637/jss.v034.i01
lattice 0.20-41
   Sarkar, Deepayan (2008) Lattice: Multivariate Data Visualization with R.
   Springer, New York. ISBN 978-0-387-75968-5

```

8 1 Software Environment and P...

```
mice 3.8.0
  Stef van Buuren, Karin Groothuis-Oudshoorn (2011). mice: Multivariate
  Imputation by Chained Equations in R. Journal of Statistical Software,
  45(3), 1-67. URL https://www.jstatsoft.org/v45/i03/.
readxl 1.3.1
  Hadley Wickham and Jennifer Bryan (2019). readxl: Read Excel Files. R
  package version 1.3.1. https://CRAN.R-project.org/package=readxl
finalfit 1.0.1
  Ewen Harrison, Tom Drake and Riinu Ots (2020). finalfit: Quickly Create
  Elegant Regression Results Tables and Plots when Modelling. R package
  version 1.0.1. https://CRAN.R-project.org/package=finalfit
MatchIt 3.0.2
  Daniel E. Ho, Kosuke Imai, Gary King, Elizabeth A. Stuart (2011). MatchIt:
  Nonparametric Preprocessing for Parametric Causal Inference. Journal of
  Statistical Software, Vol. 42, No. 8, pp. 1-28. URL
  http://www.jstatsoft.org/v42/i08/
tableone 0.11.1
  Kazuki Yoshida (2020). tableone: Create 'Table 1' to Describe Baseline
  Characteristics. R package version 0.11.1.
  https://CRAN.R-project.org/package=tableone
forcats 0.5.0
  Hadley Wickham (2020). forcats: Tools for Working with Categorical
  Variables (Factors). R package version 0.5.0.
  https://CRAN.R-project.org/package=forcats
stringr 1.4.0
  Hadley Wickham (2019). stringr: Simple, Consistent Wrappers for Common
  String Operations. R package version 1.4.0.
  https://CRAN.R-project.org/package=stringr
dplyr 0.8.5
  Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2020).
  dplyr: A Grammar of Data Manipulation. R package version 0.8.5.
  https://CRAN.R-project.org/package=dplyr
purrr 0.3.4
  Lionel Henry and Hadley Wickham (2020). purrr: Functional Programming
  Tools. R package version 0.3.4. https://CRAN.R-project.org/package=purrr
readr 1.3.1
  Hadley Wickham, Jim Hester and Romain François (2018). readr: Read
  Rectangular Text Data. R package version 1.3.1.
  https://CRAN.R-project.org/package=readr
tidyr 1.0.2
  Hadley Wickham and Lionel Henry (2020). tidyr: Tidy Messy Data. R package
  version 1.0.2. https://CRAN.R-project.org/package=tidyr
tibble 3.0.0
  Hadley Wickham and Lionel Henry (2020). tidyr: Tidy Messy Data. R package
  version 1.0.2. https://CRAN.R-project.org/package=tidyr
ggplot2 3.3.0
  H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag
  New York, 2016.
tidyverse 1.3.0
  Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source
  Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
forestplot 1.9
  Max Gordon and Thomas Lumley (2019). forestplot: Advanced Forest Plot Using
  'grid' Graphics. R package version 1.9. https://CRAN.R-project.org/package=forestplot
MatchThem 0.9.3
  Farhad Pishgar and Noah Greifer (2020). MatchThem: Matching and Weighting
  Multiply Imputed Datasets. R package version 0.9.3. https://CRAN.R-project.org/package=MatchThem
```

1.1 Load Packages and Data

9

```
miceadds 3.9-14
  Robitzsch, A., & Grund, S. (2020). miceadds: Some Additional Multiple
  Imputation Functions, Especially for 'mice'. R package version 3.9-14.
  https://CRAN.R-project.org/package=miceadds
cobalt 4.2.2
  Noah Greifer (2020). cobalt: Covariate Balance Tables and Plots. R package
  version 4.2.2. https://CRAN.R-project.org/package=cobalt
```

1.1 Load Packages and Data

1.1.1 Load Packages:

```
library(MKmisc)
library(tidyverse)
library(tableone)
library(MatchIt)
library(finalfit)
library(readxl)
library(cobalt)
library(mice)
library(miceadds)
library(Hmisc)
library(epiR)
library(MatchThem)
library(ordinal)
library(forestplot)
```

1.2 Power Calculation

1.2.0.0.1 This code calculates the sample size (positive and negative by gold standard test) needed to evaluate a diagnostic test with 56% sensitivity at 80% power with alpha 0.05. The 56% value is the lower confidence reported by Wong et al. and lower sensitivities typically require higher sample sizes, the result is the same whether specificity or sensitivities are passed as arguments, the previously published specificities are higher than sensitivities so for a generous estimate, the sensitivity was used.

```
power <- power.diagnostic.test(sens = 0.56,
  sig.level = 0.05, delta = 0.1, power = 0.8) %>%
  print()
```


10

1 Software Environment and P...

```
Diagnostic test exact power calculation
```

```
    sens = 0.56  
      n  = 165  
     n1 = 165  
    delta = 0.1  
sig.level = 0.05  
    power = 0.8  
     prev = NULL
```

```
NOTE: n is number of cases, n1 is number of controls
```

2 Load Data:

```
data <- read_csv("FullDataWithCT.csv", col_types = cols(Age = col_integer(),
  Albumin = col_number(), CK = col_number(),
  CT = col_character(), CRP = col_number(),
  DDimer = col_number(), DateOfDeath = col_date(format = "%d/%m/%Y"),
  DateOfDischarge = col_date(format = "%d/%m/%Y"),
  DateOfVisit = col_date(format = "%d/%m/%Y"),
  DateOfSymptomOnset = col_date(format = "%d/%m/%Y"),
  DiastolicBP = col_number(), FiO2 = col_skip(),
  GCS = col_number(), HR = col_number(),
  MRN = col_skip(), NEWS = col_number(),
  `NEWS2(noFiO2)` = col_skip(), Neutrophils = col_number(),
  RR = col_number(), Sats = col_number(),
  `Supplemental Oxygen` = col_skip(), SystolicBP = col_number(),
  Temperature = col_number(), Troponin = col_number(),
  CTBSTI = col_integer()))
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

3 Data Cleaning

3.0.0.0.1 Format data into factors/ differences between dates:

```
data <- mutate_if(data, is.character, as.factor)
data$DayOfSymptoms <- difftime(data$DateOfVisit,
  data$DateOfSymptomOnset, units = "days")
data$TimeToDeath <- abs(difftime(data$DateOfDeath,
  data$DateOfVisit, units = "days"))
data$DayOfSymptoms <- as.numeric(data$DayOfSymptoms)
data$TimeToDeath <- as.numeric(data$TimeToDeath)
```

3.0.0.1 Recode ethnicities as too many options:

3.0.0.1.0.1 This code collapses the ethnicity categories into 'White', 'Black', 'South Asian', 'Other Asian', 'Mixed' or 'Other';

```
data$Ethnicity <- fct_collapse(data$Ethnicity,
  White = c("White - British", "White - Irish",
    "White - Any Other White Background"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  Black = c("Black - Any Other Black Background",
    "Black or Black British - African",
    "Black or Black British - African",
    "Black or Black British - Caribbean"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  `South Asian` = c("Asian or Asian British - Bangladeshi",
    "Asian or Asian British - Indian",
    "Asian or Asian British - Pakistani"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  `Other Asian` = c("Asian - Any Other Asian Background",
    "Other - Chinese"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  Mixed = c("mixed - Any Other mixed Background",
    "Mixed - Any Other Mixed Background",
    "Mixed - White and Asian", "Mixed - White and Black African",
    "mixed - White and Black Caribbean",
    "Mixed - White and Black Caribbean"))
```

3.0.0.1.0.2 New XR positive column for “Classic Covid” or not:

```
data$XRPositive <- ifelse(data$XR Chest ==  
  "Classic COVID", "Positive", "Negative")  
data$XRPositive <- as.factor(data$XRPositive)
```

3.0.1 Follow Up Swabs + Initial Swabs Positive:

3.0.1.0.0.1 Creates new column ‘OverallPos’ which includes initial RT-PCR swab and follow-up swabs in 30 days of attendance, if any are positive the value will be positive in this column

```
data$OverallPos <- case_when(data$RTPCR ==  
  "Positive" | data$FollowUpPos == "Positive" ~  
  "Positive")  
data$OverallPos <- replace_na(data$OverallPos,  
  "Negative")
```

3.0.1.0.0.2 Create new vector with all variable names (i.e. the column headers)

```
explanatory <- names(data)
```

3.0.2 Paired XR and RT-PCR data

3.0.2.1 Creates new variable ‘completedata’ which contains only patients who had both CXR and RT-PCR in ED

```
completedata <- filter(data, !is.na(data$XRPositive) &  
  !is.na(data$RTPCR))
```

3.0.2.1.1 Remove missing data variable

```
completedata <- completedata[-c(31)]
```

3.0.2.2 Format complete data variables

```
completedata$OverallPos <- as.factor(completedata$OverallPos)

completedata$ThirtyDayFU <- as.factor(completedata$ThirtyDayFU)
completedata$TimeToDeath <- abs(difftime(completedata$DateOfDeath,
completedata$DateOfVisit, units = "days"))

completedata$TimeToDeath <- as.numeric(completedata$TimeToDeath)
```

3.0.2.2.0.1 Set 'XRchest' as ordinal variable on scale of 'Alternative pathology' as lowest value and 'Classical COVID' as highest

```
completedata$XRchest <- ordered(completedata$XRchest,
levels = c("Alternative pathology", "No abnormalities",
"Indeterminate", "Classic COVID"))
```

3.0.2.2.0.2 Convert CT BSTI grade column into factor:

```
completedata$CTBSTI <- as.factor(completedata$CTBSTI)
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

4 Demographic table of raw data

4.0.0.0.1 This code creates an unformatted demographic table (table 2 in manuscript), for the raw data, stratified by RT-PCR status, significance testing between RT-PCR +ve and -ve groups is carried out automatically using chi squared, t-tests, ANOVA etc.; there is also a column for the proportion of missing data

```

CreateTableOne(vars = explanatory,
               strata = 'OverallPos',
               data = completedata) -> demogtable

#### List nonnormal factors for summarisation as median / IQR and non
      parametric statistical test

explanatorynormal<-c("Sats", "RR", "GCS", "SystolicBP", "Temperature", "HR",
                    "Neutrophils",
                    "DDimer", "Albumin", "CRP", "CK", "Troponin")
as.data.frame(print(demogtable, nonnormal = explanatorynormal, missing =
                    TRUE))->demogtable

write.csv(demogtable, file = "Demogtable.csv")

```

Age (mean (SD))	62.74 (17.72)	66.18 (17.58)
0.001		
Ethnicity (%)		
0.097		
Other Asian	29 (8.0)	72 (11.8)
South Asian	27 (7.5)	38 (6.2)
Black	41 (11.4)	91 (14.9)
Mixed	6 (1.7)	6 (1.0)
Other - Any Other Ethnic Group	56 (15.5)	105 (17.2)
White	202 (56.0)	297 (48.8)
Sex = Male (%)	233 (53.6)	480 (62.9)
0.002		
Sats (median [IQR])	95.00 [92.00, 98.00]	93.00 [88.00,
96.00]	<0.001 nonnorm	
RR (median [IQR])	22.00 [20.00, 28.00]	26.00 [20.00,
32.00]	<0.001 nonnorm	
GCS (median [IQR])	15.00 [15.00, 15.00]	15.00 [15.00,
15.00]	0.043 nonnorm	
SystolicBP (median [IQR])	134.00 [119.00, 151.50]	130.00 [115.00,
145.00]	0.009 nonnorm	
DiastolicBP (mean (SD))	79.54 (16.40)	75.61 (14.51)
<0.001		
HR (median [IQR])	96.00 [83.00, 110.00]	94.00 [81.00,
108.00]	0.092 nonnorm	

18

4 Demographic table of raw data

Temperature (median [IQR]) 38.40] <0.001 nonnorm	37.10 [36.60, 38.00]	37.70 [37.00,
XR Chest (%) <0.001		
Alternative pathology	4 (0.9)	3 (0.4)
No abnormalities	178 (40.9)	136 (17.8)
Indeterminate	83 (19.1)	169 (22.1)
Classic COVID	170 (39.1)	455 (59.6)
CTPA = PE (%) 0.127	16 (30.2)	28 (45.9)
Comorbidity = Yes (%) 0.669	297 (79.0)	482 (80.3)
Dyspnoea = Yes (%) 0.034	274 (69.4)	497 (75.5)
Neutrophils (median [IQR]) 7.61] <0.001 nonnorm	6.42 [4.55, 9.11]	5.25 [3.69,
DDimer (median [IQR]) 2428.50] 0.204 nonnorm	1250.00 [619.00, 3059.00]	1105.00 [626.00,
Albumin (median [IQR]) 40.00] <0.001 nonnorm	39.00 [35.00, 42.00]	37.00 [34.00,
CRP (median [IQR]) 158.00] <0.001 nonnorm	51.00 [13.00, 117.00]	83.00 [42.00,
CK (median [IQR]) 342.75] <0.001 nonnorm	91.00 [54.00, 169.00]	146.50 [78.00,
Troponin (median [IQR]) 53.00] 0.278 nonnorm	19.00 [7.00, 53.00]	20.00 [9.00,
Admitted = Discharged (%) 0.003	104 (24.0)	128 (16.8)
AdmittedToITU = Yes (%) 0.005	5 (1.3)	32 (4.8)
RTPCR = Positive (%) <0.001	0 (0.0)	738 (96.7)
CT = 1 (%) 0.011	37 (57.8)	26 (86.7)
NEWS (mean (SD)) 0.032	4.36 (3.06)	5.48 (2.71)
ThirtyDayFU (%) <0.001		
1	219 (78.2)	367 (58.3)
2	14 (5.0)	49 (7.8)
3	18 (6.4)	60 (9.5)
4	29 (10.4)	154 (24.4)
CTBSTI (%) <0.001		
0	23 (22.1)	6 (3.3)
1	52 (50.0)	157 (85.8)
2	14 (13.5)	14 (7.7)
3	15 (14.4)	6 (3.3)
DayOfSymptoms (mean (SD)) 0.368	9.84 (9.63)	8.56 (15.80)
TimeToDeath (mean (SD)) 0.618	50.33 (77.93)	57.76 (70.02)
XRPositive = Positive (%) <0.001	170 (39.1)	455 (59.6)
OverallPos = Positive (%)	0 (0.0)	763 (100.0)

1
2
3
4
5
6 **4.0.0.0.2 Limited dataset comprising relevant data and those without**
7 **significant missingness:**
8

```
9  
10 limcompletedata <- dplyr::select(completedata,  
11   c("Age", "XRchest", "Ethnicity", "Sex",  
12     "RR", "Sats", "GCS", "Temperature",  
13     "HR", "SystolicBP", "DiastolicBP",  
14     "Neutrophils", "DDimer", "CRP", "Troponin",  
15     "Albumin", "CK", "OverallPos", "Admitted",  
16     "AdmittedToITU", "ThirtyDayFU", "Dyspnoea",  
17     "Comorbidity", "XRPositive"))  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

5 Imputation

5.0.0.0.0.1 This code generates 15 imputed datasets using the permuted mean matching method, based on the 'limcompletedata' dataset which has filtered the most relevant fields, with minimal missing data initially

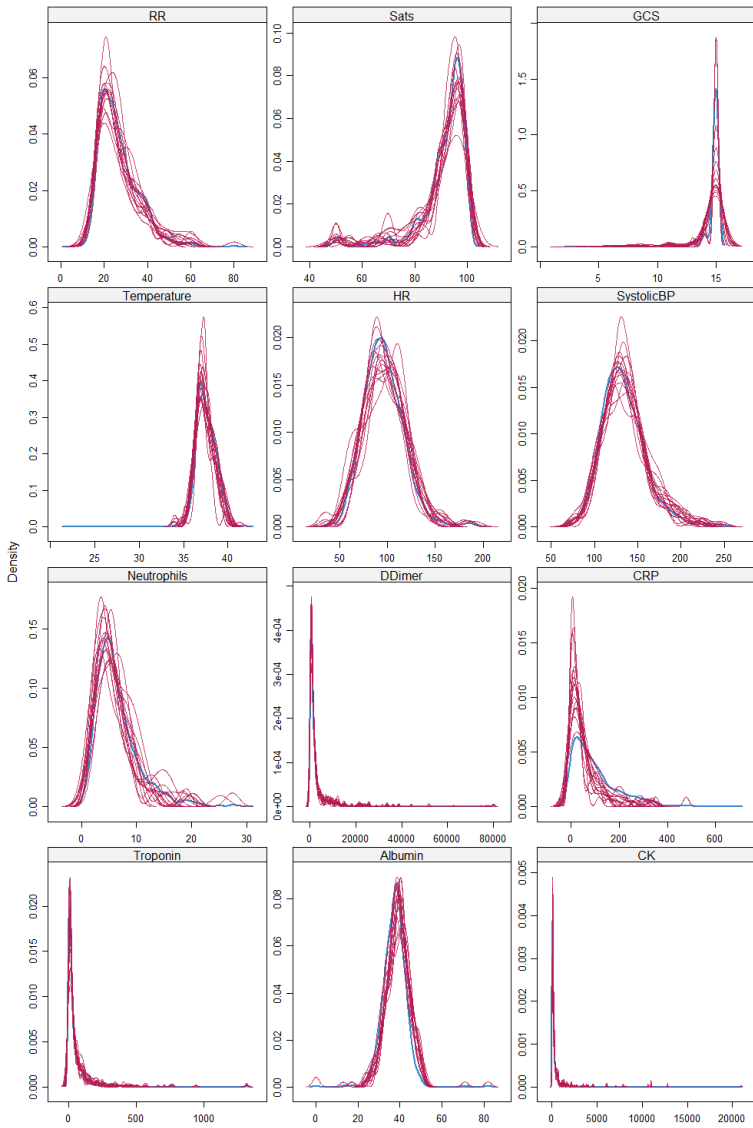
```
imputed <- mice(limcompletedata, m = 15,  
method = "pmm")
```

5.0.0.0.0.2 Imputation Diagnostics Density plot, this corresponds to supplementary figure 1:

```
densityplot(imputed)
```

22

5 Imputation



6 Propensity Score Matching

6.0.0.0.1 This code matches data in the imputed datasets on whether the XR was reported classical COVID or not, the matching is done based on the covariates Sex, Age, Comorbidity, Ethnicity and Respiratory Rate

```

library(MatchThem)
#### MatchThem package requires dependent variable to be coded as 0 or 1
imputed[["data"]][["XRPositive"]] %>% recode_factor("Positive" = "1",
"Negative" = "0") ->imputed[["data"]][["XRPositive"]]
matchthem(
  XRPositive ~ Sex + Age + Comorbidity + Ethnicity + RR,
  data = imputed,
  method = 'nearest',
  verbose = FALSE,
  replace = FALSE,
  ratio = 1,
  caliper = 0.2,
  m.order = "random",) -> matchedtest
### Set XRChest to unordered for binomial analyses
matchedtest[["datasets"]]c(1:15)[["XRChest"]] %>% factor(ordered = FALSE) ->
  matched2[["datasets"]]c(1:15)[["XRChest"]]

```

6.1 Match Balance Diagnostics

6.1.0.0.1 Creates plots and table with mean difference and distribution of values in covariates between XR +ve and -ve groups after matching across all imputed datasets:

```

#### Supplementary tables 1,2 and 3:
bal.tab(matchedtest)
#### Supplementary figure 2
bal.plot(matchedtest)
#### Supplementary figure 3:
bal.plot(matchedtest, var.name = "Age", type = "histogram",
  which = "both")
bal.plot(matchedtest, var.name = "Sex", type = "histogram",
  which = "both")
bal.plot(matchedtest, var.name = "Ethnicity",

```

```
1     type = "histogram", which = "both")
2
3
4     bal.plot(matchedtest, var.name = "RR", type = "histogram",
5         which = "both")
6     bal.plot(matchedtest, var.name = "Comorbidity",
7         type = "histogram", which = "both")
8     ### Supplementary figure 4:
9     love.plot(matchedtest)
```

7 Matched Demographics Table:

7.0.0.0.1 Stack matched imputed datasets into one large dataset and split into COVID +ve and -ve groups:

```
### 'all=FALSE' gets matched data only
stacked <- MatchThem::complete(matchedtest,
  n = c(1:15), all = FALSE)
stacked <- stacked %>% filter(.imp > 0)
```

7.0.0.0.2 Creates demographics table as above, but on propensity matched imputed datasets, corresponds to Table 4:

```
table4 <- CreateTableOne(strata = "OverallPos",
  data = stacked)
#### Means and SD kept as is, mean counts
#### calculated after dividing by 15 (as 15
#### imputed datasets)
```

7.0.0.0.3 Creates demographic table stratified by XR Positive or Negative on matched imputed datasets, corresponds to Table 5:

```
table5 <- CreateTableOne(strata = "XRPositive",
  data = stacked)
#### Means and SD kept as is, mean counts
#### calculated after dividing by 15 (as 15
#### imputed datasets)
```

7.0.0.0.4 Summary statistics for pooled data:

```
### Normal means sd
explanatorynorm <- c("Age", "Temperature",
  "HR", "SystolicBP")
summarynormalOverallPos <- stacked %>% group_by(OverallPos) %>%
```


26

7 Matched Demographics Table:

```
summarise_at(vars(explanatorynorm), list(mean.default,
sd))
summarynormalXRPositive <- stacked %>% group_by(XRPositive) %>%
summarise_at(vars(explanatorynorm), list(mean.default,
sd))

### Non normal medians and IQR
summarynormalOverallPos <- stacked %>% group_by(OverallPos) %>%
summarise_at(vars(explanatorynormal),
list(median, IQR))
summarynormalXRPositive <- stacked %>% group_by(XRPositive) %>%
summarise_at(vars(explanatorynormal),
list(median, IQR))
```

8 Diagnostic Accuracy

8.0.0.1 This section generates the diagnostic accuracy statistics (e.g. sensitivity, specificity) for CXR and CT with RT-PCR as the reference standard using the matched imputed datasets

8.0.0.2 This code creates a contingency table of False/ True Positives and Negatives for Chest X-ray taken from the demographic tables above:

```
contingxr <- matrix(c(305, 243, 125, 187),
  nrow = 2, ncol = 2)

colnames(contingxr) <- c("PCR+", "PCR-")

rownames(contingxr) <- c("XR+", "XR-")
```

8.0.0.2.1 This function calculates diagnostic accuracy test statistics:

```
xraccuracy <- epi.tests(contingxr, conf.level = 0.95)
```

8.0.0.3 Giving the diagnostic accuracy output for CXR in table 3:

```
xraccuracy
```

	Outcome +	Outcome -	Total
Test +	305	125	430
Test -	243	187	430
Total	548	312	860

Point estimates and 95 % CIs:

Apparent prevalence	0.50 (0.47, 0.53)
True prevalence	0.64 (0.60, 0.67)

Sensitivity	0.56 (0.51, 0.60)
Specificity	0.60 (0.54, 0.65)
Positive predictive value	0.71 (0.66, 0.75)
Negative predictive value	0.43 (0.39, 0.48)
Positive likelihood ratio	1.39 (1.19, 1.62)
Negative likelihood ratio	0.74 (0.65, 0.84)

8.0.0.3.0.1 NB diagnostic accuracy values in table available in list view of xrraccuracy variable

8.1 CT Data and Accuracy

8.1.0.0.0.1 Only those with CT and RT PCR:

```
CTdata <- filter(data, is.na(data$CTBSTI) ==
  FALSE & is.na(data$RTPCR) == FALSE)
```

8.1.0.0.0.2 Select relevant variables

```
CTdata <- dplyr::select(CTdata, c("Age",
  "XRchest", "Ethnicity", "Sex", "RR",
  "Sats", "GCS", "Temperature", "HR", "SystolicBP",
  "DiastolicBP", "Neutrophils", "DDimer",
  "CRP", "Troponin", "OverallPos", "Admitted",
  "AdmittedToITU", "ThirtyDayFU", "Dyspnoea",
  "Comorbidity", "XRPositive", "OverallPos",
  "CTBSTI"))
```

8.1.0.0.0.3 Set RT-PCR as factor:

```
CTdata$OverallPos <- as.factor(CTdata$OverallPos)
```

8.1 CT Data and Accuracy

29

8.1.0.0.4 Rename 1 and 0 to Positive and Negative:

```
CTdata$CTPositive <- ifelse(CTdata$CTBSTI ==
  "1", "Positive", "Negative")
CTdata$CTPositive <- as.factor(CTdata$CTPositive)
```

8.1.0.0.5 Regression with CT as outcome variable:

```
CT <- finalfit(
  CTdata,
  "OverallPos",
  c(
    "Age",
    "Sex",
    "RR",
    "GCS",
    "CTPositive",
    "Temperature",
    "HR",
    "SystolicBP",
    "DiastolicBP",
    "Sats",
    "Dyspnoea",
    "Comorbidity"
  ),
  confint_level = 0.95
)
```

8.1.0.0.6 Contingency table of True/False Positives and Negatives for CT taken from Regression table:

```
contingct <- matrix(c(CT[7, 4], CT[6, 4],
  CT[7, 3], CT[6, 3]), nrow = 2, ncol = 2)
colnames(contingct) <- c("PCR+", "PCR-")
rownames(contingct) <- c("CT+", "CT-")
contingct <- substr(contingct, start = 1,
  stop = 3)
contingct <- sapply(contingct, as.numeric)
contingct <- matrix(contingct, nrow = 2,
  ncol = 2)
colnames(contingct) <- c("PCR+", "PCR-")
rownames(contingct) <- c("CT+", "CT-")
```

30

8 Diagnostic Accuracy

8.1.0.0.7 Diagnostic accuracy statistics for CT

```

epi.tests(contingct, conf.level = 0.95) -> ctaccuracy
      Outcome +   Outcome -   Total
Test +         162         55     217
Test -          29         56      85
Total          191        111     302

```

Point estimates and 95 % CIs:

```

-----
Apparent prevalence           0.72 (0.66, 0.77)
True prevalence               0.63 (0.58, 0.69)
Sensitivity                   0.85 (0.79, 0.90)
Specificity                   0.50 (0.41, 0.60)
Positive predictive value     0.75 (0.68, 0.80)
Negative predictive value     0.66 (0.55, 0.76)
Positive likelihood ratio     1.71 (1.41, 2.08)
Negative likelihood ratio     0.30 (0.21, 0.44)
-----

```

8.1.0.0.8 NB Diagnostic accuracy values found in list view rather than output

8.2 CT and XR accuracy comparison

8.2.0.1 In this section mean differences of diagnostic accuracy statistics between CT and Chest X-ray with confidence intervals and p-values are calculated

8.2.1 Sensitivity

8.2 CT and XR accuracy comp...

31

8.2.1.0.0.1 Upper confidence limit for difference in sensitivity

```
ubsens <- (ctaccuracy[["elements"]][["se.up"]] -  
xraccuracy[["elements"]][["se.low"]])
```

8.2.1.0.0.2 Lower confidence limit for difference in sensitivity

```
lbsens <- (ctaccuracy[["elements"]][["se.low"]] -  
xraccuracy[["elements"]][["se.up"]])
```

8.2.1.0.0.3 Mean difference in sensitivity

```
meansens <- ctaccuracy[["elements"]][["se"]] -  
xraccuracy[["elements"]][["se"]]
```

8.2.1.0.0.4 Standard error for sensitivity

```
sesens <- (ubsens - lbsens)/(2 * 1.96)
```

8.2.1.0.0.5 value for difference in sensitivity

```
zsens <- meansens/sesens
```

8.2.1.0.0.6 P-value for difference in sensitivity

```
psens <- exp(-0.717 * zsens - 0.416 * zsens^2)
```

32

8 Diagnostic Accuracy

8.2.1.0.0.7 Format values into 'mean difference (95% CI) p-value' rounded to 2 d.p.

```
diffsens <- sprintf("%s (%s-%s)", round(meansens,
  digits = 2), round(lbsens, digits = 2),
  round(ubsens, digits = 2))
diffsensp <- c(diffsens, psens)
```

8.2.1.0.0.8 Subsequent analyses in this section follow the code above

```
## Specificity
ubspec <- (ctaccuracy[["elements"]][["sp.up"]] -
  xraccuracy[["elements"]][["sp.low"]])
lbspec <- (ctaccuracy[["elements"]][["sp.low"]] -
  xraccuracy[["elements"]][["sp.up"]])
meanspec <- ctaccuracy[["elements"]][["sp"]] -
  xraccuracy[["elements"]][["sp"]]
sespec <- (ubspec - lbspec)/(2 * 1.96)
zspec <- meanspec/sespec
pspec <- exp(-0.717 * zspec - 0.416 * zspec^2)
diffspec <- sprintf("%s (%s-%s)", round(meanspec,
  digits = 2), round(lbspec, digits = 2),
  round(ubspec, digits = 2))
diffspecp <- c(diffspec, pspec)

ubda <- (ctaccuracy[["elements"]][["da.up"]] -
  xraccuracy[["elements"]][["da.low"]])
lbda <- (ctaccuracy[["elements"]][["da.low"]] -
  xraccuracy[["elements"]][["da.up"]])
meanda <- ctaccuracy[["elements"]][["da"]] -
  xraccuracy[["elements"]][["da"]]
seda <- (ubda - lbda)/(2 * 1.96)
zda <- meanda/seda
pda <- exp(-0.717 * zda - 0.416 * zda^2)
diffda <- sprintf("%s (%s-%s)", round(meanda,
  digits = 2), round(lbda, digits = 2),
  round(ubda, digits = 2))
diffdap <- c(diffda, pda)

## Positive Likelihood Ratio
ublrpos <- (ctaccuracy[["elements"]][["lrpos.up"]] -
  xraccuracy[["elements"]][["lrpos.low"]])
lblrpos <- (ctaccuracy[["elements"]][["lrpos.low"]] -
  xraccuracy[["elements"]][["lrpos.up"]])
meanlrpos <- ctaccuracy[["elements"]][["lrpos"]] -
  xraccuracy[["elements"]][["lrpos"]]
selrpos <- (ublrpos - lblrpos)/(2 * 1.96)
zlrpos <- meanlrpos/selrpos
plrpos <- exp(-0.717 * zlrpos - 0.416 * zlrpos^2)
difflrpos <- sprintf("%s (%s-%s)", round(meanlrpos,
  digits = 2), round(lblrpos, digits = 2),
```

8.2 CT and XR accuracy comp...

33

```

1   round(ublrrpos, digits = 2))
2   diffplrposp <- c(diffplrpos, plrpos)
3   ## Negative Likelihood Ratios
4   ublrneg <- (ctaccuracy[["elements"]][["lrneg.up"]] -
5     xraccuracy[["elements"]][["lrneg.low"]])
6   lblrneg <- (ctaccuracy[["elements"]][["lrneg.low"]] -
7     xraccuracy[["elements"]][["lrneg.up"]])
8   meanlrneg <- ctaccuracy[["elements"]][["lrneg"]] -
9     xraccuracy[["elements"]][["lrneg"]]
10  selrneg <- (ublrneg - lblrneg)/(2 * 1.96)
11  zlrneg <- meanlrneg/selrneg
12  plrneg <- exp(-0.717 * zlrneg - 0.416 * zlrneg^2)
13  diffflrneg <- sprintf("%s (%s-%s)", round(meanlrneg,
14    digits = 2), round(lblrneg, digits = 2),
15    round(ublrneg, digits = 2))
16  diffflrnegp <- c(diffflrneg, plrneg)
17
18  ## Positive Predictive Value
19  ppv <- (ctaccuracy[["elements"]][["ppv.low"]] -
20    xraccuracy[["elements"]][["ppv.up"]])
21  meanppv <- ctaccuracy[["elements"]][["ppv"]] -
22    xraccuracy[["elements"]][["ppv"]]
23  seppv <- (ubppv - lbppv)/(2 * 1.96)
24  zppv <- meanppv/seppv
25  pppv <- exp(-0.717 * zppv - 0.416 * zppv^2)
26  diffppv <- sprintf("%s (%s-%s)", round(meanppv,
27    digits = 2), round(lbppv, digits = 2),
28    round(ubppv, digits = 2))
29  diffppvp <- c(diffppv, pppv)
30
31  ## Negative Predictive Value
32  npv <- (ctaccuracy[["elements"]][["npv.low"]] -
33    xraccuracy[["elements"]][["npv.up"]])
34  meannpv <- ctaccuracy[["elements"]][["npv"]] -
35    xraccuracy[["elements"]][["npv"]]
36  senpv <- (ubnpv - lbnpv)/(2 * 1.96)
37  znpv <- meannpv/senpv
38  pnpv <- exp(-0.717 * znpv - 0.416 * znpv^2)
39  diffnpv <- sprintf("%s (%s-%s)", round(meannpv,
40    digits = 2), round(lbnpv, digits = 2),
41    round(ubnpv, digits = 2))
42  diffnpvp <- c(diffnpv, pnpv)
43
44  ## Apparent Prevalence
45  meantp <- ctaccuracy[["elements"]][["tp"]] -
46    xraccuracy[["elements"]][["tp"]]
47  setp <- (ubtp - lbtp)/(2 * 1.96)
48  ztp <- meantp/setp
49  ptp <- exp(-0.717 * ztp - 0.416 * ztp^2)
50  diffftp <- sprintf("%s (%s-%s)", round(meantp,
51    digits = 2), round(lbtp, digits = 2),
52    round(ubtp, digits = 2))
53  diffftp <- c(diffftp, ptp)
54
55  ## True Prevalence
56  meanap <- ctaccuracy[["elements"]][["ap"]] -
57    xraccuracy[["elements"]][["ap"]]

```



```

seap <- (ubap - lbap)/(2 * 1.96)
zap <- meanap/seap
pap <- exp(-0.717 * zap - 0.416 * zap^2)
diffap <- sprintf("%s (%s-%s)", round(meanap,
  digits = 2), round(lbap, digits = 2),
  round(ubap, digits = 2))
diffapp <- c(diffap, pap)

```

8.3 Intermodality Agreement

8.3.0.0.1 This section contains code to analyse the level of agreement in the unmatched CT dataset which contains only data with CT, XR and RT-PCR

8.3.0.0.2 First- comparing CT and XR agreement

```

library(irr)
kappa2(c(CTdata$XRPositive, CTdata$CTPositive),
  weight = "squared")
d <- CTdata %>% select(c("CTPositive", "XRPositive"))
View(d)
kappa2(d, weight = "squared")

```

8.3.0.0.3 Output:

```

Cohen's Kappa for 2 Raters (Weights: squared)

Subjects = 287
Raters = 2
Kappa = 0.406

z = 7.14
p-value = 9.37e-13

```

8.3.0.0.4 The following code compares RT-PCR, CT and XR

```

d2 <- CTdata %>% select(c("CTPositive", "XRPositive",
  "OverallPos"))
View(d2)
kappam.fleiss(d2)

```

8.3 Intermodality Agreement

35

8.3.0.0.5 Output:

```

Fleiss' Kappa for m Raters

Subjects = 287
Raters = 3
Kappa = 0.361

z = 10.6
p-value = 0

```

8.3.1 Diagnostic Accuracy Analysis when Indeterminate Reports of CXR and CT are taken as positive

8.3.1.1 XR Indeterminates

8.3.1.1.0.1 New column for positive if indeterminate

```

stacked$XRIndPositive <- ifelse(stacked$XR Chest ==
  "Classic COVID" | stacked$XR Chest ==
  "Indeterminate", "Positive", "Negative")
stacked$XRIndPositive <- as.factor(stacked$XRIndPositive)
stackedpos <- stacked %>% filter(OverallPos ==
  "Positive")
stackedneg <- stacked %>% filter(OverallPos ==
  "Negative")
summary(stackedpos$XRIndPositive)
summary(stackedneg$XRIndPositive)

contingxrind <- matrix(c(441, 107, 186, 126),
  nrow = 2, ncol = 2)
colnames(contingxrind) <- c("PCR+", "PCR-")

rownames(contingxrind) <- c("XR+", "XR-")
xrindaccuracy <- epi.tests(contingxrind)

```

36

8 Diagnostic Accuracy

8.3.1.1.0.2 In this section mean differences of diagnostic accuracy statistics between CT (when CT indeterminates are not counted as positive) and Chest X-ray with confidence intervals and p-values are calculated, follows the same pattern as code previously

```
##### Sensitivity Upper confidence Limit for
##### difference in sensitivity

ubsens <- (ctaccuracy[["elements"]][["se.up"]] -
  xrindaccuracy[["elements"]][["se.low"]])
## Lower confidence limit for difference
## in sensitivity
lbsens <- (ctaccuracy[["elements"]][["se.low"]] -
  xrindaccuracy[["elements"]][["se.up"]])
## Mean difference in sensitivity
meansens <- ctaccuracy[["elements"]][["se"]] -
  xrindaccuracy[["elements"]][["se"]]
## Standard error for sensitivity
sesens <- (ubsens - lbsens)/(2 * 1.96)
## Z value for difference in sensitivity
zsens <- meansens/sesens
## P-value for difference in sensitivity
psens <- exp(-0.717 * zsens - 0.416 * zsens^2)
### Format values into 'mean difference
### (95% CI) p-value' rounded to 2 d.p.
diffsens <- sprintf("%s (%s-%s)", round(meansens,
  digits = 2), round(lbsens, digits = 2),
  round(ubsens, digits = 2))
diffsensp <- c(diffsens, psens)

### Subsequent analyses in this section
### follow the code above Specificity
ubspec <- (ctaccuracy[["elements"]][["sp.up"]] -
  xrindaccuracy[["elements"]][["sp.low"]])
lbspec <- (ctaccuracy[["elements"]][["sp.low"]] -
  xrindaccuracy[["elements"]][["sp.up"]])
meanspec <- ctaccuracy[["elements"]][["sp"]] -
  xrindaccuracy[["elements"]][["sp"]]
sespec <- (ubspec - lbspec)/(2 * 1.96)
zspec <- meanspec/sespec
pspec <- exp(-0.717 * zspec - 0.416 * zspec^2)
diffspec <- sprintf("%s (%s-%s)", round(meanspec,
  digits = 2), round(lbspec, digits = 2),
  round(ubspec, digits = 2))
diffspecp <- c(diffspec, pspec)

ubda <- (ctaccuracy[["elements"]][["da.up"]] -
  xrindaccuracy[["elements"]][["da.low"]])
lbda <- (ctaccuracy[["elements"]][["da.low"]] -
  xrindaccuracy[["elements"]][["da.up"]])
meanda <- ctaccuracy[["elements"]][["da"]] -
  xrindaccuracy[["elements"]][["da"]]
seda <- (ubda - lbda)/(2 * 1.96)
```

8.3 Intermodality Agreement

37

```

1  zda <- meanda/seda
2
3
4  pda <- exp(-0.717 * zda - 0.416 * zda^2)
5  diffda <- sprintf("%s (%s-%s)", round(meanda,
6    digits = 2), round(lbda, digits = 2),
7    round(ubda, digits = 2))
8  diffdap <- c(diffda, pda)
9  ## Positive Likelihood Ratio
10 ublrpos <- (ctaccuracy[["elements"]][["lrpos.up"]] -
11   xrindaccuracy[["elements"]][["lrpos.low"]])
12 lblrpos <- (ctaccuracy[["elements"]][["lrpos.low"]] -
13   xrindaccuracy[["elements"]][["lrpos.up"]])
14 meanlrpos <- ctaccuracy[["elements"]][["lrpos"]] -
15   xrindaccuracy[["elements"]][["lrpos"]]
16 selrpos <- (ublrpos - lblrpos)/(2 * 1.96)
17 zlrpos <- meanlrpos/selrpos
18 plrpos <- exp(-0.717 * zlrpos - 0.416 * zlrpos^2)
19 diffplrpos <- sprintf("%s (%s-%s)", round(meanlrpos,
20   digits = 2), round(lblrpos, digits = 2),
21   round(ublrpos, digits = 2))
22 diffplrposp <- c(diffplrpos, plrpos)
23 ## Negative Likelihood Ratios
24 ublrneg <- (ctaccuracy[["elements"]][["lrneg.up"]] -
25   xrindaccuracy[["elements"]][["lrneg.low"]])
26 lblrneg <- (ctaccuracy[["elements"]][["lrneg.low"]] -
27   xrindaccuracy[["elements"]][["lrneg.up"]])
28 meanlrneg <- ctaccuracy[["elements"]][["lrneg"]] -
29   xrindaccuracy[["elements"]][["lrneg"]]
30 selrneg <- (ublrneg - lblrneg)/(2 * 1.96)
31 zlrneg <- meanlrneg/selrneg
32 plrneg <- exp(-0.717 * zlrneg - 0.416 * zlrneg^2)
33 diffplrneg <- sprintf("%s (%s-%s)", round(meanlrneg,
34   digits = 2), round(lblrneg, digits = 2),
35   round(ublrneg, digits = 2))
36 diffplrnegp <- c(diffplrneg, plrneg)
37
38 ## Positive Predictive Value
39 ppv <- (ctaccuracy[["elements"]][["ppv.low"]] -
40   xrindaccuracy[["elements"]][["ppv.up"]])
41 meanppv <- ctaccuracy[["elements"]][["ppv"]] -
42   xrindaccuracy[["elements"]][["ppv"]]
43 seppv <- (ubppv - lbppv)/(2 * 1.96)
44 zppv <- meanppv/seppv
45 pppv <- exp(-0.717 * zppv - 0.416 * zppv^2)
46 diffppv <- sprintf("%s (%s-%s)", round(meanppv,
47   digits = 2), round(lbppv, digits = 2),
48   round(ubppv, digits = 2))
49 diffppvp <- c(diffppv, pppv)
50
51 ## Negative Predictive Value
52 npv <- (ctaccuracy[["elements"]][["npv.low"]] -
53   xrindaccuracy[["elements"]][["npv.up"]])
54 meannpv <- ctaccuracy[["elements"]][["npv"]] -
55   xrindaccuracy[["elements"]][["npv"]]
56 senpv <- (ubnpv - lbnpv)/(2 * 1.96)
57 znpv <- meannpv/senpv
58 pnpv <- exp(-0.717 * znpv - 0.416 * znpv^2)
59 diffnpv <- sprintf("%s (%s-%s)", round(meannpv,
60   digits = 2), round(lbnpv, digits = 2),

```

```

4     round(ubnpv, digits = 2))
5     diffnpvp <- c(diffnpv, pnpv)

6
7     ## True Prevalence
8     meantp <- ctaccuracy[["elements"]][["tp"]] -
9       xrindaccuracy[["elements"]][["tp"]]
10    setp <- (ubtp - lbtp)/(2 * 1.96)
11    ztp <- meantp/setp
12    ptp <- exp(-0.717 * ztp - 0.416 * ztp^2)
13    diffftp <- sprintf("%s (%s-%s)", round(meantp,
14      digits = 2), round(lbtp, digits = 2),
15      round(ubtp, digits = 2))
16    diffftpp <- c(diffftp, ptp)

17    ## Apparent Prevalence
18    meanap <- ctaccuracy[["elements"]][["ap"]] -
19      xrindaccuracy[["elements"]][["ap"]]
20    seap <- (ubap - lbap)/(2 * 1.96)
21    zap <- meanap/seap
22    pap <- exp(-0.717 * zap - 0.416 * zap^2)
23    diffpap <- sprintf("%s (%s-%s)", round(meanap,
24      digits = 2), round(lbap, digits = 2),
25      round(ubap, digits = 2))
26    diffpapp <- c(diffpap, pap)

```

8.3.1.2 CT Indeterminates

8.3.1.2.0.1 New column for positive if indeterminate

```

27
28
29     CTdata$CTIndPositive <- ifelse(CTdata$CTBSTI ==
30       "1" | CTdata$CTBSTI == "2", "Positive",
31       "Negative")
32     CTdata$CTIndPositive <- as.factor(CTdata$CTIndPositive)
33     valuesctind <- CTdata %>% group_by(OverallPos,
34       CTIndPositive) %>% summarise(n = n())
35     ctcontingind <- matrix(data = c(178, 13,
36       70, 41), nrow = 2, ncol = 2)
37
38     colnames(ctcontingind) <- c("PCR+ve", "PCR-ve")
39     rownames(ctcontingind) <- c("CT+ve", "CT-ve")
40     ctindaccuracy <- epi.tests(ctcontingind)

```

9 Pooled Regression after Multiple Imputation and Propensity Score Matching

9.0.0.0.0.1 Binomnal Logistic regression with RT-PCR as dependent variable

```

overallposmatchimp <- matchedtest %>% with(glm(formula(ff_formula(dependent =
  "OverallPos",
  explanatory = c("Age", "Ethnicity", "Sex",
    "RR", "GCS", "Temperature", "HR",
    "SystolicBP", "Neutrophils", "DDimer",
    "CRP", "Troponin", "Albumin", "CK",
    "Sats", "Admitted", "AdmittedToITU",
    "ThirtyDayFUTwo", "Dyspnoea", "Comorbidity",
    "XRchest"))), family = "binomial"),
  all = FALSE)
P <- overallposmatchimp %>% pool()
multivarpooledeoverallpos = P %>% fit2df(estimate_name = "OR (multiple
  imputation)",
  exp = TRUE)

```

9.0.0.0.0.2 'multivarpooledeoverallpos' produces multivariate odds ratios for each explanatory variable, corresponding to Table 4

9.0.1 Pooled Univariate Odds Ratios for OverallPos as dependent variable

9.0.1.0.0.1 This code is run with each of the explanatory variables in table 4 as arguments to produce their respective odds Ratios in table 4

```

overallposmatchimpunivar <- matchedtest %>%
  with(glm(formula(ff_formula(dependent = "OverallPos",

```

40

9 Pooled Regression after Multi...

```

    explanatory = "XRchest")), family = "binomial"))
P <- overallposmatchimpunivar %>% pool()
univarpooledoverallpos = univaroverallpos <- P %>%
  fit2df(estimate_name = "OR (univariate)",
        exp = TRUE)
univaroverallpos

```

9.0.2 Binomial Logistic Regression with Positive Chest X-ray Report as Dependent Variable

9.0.2.0.0.1 This code follows the format above to produce univariate and multivariate odds ratios for each explanatory variable for having a positive XR report

9.0.3 Univariate XRPositive as dependent

9.0.3.0.0.1 (different explanatory variables passed into function to produce Odds ratios for each)

```

XRchestmatchimp <- matchedtest %>% with(glm(formula(ff_formula(dependent =
  "XRPositive",
  explanatory = "Comorbidity")), family = "binomial"))
P <- XRchestmatchimp %>% pool()
multivarpooledXRchest = univarXRchest <- P %>%
  fit2df(estimate_name = "OR (univariate)",
        exp = TRUE)
univarXRchest

```

9.0.4 Multivariate XRPositive as dependent

```

XRchestmatchimp <- matchedtest %>% with(glm(formula(ff_formula(dependent =
  "XRPositive",
  explanatory = c("Age", "OverallPos",
  "Ethnicity", "Sex", "RR", "GCS",
  "Temperature", "HR", "SystolicBP",
  "Neutrophils", "DDimer", "CRP", "Troponin",
  "Albumin", "CK", "Sats", "Admitted",
  "AdmittedToITU", "ThirtyDayFUTwo",
  "Dyspnoea", "Comorbidity"))), family = "binomial"))
P <- XRchestmatchimp %>% pool()
multivarpooledXRchest = multivarXRchest <- P %>%
  fit2df(estimate_name = "OR (multivariate)",

```

9.1 Forest Plots

41

```

exp = TRUE)
multivarXRchest

```

9.0.5 Pooled Ordinal Logistic Regression with XRPositive as dependent

9.0.5.0.1 This code also produces multivariate odds ratios for table 5, however, uses ordinal linear regression after the CXR report variable is converted to an ordered categorical variable, with alternative pathology as the lowest and classic covid as the highest value (see table 3)

```

XRchestmatchimpord <- matchedtest %>% with(clm(formula = XRchest ~
  Age + OverallPos + Ethnicity + Sex +
  RR + GCS + Temperature + HR + SystolicBP +
  Neutrophils + DDimer + CRP + Troponin +
  Sats + Admitted + AdmittedToITU +
  ThirtyDayFUTwo + Dyspnoea + Comorbidity))
P <- pool(object = XRchestmatchimpord[["analyses"]])
multivarXRchestord = multivarXRchestord <- P %>%
  fit2df(estimate_name = "OR (multivariate)",
  exp = TRUE)
multivarXRchestord

```

9.1 Forest Plots

9.1.0.0.1 Creates forest plots for post matched regression tables above:

```

Figure1Forest <- read_excel("Figure1Forest.xlsx",
  col_types = c("text", "numeric", "numeric",
  "numeric", "text", "text"))

tabletext1 <- cbind(Figure1Forest$explanatory,
  Figure1Forest$summary)
forestplot(tabletext1, Figure1Forest$Mean,
  Figure1Forest$Lower, Figure1Forest$Upper,
  is.summary = FALSE, clip = c(0, 2), xlab = "<U+2190> Decreased Odds SARS-
  CoV 2 Increased Odds SARS-CoV 2 <U+2192>",
  zero = 1, cex = 0.9, lineheight = unit(6,
  "mm"), boxsize = 0.4, colgap = unit(6,
  "mm"), lwd.ci = 2, ci.vertices = TRUE,
  ci.vertices.height = 0.4, title = "Odds Ratio of Positivity for SARS-CoV 2
  by RT-PCR",
  txt_gp = fpTxtGp(label = gpar(cex = 1.25),
  ticks = gpar(cex = 1.1), xlab = gpar(cex = 1.2),

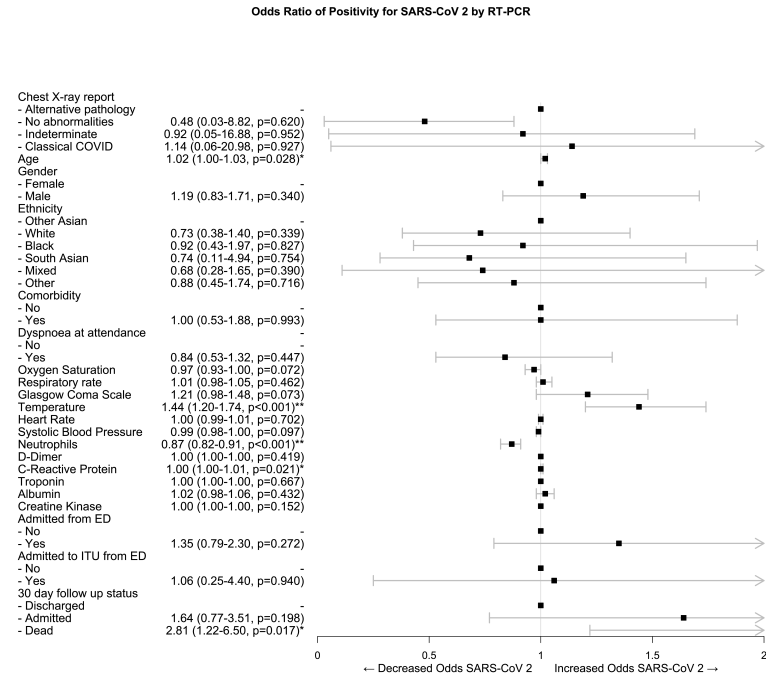
```


42

9 Pooled Regression after Multi...

```
title = gpar(cex = 1.2)), graphwidth = unit(200,
"mm"))
```

9.1.0.0.2 Figure 2:



9.1.0.0.3 Figure 3 (XR dependent):

```
Figure2Forest <- read_excel("Figure2Forest.xlsx",
col_types = c("text", "numeric", "numeric",
"numeric", "text", "text"))

tabletext2<-cbind(Figure2Forest$explanatory,Figure2Forest$summary)
forestplot (tabletext2, Figure2Forest$Mean,
Figure2Forest$Lower, Figure2Forest$Upper, is.summary = FALSE,
clip = c(0, 2),
```

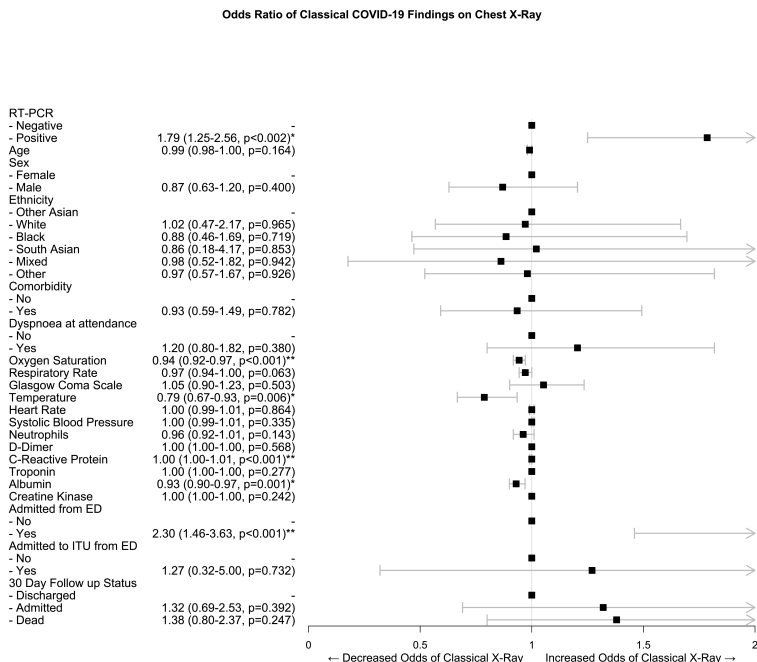
9.1 Forest Plots

43

```

xlab="\u2190 Decreased Odds of Classical X-Ray      Increased Odds
of Classical X-Ray \u2192",
zero=1, cex=0.9, lineheight = unit(6,"mm"), boxsize=0.5,
colgap=unit(6,"mm"),
lwd.ci=2, ci.vertices=TRUE, ci.vertices.height = 0.4,
title="Odds Ratio of Classical COVID-19 Findings on Chest X-Ray",
txt_gp=fpTxtGp(label=gpar(cex=1.25),
ticks=gpar(cex=1.1),
xlab=gpar(cex = 1.2),
title=gpar(cex = 1.2)),
graphwidth = unit(200,"mm")
)

```



9.2 Correlation Matrix

9.2.0.0.1 This section creates a plot of correlation between all the variables in the raw data

```
library(corrplot)
library(Hmisc)
```

9.2.0.0.2 Relevel factors so relevant value is first

```
data$XRPositive <- relevel(data$XRPositive,
  "Negative")

data$Admitted <- relevel(data$Admitted, "Discharged")
data$AdmittedToITU <- relevel(data$AdmittedToITU,
  "No")
```

9.2.0.0.3 New variable for correlation matrix

```
cor <- data
```

9.2.0.0.4 Remove variables with high missings/ data which won't work e.g. date, RT-PCR removed as it only represents initial ED swab, OverallPos used instead as this includes susequent swabs in 30 days

```
cor <- subset(data, select = -c(CT, DateOfDeath, DateOfDischarge, RTPCR,
  DateOfVisit, DateOfSymptomOnset, FollowUpPos, TimeToDeath, NEWS))'
```

9.2.0.0.5 Format and re-name values

```
cor$CTPositive <- ifelse(cor$CTBSTI == "1",
  "Positive", "Negative")
cor$CTPositive <- as.factor(cor$CTPositive)
cor$CTPositive <- relevel(cor$CTPositive,
```

9.2 Correlation Matrix

45

```
"Negative")
cor$Death <- as.factor(ifelse(cor$ThirtyDayFU ==
  "4", "Dead", "Alive"))
cor$Death <- relevel(cor$Death, "Alive")
cor$OverallPos <- as.factor(cor$OverallPos)
cor <- sapply(cor, as.numeric)
```

9.2.0.0.6 Create new numerical correlation matrix

```
cormatrixall <- cor(cor, method = "spearman",
  use = "pairwise.complete.obs")
```

9.2.0.0.7 This variable also contains p-values so identification of only significant correlations is possible:

```
cormatrixall2 <- rcorr(as.matrix(cor), type = "spearman")
```

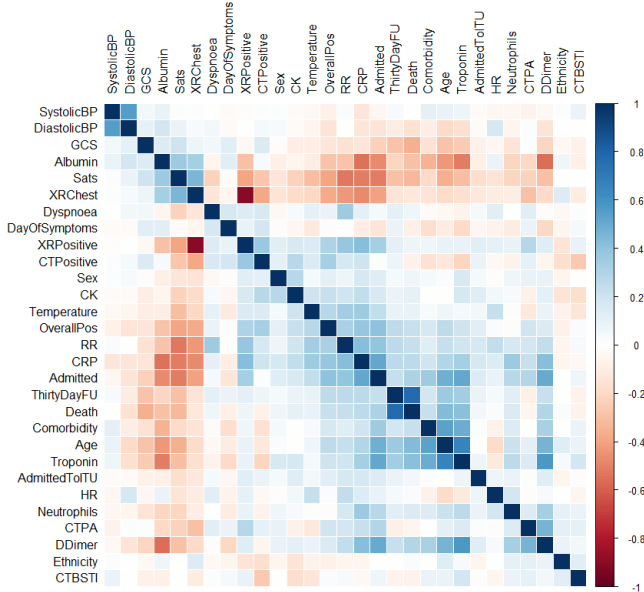
9.2.0.0.8 Function to create and format correlation matrix plot

```
corrplot(cormatrixall2$r, method = "color",
  type = "full", order = "hclust", p.mat = cormatrixall2$p,
  sig.level = 0.05, insig = "blank", tl.col = "black",
  outline = "white", title = "Correlation Matrix of Explanatory and Outcome
  Variables",
  line = -1, cex.main = 2, adj.main = 0.5)
```

46

9 Pooled Regression after Multi...

Correlation Matrix of Explanatory and Outcome Variables



9.3 STARD Flow Diagram

9.3.0.0.1 See instructions from <https://www.r-bloggers.com/flow-charts-in-r/>

9.3.0.0.2 Produces flow charts in Figure 1, (images need to be stretched out, output as svgs)

```
library(grid)
library(Gmisc)

grid.newpage()
# set some parameters to use repeatedly
leftx <- 0.25
```

9.3 STARD Flow Diagram

47

```

1 midx <- 0.5
2 rightx <- 0.75
3 width <- 0.4
4 gp <- gpar(fill = "white")
5 # create boxes
6 (totalattendance <- boxGrob("Number of Patients Attending Emergency Department
7 (ED) in Study Period\n n = 1862",
8 x = midx, y = 0.9, box_gp = gp, width = 0.7))
9
10 (numberwithxr <- boxGrob("Total Number of Patients with Chest X-ray\n n =
11 1772",
12 x = midx, y = 0.75, box_gp = gp, width = width))
13 # connect boxes like this
14 connectGrob(totalattendance, numberwithxr,
15 "v")
16
17 (numberwithoutxr <- boxGrob("No Chest X-ray\n n = 90",
18 x = rightx, y = 0.825, box_gp = gp, width = unit(2,
19 "inch"), height = 0.05))
20
21 connectGrob(totalattendance, numberwithoutxr,
22 "-")
23
24 (XRPos <- boxGrob("Chest X-ray Positive for COVID-19 \n n = 750",
25 x = leftx, y = 0.6, box_gp = gp, width = width))
26 (XRNeg <- boxGrob("Chest X-ray Negative for COVID-19\n n = 1022",
27 x = rightx, y = 0.6, box_gp = gp, width = width))
28
29 connectGrob(numberwithxr, XRPos, "N")
30 connectGrob(numberwithxr, XRNeg, "N")
31
32 (RTPCRXRPos <- boxGrob("Chest X-Ray Positive with RT-PCR swab\n n = 625",
33 x = leftx, y = 0.4, box_gp = gp, width = width))
34 (RTPCRXRNeg <- boxGrob("Chest X-Ray Negative with RT-PCR swab \n n = 573",
35 x = rightx, y = 0.4, box_gp = gp, width = width))
36
37 connectGrob(XRPos, RTPCRXRPos, "N")
38 connectGrob(XRNeg, RTPCRXRNeg, "N")
39
40 (NoRTPCRXRPos <- boxGrob("No RT-PCR Swab\n n = 125",
41 x = 0.4, y = 0.5, box_gp = gp, width = unit(1.5,
42 "inch")))
43 (NoRTPCRXRNeg <- boxGrob("No RT-PCR Swab\n n = 449",
44 x = 0.9, y = 0.5, box_gp = gp, width = unit(1.5,
45 "inch")))
46
47 connectGrob(XRPos, NoRTPCRXRPos, "-")
48 connectGrob(XRNeg, NoRTPCRXRNeg, "-")
49
50 (MatchedXRPos <- boxGrob("Chest X-Ray Positive \n after Propensity Score
51 Matching\n n = 430",
52 x = leftx, y = 0.225, box_gp = gp, width = width))
53 (MatchedXRNeg <- boxGrob("Chest X-Ray Negative \n after Propensity Score
54 Matching \n n = 430",
55 x = 0.65, y = 0.25, box_gp = gp, width = unit(4.2,
56 "inch")))
57
58 connectGrob(RTPCRXRPos, MatchedXRPos, "N")
59 connectGrob(RTPCRXRNeg, MatchedXRNeg, "N")

```

```

(UnmatchedXRPos <- boxGrob("Unmatched\n n = 195",
  x = 0.4, y = 0.325, box_gp = gp, width = unit(1.5,
  "inch"))))
(UnmatchedXRNeg <- boxGrob("Unmatched\n n = 143",
  x = 0.9, y = 0.325, box_gp = gp, width = unit(1.5,
  "inch"))))

connectGrob(RTPCRXRPos, UnmatchedXRPos, "-")
connectGrob(RTPCRXRNeg, UnmatchedXRNeg, "L")

(DiagXRPositive <- boxGrob("COVID-19 Positive n=305\n COVID-19 Negative n=125",
  x = leftx, y = 0.1, box_gp = gp, width = width))
(DiagXRNegative <- boxGrob("COVID-19 Positive n=243 \n COVID-19 Negative
n=187",
  x = rightx, y = 0.1, box_gp = gp, width = width))

connectGrob(MatchedXRPos, DiagXRPositive,
  "N")
connectGrob(MatchedXRNeg, DiagXRNegative,
  "vertical")

(XRInd <- boxGrob("Chest X-Ray Indeterminate \n n = 197",
  x = 0.88, y = 0.25, box_gp = gp, width = unit(2.5,
  "inch"))))

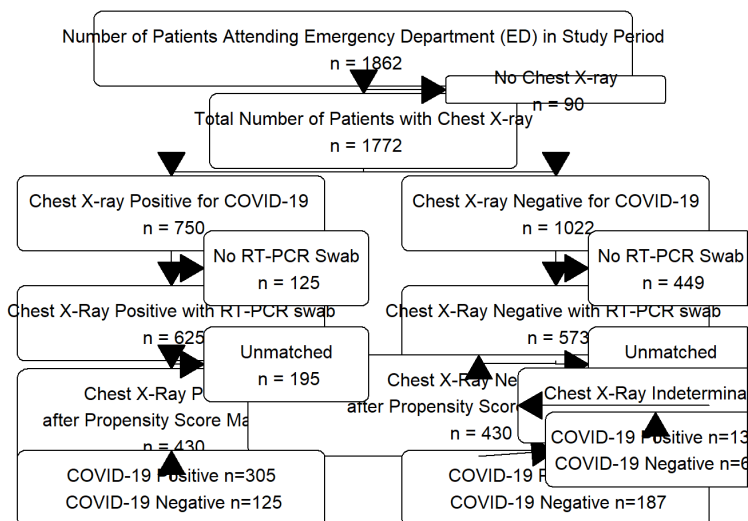
connectGrob(MatchedXRNeg, XRInd, "horizontal")

(DiagXRInd <- boxGrob("COVID-19 Positive n=136\n COVID-19 Negative n=63",
  x = 0.88, y = 0.17, box_gp = gp, width = unit(2,
  "inch"))))
connectGrob(XRInd, DiagXRInd, "vertical")

```

9.3 STARD Flow Diagram

49



```
##### CT Flow Chart####
```

```
grid.newpage()
(totalattendance <- boxGrob("Number of Patients Attending Emergency Department
(ED) in Study Period\n n = 1862",
  x = midx, y = 0.9, box_gp = gp, width = 0.7))

(numberwithCT <- boxGrob("Total Number with Chest Computed Tomography (CT)\n n
= 319",
  x = midx, y = 0.75, box_gp = gp, width = width))
connectGrob(totalattendance, numberwithCT,
  "vertical")

(numberwithoutCT <- boxGrob("No Chest CT\n n = 1543",
  x = rightx, y = 0.825, box_gp = gp, width = unit(2,
  "inch"), height = 0.05))

connectGrob(totalattendance, numberwithoutCT,
  "-")

(CTPos <- boxGrob("CT Positive for COVID-19 \n n = 232",
  x = leftx, y = 0.6, box_gp = gp, width = width))
(CTNeg <- boxGrob("CT Negative for COVID-19\n n = 87",
  x = rightx, y = 0.6, box_gp = gp, width = width))

connectGrob(numberwithCT, CTPos, "N")
connectGrob(numberwithCT, CTNeg, "N")

(RTPCRCTPos <- boxGrob("CT Positive with RT-PCR swab\n n = 217",
  x = leftx, y = 0.4, box_gp = gp, width = width))
```

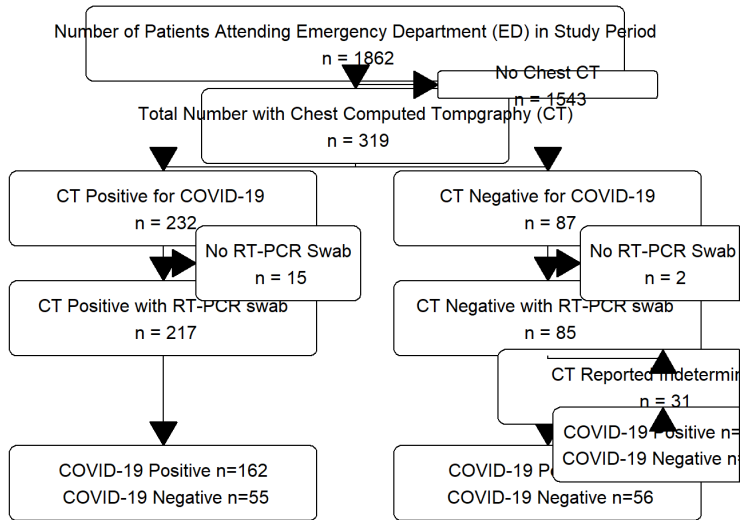

50

9 Pooled Regression after Multi...

```
(RTPCRCTNeg <- boxGrob("CT Negative with RT-PCR swab \n n = 85",  
  x = rightx, y = 0.4, box_gp = gp, width = width))  
  
connectGrob(CTPos, RTPCRCTPos, "N")  
connectGrob(CTNeg, RTPCRCTNeg, "N")  
  
(NoRTPCRCTPos <- boxGrob("No RT-PCR Swab\n n = 15",  
  x = 0.4, y = 0.5, box_gp = gp, width = unit(1.5,  
  "inch")))  
(NoRTPCRCTNeg <- boxGrob("No RT-PCR Swab\n n = 2",  
  x = 0.9, y = 0.5, box_gp = gp, width = unit(1.5,  
  "inch")))  
  
connectGrob(CTPos, NoRTPCRCTPos, "-")  
connectGrob(CTNeg, NoRTPCRCTNeg, "-")  
  
(DiagCTPositive <- boxGrob("COVID-19 Positive n=162\n COVID-19 Negative n=55",  
  x = leftx, y = 0.1, box_gp = gp, width = width))  
(DiagCTNegative <- boxGrob("COVID-19 Positive n=29\n COVID-19 Negative n=56",  
  x = rightx, y = 0.1, box_gp = gp, width = width))  
  
connectGrob(RTPCRCTPos, DiagCTPositive, "N")  
connectGrob(RTPCRCTNeg, DiagCTNegative, "N")  
  
(CTInd <- boxGrob("CT Reported Indeterminate \n n = 31",  
  x = 0.9, y = 0.275, box_gp = gp, width = unit(3,  
  "inch")))  
  
connectGrob(RTPCRCTNeg, CTInd, "N")  
  
(DiagCTInd <- boxGrob("COVID-19 Positive n=16\n COVID-19 Negative n=15",  
  x = 0.9, y = 0.17, box_gp = gp, width = unit(2,  
  "inch")))  
connectGrob(CTInd, DiagCTInd, "vertical")
```

9.3 STARD Flow Diagram

51



```

24  ### Labels###
25  grid.newpage()
26  (indextest <- boxGrob("Index Tests", x = midx,
27    y = 0.9, box_gp = gpar(fill = "light blue"),
28    width = 0.7))
29  (reftest <- boxGrob("Index Tests and Reference Standards",
30    x = midx, y = 0.4, box_gp = gpar(fill = "light blue"),
31    width = 0.7))
32  (finaldiag <- boxGrob("Final Diagnoses",
33    x = midx, y = 0.1, box_gp = gpar(fill = "light blue"),
34    width = 0.7))
  
```

52

9 Pooled Regression after Multi...

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

Index Tests

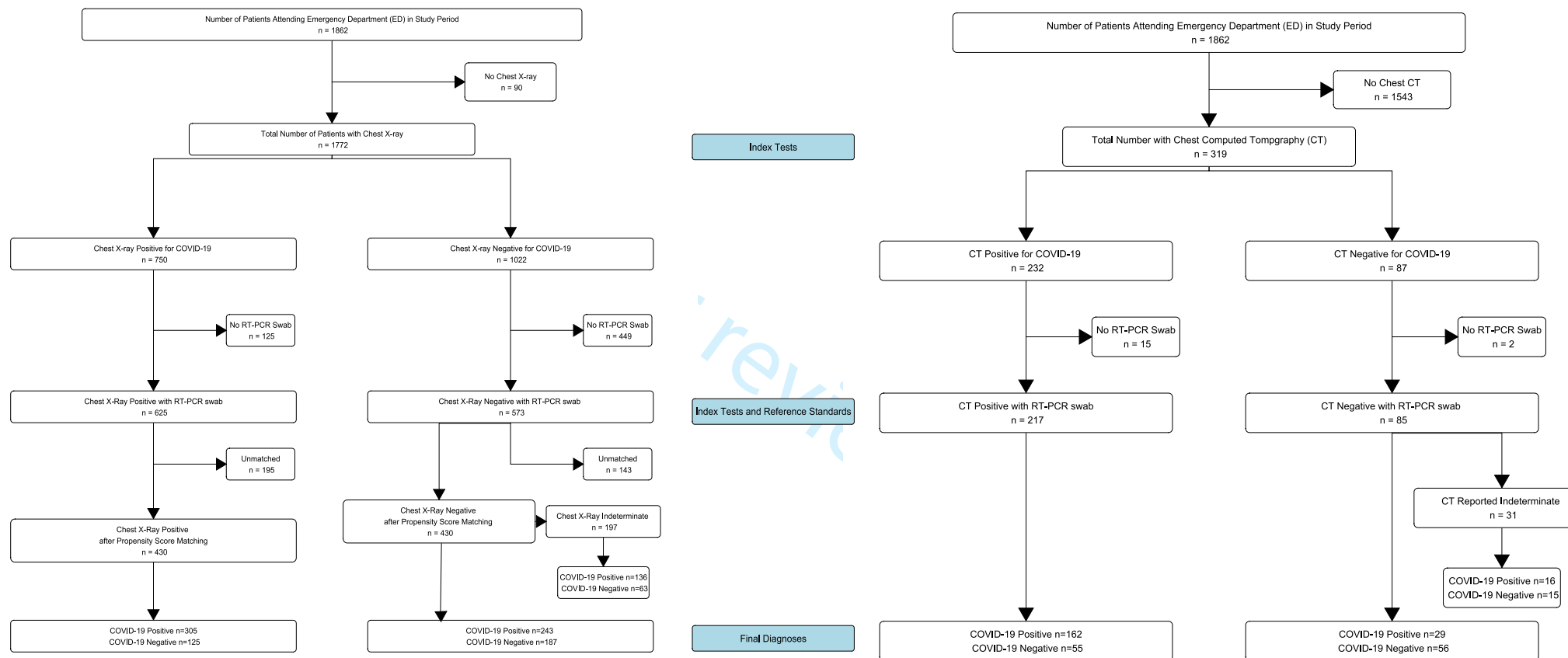
Index Tests and Reference Standards

Final Diagnoses

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	5
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
<i>Participants</i>	6	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	5,20
	11	Rationale for choosing the reference standard (if alternatives exist)	N/A
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	5
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	20
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	12
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	6,7
	15	How indeterminate index test or reference standard results were handled	5
	16	How missing data on the index test and reference standard were handled	N/A, excluded
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	N/A
	18	Intended sample size and how it was determined	7
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	22, diagram below
	20	Baseline demographic and clinical characteristics of participants	21
	21a	Distribution of severity of disease in those with the target condition	21
	21b	Distribution of alternative diagnoses in those without the target condition	N/A
	22	Time interval and any clinical interventions between index test and reference standard	N/A
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	22
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	22
	25	Any adverse events from performing the index test or the reference standard	N/A
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12
	27	Implications for practice, including the intended use and clinical role of the index test	14
OTHER INFORMATION			
	28	Registration number and name of registry	N/A
	29	Where the full study protocol can be accessed	N/A
	30	Sources of funding and other support; role of funders	N/A



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Supplementary Figure- STARD Flow Diagram

For peer review only

BMJ Open

Diagnostic Accuracy of X-ray versus CT in COVID-19: A Propensity Matched Database Study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-042946.R2
Article Type:	Original research
Date Submitted by the Author:	30-Sep-2020
Complete List of Authors:	Borakati, Aditya; Royal Free Hospital, Emergency Department; University College London, Division of Surgery and Interventional Science Perera, Adrian; Royal Free Hospital, Emergency Department Johnson, James; Royal Free Hospital, Emergency Department Sood, Tara; Royal Free Hospital, Emergency Department
Primary Subject Heading:	Emergency medicine
Secondary Subject Heading:	Radiology and imaging, Medical management, Infectious diseases, Respiratory medicine, Diagnostics
Keywords:	COVID-19, Chest imaging < RADIOLOGY & IMAGING, ACCIDENT & EMERGENCY MEDICINE, GENERAL MEDICINE (see Internal Medicine), Diagnostic radiology < RADIOLOGY & IMAGING

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Diagnostic Accuracy of X-ray versus CT in COVID-19: A Propensity Matched Database Study

Authors: Aditya Borakati (0000-0003-0457-4944), Adrian Perera (0000-0002-1753-2586), James Johnson (0000-0001-7796-659X), Tara Sood (0000-0003-1620-916X)

Emergency Department, Royal Free Hospital, Pond Street, London, UK NW3 2QG

Aditya Borakati, Academic Foundation Doctor, Division of Surgery and Interventional Science, Royal Free Campus, University College London

Adrian Perera, Emergency Medicine and Intensive Care Medicine Specialist Registrar, Emergency Department, Royal Free Hospital

James Johnson, Foundation Doctor, Emergency Department, Royal Free Hospital

Tara Sood, Consultant in Emergency Medicine, Emergency Department, Royal Free Hospital

Correspondence:

Dr Aditya Borakati

Academic Foundation Doctor

Emergency Department, Royal Free Hospital, London, UK NW3 2QG

a.borakati@doctors.org.uk

Author contribution (CRediT) statement:

Aditya Borakati: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project Administration

Adrian Perera: Conceptualization, Methodology, Investigation, Writing- Review & Editing, Supervision, Project Administration

James Johnson: Investigation

Tara Sood: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project Administration

Aditya Borakati is the overall guarantor of this work.

Word count: 4236

Abstract

Objectives: To identify the diagnostic accuracy of common imaging modalities, chest X-ray (CXR) and computed tomography (CT) for diagnosis of COVID-19 in the general emergency population in the UK and to find the association between imaging features and outcomes in these patients.

Design: Retrospective analysis of electronic patient records

Setting: Tertiary academic health science centre and designated centre for high consequence infectious diseases in London, UK.

Participants: 1,198 patients who attended the emergency department with paired RT-PCR swabs for SARS-CoV 2 and CXR between 16th March and 16th April 2020

Main outcome measures: Sensitivity and specificity of CXR and CT for diagnosis of COVID-19 using the British Society of Thoracic Imaging reporting templates. Reference standard was any reverse transcriptase polymerase chain reaction (RT-PCR) positive naso-oropharyngeal swab within 30 days of attendance. Odds ratios of CXR in association with vital signs, laboratory values and 30-day outcomes were calculated.

Results: Sensitivity and specificity of CXR for COVID-19 diagnosis were 0.56 (95% CI 0.51-0.60) and 0.60 (95% CI 0.54-0.65), respectively. For CT scans these were 0.85 (95% CI 0.79-0.90) and 0.50 (95% CI 0.41-0.60), respectively. This gave a statistically significant mean increase in sensitivity with CT compared with CXR, of 29% (95% CI 19%-38%, $p < 0.0001$). Specificity was not significantly different between the two modalities.

Chest X-ray findings were not statistically significantly or clinically meaningfully associated with vital signs, laboratory parameters or 30-day outcomes.

Conclusions: Computed tomography has substantially improved diagnostic performance over CXR in COVID-19. CT should be strongly considered in the initial assessment for suspected COVID-19. This gives potential for increased sensitivity and considerably faster turnaround time, where capacity allows and balanced against excess radiation exposure risk.

Strengths and limitations

-Large, appropriately powered, study population consisting of all patients attending the emergency department rather than those solely with confirmed COVID-19; this allowed assessment of specificity for the imaging modalities and applicability to the general population who may attend medical personnel with other complaints, but have underlying SARS-CoV 2 infection

1
2
3
4 -Comprehensive statistical analyses were conducted to address confounding in
5 reporting of X-rays including propensity score matching and logistic regression to give a
6 'doubly robust' model
7
8

9
10 -Low amount of missing data and for secondary covariates only; multiple imputation was
11 performed with a good fit, however, observed data would be preferable to imputed data
12

13
14 -Single centre, retrospective study; potential for inter-reporter and inter-centre variability
15 in reporting
16

17
18
19 -Large proportion of patients excluded due to not having an RT-PCR swab,
20 predominantly, those with imaging reported as negative, this may bias the results
21 towards increased sensitivity and specificity
22
23

24
25 **Key words:** X-Rays, Computed Tomography, COVID-19, severe acute respiratory
26 syndrome coronavirus 2, Emergency Medicine, Diagnostic Imaging
27

28
29 **Statistical review:** The statistical methods in this manuscript and associated code have
30 been reviewed by Dr Federico Ricciardi of the Department of Statistical Science at
31 University College London and confirmed as robust and accurate.
32

33
34 **Ethical approval:** This study was registered with the local institutional review board
35 as a service evaluation using anonymised data only. No formal ethics committee
36 review was required.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Declarations of Interests:** The authors have no relevant conflicts of interest to declare.
4 All authors have completed the [Unified Competing Interest form](#) (available on request
5 from the corresponding author) and declare: no support from any organisation for the
6 submitted work; no financial relationships with any organisations that might have an
7 interest in the submitted work in the previous three years, no other relationships or
8 activities that could appear to have influenced the submitted work.
9
10

11 **Transparency declaration:** The lead author (AB) affirms that this manuscript is an
12 honest, accurate, and transparent account of the study being reported; that no
13 important aspects of the study have been omitted; and that any discrepancies from the
14 study as planned (and, if relevant, registered) have been explained.
15
16

17 **Funding:** No funding was received for this study.
18

19 **Copyright:** The Corresponding Author has the right to grant on behalf of all authors
20 and does grant on behalf of all authors, a worldwide licence to the Publishers and its
21 licensees in perpetuity, in all forms, formats and media (whether known now or created
22 in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii)
23 translate the Contribution into other languages, create adaptations, reprints, include
24 within collections and create summaries, extracts and/or, abstracts of the Contribution,
25 iii) create any other derivative work(s) based on the Contribution, iv) to exploit all
26 subsidiary rights in the Contribution, v) the inclusion of electronic links from the
27 Contribution to third party material where-ever it may be located; and, vi) licence any
28 third party to do any or all of the above
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

SARS-CoV 2 and its resulting disease, COVID-19, have propagated exponentially worldwide, with over 10 million cases in 188 countries at the time of writing [1,2].

The gold standard for diagnosis of the virus is the detection of viral RNA through reverse transcriptase polymerase chain reaction (RT-PCR) of respiratory tract samples. However, this method has several limitations including: (1) low sensitivity at 59-71% [3,4], (2) relatively slow turnaround times ranging from a few hours to several days [5], (3) high expense and (4) limited capacity for testing in many countries.

Computed tomography (CT) has been shown to be more sensitive than RT-PCR for diagnosis of COVID-19 [3,4], while being significantly faster and cheaper. This comes with a large radiation dose and capacity is still lacking in many countries.

Plain film chest X-ray (CXR) is ubiquitous worldwide, with a 30-70x lower dose of radiation[6] and is commonly performed as an initial investigation in COVID-19.

Studies have so far only evaluated imaging in those with confirmed infection, it is therefore, not possible to calculate the specificity of these modalities. In the context of the global pandemic, infection may be widespread in the community, often with subclinical infection [7,8]. A reliable and rapid method to detect infection in the general population, who may present to medical personnel with other complaints, is needed.

Despite its extensive use, the specificity and sensitivity of CXR in the general emergency population for diagnosis of COVID-19 is unknown, nor how imaging features correlate with severity.

This study evaluated the performance of CXR in diagnosing COVID-19 in the emergency department (ED) of a tertiary care hospital.

Methods

This study was conducted at the Royal Free Hospital, London, UK, an academic health science centre and nationally designated centre for High Consequence Infectious Diseases [9].

All individuals attending the emergency department who had paired posterior-anterior chest radiographs and RT-PCR nasopharyngeal swabs for COVID-19 at the time of initial attendance between 16th March 2020 and 16th April 2020 were included.

All chest radiographs were reported by a Consultant Radiologist and rated on an ordinal scale for probability of COVID-19: Alternative pathology identified, not COVID-19; Clear chest, unlikely COVID; Indeterminate findings for COVID-19; Classical findings of COVID-19, based on the British Society of Thoracic Imaging's (BSTI) reporting templates (table 1) [10]. These were reported prior to RT-PCR results being available.

RT-PCR of swabs were performed in laboratories either at our centre or at a public health laboratory (PHE Collindale, UK), according to published national standard operating procedures [11]. Subsequent RT-PCR swabs taken within 30 days of initial ED attendance were also included.

CT scans performed within 30 days of attendance were retrieved. These were also reported according to the BSTI template. CT pulmonary angiogram was performed in the ED if the D-dimer was >5000 to exclude pulmonary emboli as per the locally agreed protocol. Subsequent CT chest imaging (whether pulmonary angiogram, contrast or non-contrast) was performed on the basis of clinical suspicion.

Prospectively recorded data was extracted from the Cerner Millennium electronic patient record system (Cerner Corp., Kansas City, MO).

Primary Outcome

The primary outcome is sensitivity and specificity of initial CXR, where it is reported as having classic COVID-19 features in the ED. This is compared with RT-PCR swab as the reference standard for diagnosis of COVID-19.

In the event of multiple RT-PCR swabs during one attendance, a single positive swab was taken as an overall positive test during one admission.

Secondary Outcomes

In those patients who also had CT scans of the thorax, the diagnostic accuracy was compared with CXR, with RT-PCR again as the reference standard. Sensitivity and specificity of CXR when X-rays reported as indeterminate or atypical for COVID-19 were classed as positive was also calculated.

Chest x-ray findings were correlated with vital signs at attendance and blood results, including: neutrophil counts, D-dimer and C-reactive protein, which have been associated with poor prognosis in COVID-19 [12]. Hazard ratios for clinical outcomes including direct admission to the intensive treatment unit (ITU) from ED and 30-day mortality rates were also calculated for CXR reporting categories.

Statistical Analysis

In the event of missing data, multiple imputation was conducted using a Predictive Mean Matching algorithm, via the MICE R package, as described previously [13]. Briefly, this uses a linear regression model (or logistic regression model for categorical data), to find a random value based on already observed data, to replace missing fields [14]. Variables without missing data fields were not modified. The number of imputed datasets was similar in number to the percentage of missing data as suggested by White and colleagues [15]. Balance diagnostics with density plots are available in supplementary file 1, adequate balance was assessed via visual inspection of imputed distributions with respect to the original dataset.

The propensity for a CXR being reported as positive or negative for COVID-19 was calculated for several plausible covariates that may influence image characteristics such as Age, Gender, Ethnicity, pre-existing morbidities and the respiratory rate of the patient using a generalised linear model [16]. X-ray positive and negative groups were then matched in each imputed dataset using the nearest neighbour algorithm, with a calliper of 0.2 of the propensity score standard deviation, without replacement and in random sequential order to obtain a 1:1 match as described elsewhere [17].

The balance of the match data was assessed quantitatively with mean differences of covariates in each of the X-ray groups pre- and post-matching, with a difference of less than 0.1% considered a good match (supplementary figures 1, 2). Visual inspection of matches was also conducted to ensure balance (supplementary figures 2, 3 and 4).

After matching, outcome data were adjusted for covariates including age, gender, ethnicity and presence of co-morbidities as well as C-reactive protein, D-dimer, troponin and vital signs. This was achieved by generalised linear regression for continuous outcome data, binomial logistic regression for binary categorical outcomes, or ordinal logistic regression in the case of CXR where it is the outcome variable.

These regression models were run on each imputed dataset and outcomes were pooled together across each imputed data set according to Rubin's rules [18] to give an overall estimate.

Diagnostic Accuracy Statistics

Chest X-rays reported as classical for COVID-19 as per the BSTI guidelines were considered a positive test in the primary analysis. In a secondary analysis X-rays reported as 'Indeterminate' or 'Atypical' for COVID-19 were also considered positive. All other reports were classified as a negative test. These were compared to nasopharyngeal aspirate RT-PCR results, which were taken as the gold standard for diagnosis of COVID-19. Where more than one swab was taken during the study period (up to 30 days after initial attendance), a single positive result was taken as a positive result for calculation of diagnostic accuracy statistics.

Sensitivity, specificity, predictive values and diagnostic accuracy were calculated using the propensity matched data after imputation and pooled across imputed datasets with 95% confidence intervals. Apparent and true prevalence based on this dataset are also given for interpretation of the predictive values.

Chest CTs were also reported according to the BSTI guidelines as with X-ray. Diagnostic statistics were calculated on raw, unmatched and non-imputed data (due to a low volume of

1
2
3 data for imputation and matching) in the same manner as X-ray. Mean differences and 95%
4 confidence intervals between CT and X-ray for each of the diagnostic statistics are given, with a
5 p-value calculated from the confidence intervals.
6

7 Agreement between the modalities was assessed on the unmatched dataset, in the sample
8 where CT, CXR and RT-PCR were all available using Cohen's (for two group agreement) and
9 Fleiss' Kappa (when all 3 are compared).
10

11 **Data Presentation**

12
13 Descriptive statistics are given as means and standard deviations for normally distributed data
14 and as medians and interquartile ranges for non-normally distributed data, before and after
15 matching and multiple imputation (for the latter these statistics are pooled across imputations).
16

17 Association of explanatory variables with SARS-CoV 2 and Chest X-ray findings are given as
18 odds ratios in uni- and multi-variate configurations.
19

20 Data was considered statistically significant if $p < 0.05$. Given the large number of analyses in
21 this paper, data is separately highlighted if $p < 0.001$ as a secondary threshold to address the
22 potential for false positives with multiple testing.
23

24 Analyses were conducted using R 4.0.0 (R Foundation for Statistical Computing, Vienna,
25 Austria) and code for the analyses is given in supplementary file 2.
26

27 **Sample size calculation**

28
29 In this study, the lower confidence interval for sensitivity of CXR as reported by Wong et al.[19]
30 (56%) was used as an estimate of likely sensitivity for COVID-19 . A power of 80% at an alpha
31 of 0.05 was used to calculate the sample size for sensitivities and specificities of 56%. This
32 gave an estimated sample size of 165 in each of the COVID-19 negative and positive groups
33 by RT-PCR (total 330).
34

35 **Ethical approval**

36
37 This study was registered with the local institutional review board as a service evaluation using
38 anonymised data only. No formal ethics committee review was required.
39

40 **Reporting Guidelines**

41 This study is reported according to the STARD guidelines [20] for diagnostic accuracy studies.
42
43

44 **Patient and Public Involvement**

45
46
47 Patients and the public were not involved in the design, conduct or dissemination of this study.
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

7

For peer review only

Results

1,198 eligible patients with both CXR and RT-PCR were identified in the study period (figure 1). Their characteristics, stratified by positivity for SARS-CoV 2 infection by RT-PCR is summarized in table 2. This showed that those with confirmed SARS-CoV 2 infection were more likely to be male, older (mean age 66.2 vs 62.7), have lower saturations, higher respiratory rates, whilst being more likely to be admitted and die within 30 days. There was a significant association with X-ray images and SARS-CoV 2 at baseline, with 59.6% having classic imaging features of COVID-19 in those with positive swabs versus 39.1% in those with negative swabs. There was 8.6% missing data overall in the dataset when variables with >50% missing data were removed and 15 imputations were performed on these remaining variables only.

After multiple imputation for missing data and pooled propensity score matching for plausible covariates that may affect CXR reporting, there were 430 patients in each of the X-ray positive and X-ray negative groups, for a total of 860 patients. Adequate balance was achieved for relevant covariates with a mean difference of <0.1 between groups (supplementary file 1, table 2).

Computed tomography (CT) was performed in 302 patients with paired RT-PCR during the same time period, with a median serial interval of 4.5 days (inter quartile range 0-17) after the initial attendance in ED and of these 30.1% were within one day of attendance.

Diagnostic Accuracy

The pooled sensitivity and specificity of CXR was 0.56 (95% CI 0.51-0.60) and 0.60 (95% CI 0.54-0.65), respectively (table 3). This gave an overall diagnostic accuracy of 0.57 (95% CI 0.54-0.61) for CXR.

In comparison, sensitivity and specificity for CT was 0.85 (95% CI 0.79-0.90) and 0.50 (95% CI 0.41-0.60), respectively. This gave a statistically significant mean increase in sensitivity with CT compared with CXR by 29% (95% CI 19%-38%, $p < 0.0001$). Specificity was not significantly different between the two modalities. Diagnostic accuracy and negative predictive values were also significantly increased with CT at 0.15 and 0.22, respectively, while the negative likelihood ratio was significantly decreased at -0.44. This shows that the post-test odds of being negative for SARS-CoV 2 by RT-PCR with a negative CT is significantly lower.

Taking X-rays reported as indeterminate as positive increased the sensitivity of CXR to 0.80 (95% CI 0.77-0.84), however reduced specificity to 0.40 (95% CI 0.35-0.46). When CT scans reported as indeterminate are also considered positive the sensitivity of CT increased to 0.93 (95% CI 0.89-0.96), whilst mean specificity reduced to 0.37 (95% CI 0.28-0.47), although this was not statistically different from when indeterminate CTs are considered negative. Sensitivity of CT remained significantly higher than CXR (when indeterminates are considered positive for both) by 0.13 (95% CI 0.05-0.19, $p < 0.001$), specificity was not significantly different between the two.

When comparing only the unimputed, unmatched subset of data where CT, RT-PCR and CXR were all performed ($n=287$), the agreement between CT and CXR was poor (Cohen's kappa 0.406). Agreement between all three modalities was also poor (Fleiss' kappa 0.361).

Association of CXR with Markers of Severity and Outcomes

Association of covariates with RT-PCR results is shown in table 4 and figure 2. Those who tested positive for SARS-CoV 2 by RT-PCR were significantly more likely to have a classical X-ray (OR 1.79 95% CI 1.25-2.56, $p < 0.002$) as would be expected by the diagnostic accuracy statistics (table 4). When the CXR report is considered as an ordered scale, worsening grades of report were associated more strongly with RT-PCR positivity, with a 1.94 x increase in odds for each grade.

Positive chest X-rays for COVID-19 were significantly associated with lower oxygen saturations (OR 0.94 95% CI 0.92-0.97, $p < 0.001$) and temperatures (2.30 95% CI 1.46-3.63, $p < 0.001$) in the ED following propensity score matching and multivariate regression (table 5 and figure 3).

They also had higher rates of admission to a general ward from the ED (OR 2.30 95% CI 1.46-3.63, $p < 0.001$) but no significant association with 30 day outcomes. There was a statistically significant increase in C-reactive protein with a positive X-ray, however, this is unlikely to be clinically meaningful due to the minimal association (OR 1.00 95% CI 1.00-1.01).

Discussion

This study is the first to report the diagnostic accuracy of CXR and CT in the general emergency population during the COVID-19 pandemic.

We show that CXR has poor sensitivity and specificity for diagnosis of COVID-19, whilst CT has 29% higher sensitivity. Many international radiological guidelines advise against CT scanning for the initial assessment of COVID-19 [21–23] or where there are equivocal CXRs, whilst in other countries CT scanning is performed as a routine first line investigation. Our results suggest that CT should be considered in the initial assessment of COVID-19 and that CXR findings poorly correlate with CT findings in this setting. We also show that indeterminate and non-classical features of COVID-19 significantly increase the sensitivity of these imaging modalities, without a significant decrease in specificity. Further, we demonstrate the limited prognostic value of CXR in COVID-19.

These findings mirror what has previously been reported in the literature on individuals with confirmed COVID-19. Wong et al. [19] showed a sensitivity of 59% for initial X-ray in confirmed COVID-19 infection, similarly initial case series in China also reported a sensitivity of 59.1% [12].

A recent in press article from Italy reported a much higher sensitivity of 89% for CXR in a smaller general emergency population (n=535) without confirmed COVID-19 at attendance [24]. However, this used telephone follow up for clinical symptoms of COVID-19 as a reference standard in individuals with an initial negative RT-PCR swab and appeared to classify any abnormal X-ray as positive, which may inflate this figure. When indeterminate CXRs are counted as positive in this study, the sensitivity would be in line with this Italian data. In the US, a study of patients attending an urgent care centre with confirmed COVID-19, showed a much lower sensitivity at 41.7% for CXR where any abnormality was found on the images [25]. In this study 97/636 reports were re-classified from 'possible pneumonia' to 'normal' on second reading from a radiologist, highlighting the importance of inter-rater agreement and possibly explaining this low estimate.

Computed tomography has been reported in previous studies as being up to 98% sensitive for the diagnosis of COVID-19 in confirmed patients, when RT-PCR is used as the reference standard in confirmed patients [3,4]. These studies used any potential features of COVID-19 (e.g. ground glass opacification, crazy paving) as a positive scan, regardless of spatial distribution or features more characteristic of alternate pathology, unlike the BSTI guidelines used in this study. When we classified indeterminate CTs as positive like these latter studies, our estimates match their sensitivity values.

Consequently, a much lower specificity of 25% was found with initial RT-PCR in previous literature; however, it is reported that 10 out of 15 (67%) of these negatives subsequently tested positive. This would give an adjusted specificity of 75%, considering subsequent swabs as a reference standard, which combined with the wider CIs in these smaller studies, would bring estimates in line with the specificity in this paper. More recent meta-analyses have placed the pooled sensitivity of CT in populations with confirmed COVID-19 only, at 89.76% (95% CI 84.42%-93.84%) [26], in line with the estimates identified here.

There is limited coverage in the literature on association of X-ray findings with clinical and laboratory parameters and outcomes in the COVID-19 pandemic. This study demonstrates that classic appearances of COVID-19 were associated with initial lower saturations and lower

1
2
3 temperature. Volume opacification of the lung fields were not quantified as a surrogate of
4 severity; however, the use of the BSTI grading templates does this somewhat. When the X-ray
5 report is considered as a graded scale from low likelihood of COVID-19 and severity to high
6 likelihood and severity of disease there was no significant difference in association with vital
7 signs or laboratory parameters compared with when the X-ray report is merely considered as
8 a binary positive and negative outcome for COVID-19.
9

10
11 Borghesi and colleagues have devised a X-ray grading system, the Brixia score, for severity in
12 admitted patients with confirmed SARS-CoV 2 infection [27]. They further found a significant
13 increase in the severity of CXR by this scoring system in those who were discharged versus
14 those who died [28,29].
15

16 Here, there were no relevant associations between CXR and laboratory values. This analysis
17 also found no association with positive X-rays and 30 day outcomes after multivariate analyses,
18 unlike Borghese et al. This is also in contrast to Guan et al. who found higher rates of ITU
19 admission and death in those with positive imaging findings. However, these studies analysed
20 only those with confirmed SARS-CoV 2 infection. The divergence observed in this study may be
21 due to classifying those with 'Alternate pathology/ Indeterminate' or 'CVXC3/ CVXC2' as per the
22 BSTI templates, negative for COVID-19 in these analyses. Other studies classified X-rays with
23 any abnormality as a positive for COVID-19. These alternate distributions may still be reflective
24 of underlying COVID-19 and we show significantly higher sensitivity for both CT and CXR when
25 these are classed as positive. It may be that correlating indeterminate X-rays (in addition to
26 classical images) with vitals, laboratory markers and 30 day outcomes would yield significant
27 associations. However this may be unlikely, Xu and Zhang et al. found that those with classical
28 bilateral and diffuse involvement in upper and lower lobes had more severe disease than those
29 without [30,31].
30
31

32 There were a total of 70 confirmed pulmonary emboli (PEs) in our dataset out of 114 CT
33 pulmonary angiograms (61.0%, 5.84% of all patients attending) performed in the emergency
34 department. The incidence of venous thromboembolism is reported as ranging from 20-30% in
35 admitted confirmed SARS-CoV 2 positive patients [32]. Although we have not focused on this
36 cohort of patients in this paper for the sake of brevity and simplicity, this high incidence
37 represents a further advantage for CT over CXR.
38
39

40 CT, even with the absence of contrast has been shown to have strong accuracy in the
41 diagnosis of pulmonary emboli and many imaging features correlate with the presence of
42 pulmonary emboli. Sensitivities of non-contrast CT for diagnosis of PE have been reported at
43 96.9% and specificity at 71.9% [33,34].
44

45 We therefore see the advantages of CT scanning in COVID-19 as threefold over other
46 diagnostic techniques: 1) The rapid turnaround; 2) Increased sensitivity and 3) The possibility to
47 identify pulmonary emboli in COVID-19, which are a significant burden in this group.
48

49 This must be balanced against the excess radiation exposure with CT. Radiation from CT and
50 its association with carcinogenesis is difficult to quantify and no definitive epidemiological
51 studies have confirmed excess risk of cancer[35]. Modern CT scanners and software
52 reconstruction techniques continue to minimise radiation exposure and many ways of shielding
53 parts of the body from radiation also exist. Nevertheless, the excess risk of lifetime cancer is
54 estimated at 1 per 5,000 CT examinations[36].
55
56
57

Strengths and Limitations

This study is the largest conducted on imaging in the COVID-19 pandemic and one of the only studies conducted in the general population during the pandemic rather than only in confirmed patients. This enables greater applicability to the clinical setting where the diagnosis is uncertain, in addition to being able to calculate specificity, which is not possible in most studies. This study was planned to be powered to detect a sensitivity and specificity of 56% for CXR and greatly exceeded the sample size necessary for this.

Comprehensive statistical analyses were conducted to account for confounders in both factors influencing reporting of CXR and in factors affecting outcomes. The data was collected from prospectively maintained electronic records; however, the retrieval took place retrospectively with its inherent disadvantages. We were not able to collect data on several relevant covariates such as specific comorbidities or markers of severity such as lymphocytes. Furthermore, there was a significant amount of missing data that required multiple imputation to replace, although the fit of this imputed data was good, actual, observed data would be ideal.

Inter-rater reliability of imaging reports was not analysed in this paper and there was the potential for individual radiologists to have greater or lesser accuracy in the diagnosis of COVID-19. The literature has so far suggested a strong degree of agreement between radiologists in reporting of COVID-19 images [28].

The single centre nature of this study further limits generalisability and the potential for inter-hospital disagreement in imaging, in addition to inter-rater disagreement.

Finally, the median time for patients to receive a CT scan was 4.5 days following initial attendance to ED. Thus, the scans may not have been directly comparable to the initial CXR, both because of the progression of disease and because the SARS-CoV 2 status may have been confirmed at this point, biasing the reporting of these scans.

Future Research

Although this study used RT-PCR of nasopharyngeal swabs as a reference standard, newer methods exist for diagnosis of the disease. Serological assays for antibodies against SARS-CoV 2 are increasingly available and may represent a better gold standard in diagnosis for future research [37]. RT-PCR is limited by swabbing technique for nasopharyngeal samples and the fact that the virus is more avid in the lower respiratory tract [38]. However, many patients may not seroconvert prior to death limiting this test to survivors only.

Point of care lung ultrasound is a new technique for diagnosis of COVID-19 which may mitigate many of the issues noted with the modalities discussed so far. It has no radiation, is fast, cheap and may be able to detect lower respiratory tract disease unlike nasopharyngeal swab.

However, there is limited evidence beyond small case series on its diagnostic accuracy [39–41]. Further, like other ultrasound techniques accuracy will likely be operator dependent [42] and experience will need to be built up for robust results in evaluating suspected COVID-19.

Finally, much research has been conducted in the use of artificial intelligence techniques to correctly diagnose COVID-19 based on imaging [43–45]. These techniques would obviate capacity limitations in reporting imaging as well as eliminate inter-reporter variability. However, as with any supervised machine learning technique, large, generalisable datasets, with correctly

1
2
3 pre-classified positive and negative cases (which in turn will depend on a truly
4 accurate reference standard) are needed [46].
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Conclusion

Chest X-ray has poor sensitivity and specificity in diagnosing COVID-19 in the general population during the pandemic. CT scanning has demonstrated excellent sensitivity and should strongly be considered during the pandemic in the initial assessment of COVID-19. This needs to be balanced against the risk of excess radiation with CT, where capacity allows.

Acknowledgements

We would like to thank Scott Wilson from the Royal Free Hospital's clinical practice group analytics department for retrieving the data from the hospital's data warehouse.

We would like to thank Dr Federico Ricciardi of the Department of Statistical Science and PRIMENT Clinical Trials Unit at University College London for reviewing the statistical methods in this study.

1
2
3 **Data availability**
4

5 Anonymised data is available on reasonable request from the corresponding author. Analysis
6 scripts are attached as a supplementary file.
7

8 **Declarations of Interest**
9

10 The authors declare no conflicts of interest.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

References

- 1 COVID-19 Map. Johns Hopkins Coronavirus Resour. Cent. <https://coronavirus.jhu.edu/map.html> (accessed 30 Jun 2020).
- 2 Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020;**0**. doi:10/ggnsjk
- 3 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**. doi:10/ggmw6p
- 4 Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* 2020;**200432**. doi:10/ggnnkj
- 5 Konrad R, Eberle U, Dangel A, et al. Rapid establishment of laboratory diagnostics for the novel coronavirus SARS-CoV-2 in Bavaria, Germany, February 2020. *Eurosurveillance* 2020;**25**. doi:10/ggp6bw
- 6 Lin EC. Radiation Risk From Medical Imaging. *Mayo Clin Proc* 2010;**85**:1142–6. doi:10/c445mk
- 7 Mizumoto K, Kagaya K, Zarebski A, et al. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance* 2020;**25**:2000180. doi:10/ggn4bd
- 8 Arons MM, Hatfield KM, Reddy SC, et al. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. *N Engl J Med* 2020;**382**:2081–90. doi:10/ggszfg
- 9 High consequence infectious diseases (HCID). GOV.UK. <https://www.gov.uk/guidance/high-consequence-infectious-diseases-hcid> (accessed 24 May 2020).
- 10 Desai S. COVID-19 BSTI Reporting templates | The British Society of Thoracic Imaging. Br. Soc. Thorac. Imaging. 2020. <https://www.bsti.org.uk/covid-19-resources/covid-19-bsti-reporting-templates/> (accessed 29 Apr 2020).
- 11 NHS England. Guidance and Standard Operating Procedure: COVID-19 virus testing in NHS Laboratories. 2020. <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 24 May 2020).
- 12 Guan W, Ni Z, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020;**0**:null. doi:10/ggm6dh
- 13 Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. *J Stat Softw* 2011;**45**. doi:10/gdqc9c
- 14 Ginkel JR van, Linting M, Rippe RCA, et al. Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data. *J Pers Assess* 2020;**102**:297–308. doi:10/gftj5w

- 15 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;**30**:377–99. doi:10.1002/sim.4067
- 16 He H, McDermott MP. A robust method using propensity score stratification for correcting verification bias for binary tests. *Biostat Oxf Engl* 2012;**13**:32–47. doi:10/c4jzn6
- 17 Ho DE, Imai K, King G, et al. MatchIt : Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw* 2011;**42**. doi:10/gdwtnq
- 18 Marshall A, Altman DG, Holder RL, et al. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009;**9**:57. doi:10.1186/1471-2288-9-57
- 19 Wong HYF, Lam HYS, Fong AH-T, et al. Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients. *Radiology* 2020;:201160. doi:10/ggqbp4
- 20 Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016;**6**:e012799. doi:10.1136/bmjopen-2016-012799
- 21 Rubin GD, Ryerson CJ, Haramati LB, et al. The Role of Chest Imaging in Patient Management during the COVID-19 Pandemic: A Multinational Consensus Statement from the Fleischner Society. *Radiology* 2020;:201365. doi:10/ggrmg4
- 22 ACR Recommendations for the use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection. <https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection> (accessed 5 Jun 2020).
- 23 British Society of Thoracic Imaging. COVID-19: BSTI STATEMENT AND GUIDANCE. 2020;:1. https://www.bsti.org.uk/media/resources/files/COVID11.3.20_2.pdf (accessed 5 Jun 2020).
- 24 Schiaffino S, Tritella S, Cozzi A, et al. Diagnostic Performance of Chest X-Ray for COVID-19 Pneumonia During the SARS-CoV-2 Pandemic in Lombardy, Italy. *J Thorac Imaging* 2020;**Publish Ahead of Print**. doi:10/ggx268
- 25 Weinstock MB, Echenique A, Russell JW, et al. Chest X-Ray Findings in 636 Ambulatory Patients with COVID-19 Presenting to an Urgent Care Center: A Normal Chest X-Ray Is no Guarantee. ;:10.
- 26 Bao C, Liu X, Zhang H, et al. Coronavirus Disease 2019 (COVID-19) CT Findings: A Systematic Review and Meta-analysis. *J Am Coll Radiol* 2020;**17**:701–9. doi:10/ggr28p
- 27 Borghesi A, Zigliani A, Masciullo R, et al. Radiographic severity index in COVID-19 pneumonia: relationship to age and sex in 783 Italian patients. *Radiol Med (Torino)* 2020;**125**:461–4. doi:10/ggtvwp

- 1
2
3 28 Borghesi A, Maroldi R. COVID-19 outbreak in Italy: experimental chest X-ray scoring system
4 for quantifying and monitoring disease progression. *Radiol Med (Torino)* 2020;**125**:509–13.
5 doi:10/ggtvwn
6
- 7 29 Borghesi A, Zigliani A, Golemi S, et al. Chest X-ray severity index as a predictor of in-
8 hospital mortality in coronavirus disease 2019: A study of 302 patients from Italy. *Int J Infect*
9 *Dis* 2020;**96**:291–3. doi:10.1016/j.ijid.2020.05.021
10
- 11 30 Xu Y-H, Dong J-H, An W-M, et al. Clinical and computed tomographic imaging features
12 of novel coronavirus pneumonia caused by SARS-CoV-2. *J Infect* 2020;**80**:394–400.
13 doi:10/ggqwf3
14
- 15 31 Zhang J-J, Dong X, Cao Y-Y, et al. Clinical characteristics of 140 patients infected with
16 SARS-CoV-2 in Wuhan, China. *Allergy Published Online First*: 19 February 2020.
17 doi:10/ggpx6g
18
- 19 32 Lodigiani C, Iapichino G, Carenzo L, et al. Venous and arterial thromboembolic
20 complications in COVID-19 patients admitted to an academic hospital in Milan, Italy. *Thromb*
21 *Res* 2020;**191**:9–14. doi:10/ggvctf
22
- 23 33 Chien C-H, Shih F-C, Chen C-Y, et al. Unenhanced multidetector computed
24 tomography findings in acute central pulmonary embolism. *BMC Med Imaging*
25 2019;**19**:65. doi:10/ggzg85
26
- 27 34 Mohamed N, Othman MoustafaHM, Hassan L, et al. The accuracy of non-contrast chest
28 computed tomographic Scan in the detection of pulmonary thromboembolism. *J Curr Med*
29 *Res Pract* 2019;**4**:61. doi:10/ggzg83
30
- 31 35 McCollough CH, Bushberg JT, Fletcher JG, et al. Answers to Common Questions About the
32 Use and Safety of CT Scans. *Mayo Clin Proc* 2015;**90**:1380–92. doi:10/f3jgqx
33
- 34 36 Moser JB, Sheard SL, Edyvean S, et al. Radiation dose-reduction strategies in thoracic CT.
35 *Clin Radiol* 2017;**72**:407–20. doi:10/f95q7p
36
- 37 37 Long Q-X, Liu B-Z, Deng H-J, et al. Antibody responses to SARS-CoV-2 in patients with
38 COVID-19. *Nat Med* 2020;:1–4. doi:10.1038/s41591-020-0897-1
39
- 40 38 Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in Different Types of Clinical
41 Specimens. *JAMA* 2020;**323**:1843–4. doi:10/ggpp6h
42
- 43 39 Smith MJ, Hayward SA, Innes SM, et al. Point-of-care lung ultrasound in patients with
44 COVID-19 – a narrative review. *Anaesthesia*;n/a. doi:10/ggr2p7
45
- 46 40 Haaksma ME, Heldeweg MLA, Matta JEL, et al. Lung ultrasound findings in patients
47 with novel SARS-CoV2. *medRxiv* 2020;:2020.05.18.20105775.
48 doi:10.1101/2020.05.18.20105775
49
- 50 41 Benchoufi M, Bokobza J, Chauvin AA, et al. Lung injury in patients with or suspected
51 COVID-19 : a comparison between lung ultrasound and chest CT-scanner severity
52 assessments, an observational study. *medRxiv* 2020;:2020.04.24.20069633.
53 doi:10.1101/2020.04.24.20069633
54
- 55
56
57
58
59
60

- 1
2
3 42 Fine D, Perring S, Herbetko J, et al. Three-dimensional (3D) ultrasound imaging of the
4 gallbladder and dilated biliary tree: reconstruction from real-time B-scans. *Br J Radiol*
5 1991;**64**:1056–7. doi:10/fqr9mh
6
7 43 Shi F, Wang J, Shi J, et al. Review of Artificial Intelligence Techniques in Imaging Data
8 Acquisition, Segmentation and Diagnosis for COVID-19. *IEEE Rev Biomed Eng* 2020;:1–1.
9 doi:10/ggs2km
10
11 44 Li L, Qin L, Xu Z, et al. Artificial Intelligence Distinguishes COVID-19 from Community
12 Acquired Pneumonia on Chest CT. *Radiology* 2020;:200905. doi:10/ggpdgp
13
14 45 Wang L, Wong A. COVID-Net: A Tailored Deep Convolutional Neural Network Design
15 for Detection of COVID-19 Cases from Chest X-Ray Images. *ArXiv200309871 Cs Eess*
16 Published Online First: 11 May 2020. <http://arxiv.org/abs/2003.09871> (accessed 13 Jun
17 2020).
18
19 46 Kotsiantis SB. Use of machine learning techniques for educational proposes: a
20 decision support system for forecasting students' grades. *Artif Intell Rev* 2012;**37**:331–
21 44. doi:10/fmbng4
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Tables

Ordinal scale for study	BSTI grade	Features on X-ray
		Alternative pathology such as
	CVCX3- Non-COVID-19 pneumothorax with no features of COVID-19 identified	
1	CVCX0- Normal	No pathology seen
2	CVCX2- Indeterminate for COVID-19 or atypical features	Poor quality film or central/ basal consolidation
3	CVCX1- Classic findings of COVID-19	Peripheral ground glass opacities

Table 1- Ordinal scale used in this study based on the British Society of Thoracic Imaging (BSTI) Reporting Template [10]

	SARS-CoV 2 RT-PCR		p-value	Missing (%)
	Negative	Positive		
n (%)	435 (36.3)	763 (63.7)		
Number of Swabs (%)	810 (48.3)	868 (51.7)		
Age (mean (SD))	62.74 (17.72)	66.18 (17.58)	0.001*	0
Ethnicity			0.097	19
Other- Asian (%)	29 (8.0)	72 (11.8)		
South- Asian (%)	27 (7.5)	38 (6.2)		
Black (%)	41 (11.4)	91 (14.9)		
Mixed (%)	6 (1.7)	6 (1.0)		
Other (%)	56 (15.5)	105 (17.2)		
White (%)	202 (56.0)	297 (48.8)		
Sex – Male (%)	233 (53.6)	480 (62.9)	0.002*	0
Oxygen Saturation (median (IQR))	95 (6)	93 (8)	<0.001**	6.3
Respiratory Rate (median (IQR))	22 (8)	26 (12)	<0.001**	6.3
Glasgow Coma Scale (median (IQR))	15 (0)	15 (0)	0.043*	6.6
Systolic BP (median (IQR))	134 (32)	130 (30)	0.009*	15.8
Heart Rate (median (IQR))	96 (27)	94 (27)	0.092	6.4
Temperature (median (IQR))	37.1 (1.4)	37.7 (1.4)	<0.001**	6.7
Chest X-ray report			<0.001**	0
Alternative pathology (%)	4 (0.9)	3 (0.4)		
No abnormalities (%)	178 (40.9)	136 (17.8)		
Indeterminate (%)	83 (19.1)	169 (22.1)		
Classic COVID-19 (%)	170 (39.1)	455 (59.6)		
Presence of comorbidities (%)	297 (79.0)	482 (80.3)	0.669	18.5
Dyspnoea (%)	274 (69.4)	497 (75.5)	0.034	12.1
Neutrophils (median (IQR))	6.42 (4.56)	5.25 (3.92)	<0.001**	2.3
D-Dimer (median (IQR))	1250 (2440)	1105 (1803)	0.204	23.2
Albumin (median (IQR))	39 (7)	37 (6)	<0.001**	10
C-Reactive Protein (median (IQR))	91.0 (115)	146.5 (264.8)	<0.001**	3
Creatine Kinase (median (IQR))	51 (104)	145 (260)	<0.001**	23.3
Troponin (median (IQR))	19 (46)	20 (44)	0.278	19.1
Admitted (%)	331 (76.0)	635 (83.2)	0.003*	0.1
Admitted to ITU (%)	5 (1.3)	32 (4.8)	0.005*	12.4
Thirty Day Follow Up Status			<0.001**	24
Discharged (%)	219 (78.2)	367 (58.3)		
On Ambulatory Follow Up (%)	14 (5.0)	49 (7.8)		
Admitted (%)	18 (6.4)	60 (9.5)		
Died (%)	29 (10.4)	154 (24.4)		
CT report			<0.001**	0
No pathology identified (%)	23 (22.1)	6 (3.3)		
Classic COVID-19 findings (%)	52 (50.0)	157 (85.8)		
Indeterminate for COVID-19 (%)	14 (13.5)	14 (7.7)		
Alternative pathology identified (%)	15 (14.4)	6 (3.3)		
Day of Symptoms (mean (SD))	9.84 (9.63)	8.56 (15.80)	0.368	69.2

Table 2- Baseline characteristics of dataset stratified by overall SARS-CoV 2 RT-PCR status, including subsequent swabs during the study period- NB there were 480 additional swabs on 399 unique patients with a median of 2 and mean of 3.5 per patient; *significant at $p < 0.05$; **significant at $p < 0.001$

	Chest X-ray	CT Chest	Mean Difference	p-value
Total (n)	860	302		
True Positives (n)	305	162	-	-
False Positives (n)	125	55	-	-
True Negatives (n)	187	56	-	-
False Negatives (n)	243	29	-	-
Apparent prevalence (95% CI)	0.50 (0.47-0.53)	0.72 (0.66-0.77)	0.22 (0.04-0.21)	<0.0001**
True prevalence (95% CI)	0.64 (0.60-0.67)	0.63 (0.58-0.69)	-0.00 (-0.09-0.03)	0.111
Sensitivity (95% CI)	0.56 (0.51-0.60)	0.85 (0.79-0.90)	0.29 (0.19-0.38)	<0.0001**
Specificity (95% CI)	0.60 (0.54-0.65)	0.50 (0.41-0.60)	-0.10 (-0.25-0.04)	0.119
Positive Predictive Value (95% CI)	0.71 (0.66-0.75)	0.75 (0.68-0.80)	0.04 (-0.06-0.14)	0.492
Negative Predictive Value (95% CI)	0.43 (0.39-0.48)	0.66 (0.55-0.76)	0.22 (0.06-0.37)	0.005*
Positive Likelihood Ratio (95% CI)	1.39 (1.19-1.62)	1.71 (1.41- 2.08)	0.32 (-0.22-0.89)	0.258
Negative Likelihood Ratio (95% CI)	0.74 (0.64-0.84)	0.30 (0.21-0.44)	-0.44 (-0.64 - -0.21)	0.022*
Diagnostic Accuracy (95% CI)	0.57 (0.54-0.61)	0.72 (0.66-0.77)	0.15 (0.06-0.23)	<0.0001**

Table 3- Diagnostic Accuracy Metrics for CXR and CT Chest with RT-PCR for SARS-CoV 2, as the reference standard; *significant difference at the <0.05 level; **significant difference at the <0.0001 level

		SARS-CoV 2 RT-PCR		OR (univariable)	OR (multivariable)	
		Negative	Positive			
1	n	312	548			
2	Chest X-ray report	Alternative pathology (%)	3 (0.8)	3 (0.5)	-	-
3		No abnormalities (%)	123 (39.6)	104 (19.1)	0.76 (0.08-6.82, p=0.801)	0.48 (0.03-8.82, p=0.620)
4		Indeterminate/ atypical findings (%)	61 (19.5)	136 (4.8)	1.99 (0.22-17.81, p=0.535)	0.92 (0.05-16.88, p=0.952)
5		Classic COVID (%)	125 (40.1)	305 (55.6)	2.17 (0.24-19.19, p=0.484)	1.14 (0.06-20.98, p=0.927)
6	Age	Mean (SD)	61.8 (17.9)	67.0 (17.7)	1.02 (1.01-1.02, p<0.001)**	1.02 (1.00-1.03, p=0.028)*
7	Sex	Female (%)	138 (44.3)	212 (38.7)	-	-
8		Male (%)	174 (55.7)	336 (61.3)	1.26 (0.93-1.70, p=0.137)	1.19 (0.83-1.71, p=0.340)
9	Ethnicity	Other Asian (%)	31 (9.9)	66 (12.0)	-	-
10		White (%)	164 (52.7)	270 (49.2)	0.76 (0.44-1.31, p=0.326)	0.73 (0.38-1.40, p=0.339)
11		Black (%)	39 (12.4)	84 (15.3)	1.01 (0.52-1.98, p=0.974)	0.92 (0.43-1.97, p=0.827)
12		Mixed (%)	6 (1.8)	4 (0.8)	0.36 (0.08-1.62, p=0.184)	0.74 (0.11-4.94, p=0.754)
13		South Asian (%)	22 (7.0)	36 (6.6)	0.77 (0.34-1.76, p=0.531)	0.68 (0.28-1.65, p=0.390)
14		Other (%)	51 (16.2)	89 (16.2)	0.82 (0.43-1.55, p=0.535)	0.88 (0.45-1.74, p=0.716)
15	Comorbidity	No (%)	65 (20.8)	95 (17.4)	-	-
16		Yes (%)	247 (79.2)	453 (82.6)	1.25 (0.82-1.89, p=0.296)	1.00 (0.53-1.88, p=0.993)
17	Dyspnoea on attendance	No (%)	90 (28.8)	139 (25.4)	-	-
18		Yes (%)	222 (71.2)	409 (74.6)	1.19 (0.82-1.73, p=0.356)	0.84 (0.53-1.32, p=0.447)
19	Oxygen Saturation	Median (IQR)	96 (6)	93 (8)	0.94 (0.91-0.97, p<0.001)**	0.97 (0.93-1.00, p=0.072)
20	Respiratory rate	Median (IQR)	23 (8)	25 (8)	1.04 (1.01-1.07, p=0.002)*	1.01 (0.98-1.05, p=0.462)
21	Glasgow Coma Scale	Median (IQR)	15 (0)	15 (0)	1.02 (0.89-1.17, p=0.819)	1.21 (0.98-1.48, p=0.073)
22	Temperature	Mean (SD)	37.2 (1.4)	37.7 (1.1)	1.48 (1.26-1.73, p<0.001)**	1.44 (1.20-1.74, p<0.001)**
23	Heart Rate	Mean (SD)	96.7 (20.5)	94.9 (21.5)	1.00 (0.99-1.00, p=0.305)	1.00 (0.99-1.01, p=0.702)
24	Systolic Blood Pressure	Mean (SD)	136.2 (25.8)	132.6 (24.5)	0.99 (0.99-1.00, p=0.086)	0.99 (0.98-1.00, p=0.097)
25	Neutrophils	Median (IQR)	6.26 (4.52)	5.05 (3.93)	0.92 (0.89-0.96, p<0.001)**	0.87 (0.82-0.91, p<0.001)**
26	D-Dimer	Median (IQR)	1220 (2343)	1061 (1814)	1.00 (1.00-1.00, p=0.403)	1.00 (1.00-1.00, p=0.419)
27	C-Reactive Protein	Median (IQR)	45 (100)	77 (107)	1.00 (1.00-1.01, p<0.001)**	1.00 (1.00-1.01, p=0.021)*
28	Troponin	Median (IQR)	20 (55)	21 (46)	1.00 (1.00-1.00, p=0.890)	1.00 (1.00-1.00, p=0.667)
29	Albumin	Median (IQR)	39 (7)	37 (6)	0.97 (0.94-1.00, p=0.071)	1.02 (0.98-1.06, p=0.432)
30	Creatine Kinase	Median (IQR)	94 (131)	145 (263)	1.00 (1.00-1.00, p=0.119)	1.00 (1.00-1.00, p=0.152)
31	Admitted from ED	Admitted (%)	235 (75.2)	453 (82.7)	-	-
32		Discharged (%)	77 (24.8)	95 (17.3)	1.56 (1.06 -2.33, p=0.022)**	1.35 (0.79-2.30, p=0.272)
33	Admitted To ITU from ED	No (%)	307 (98.5)	532 (97.1)	-	-
34		Yes (%)	5 (1.5)	16 (2.9)	1.92 (0.60-6.18, p=0.274)	1.06 (0.25-4.40, p=0.940)
35	Thirty Day Follow up Status	Discharged (%)	259 (83.0)	368 (67.1)	-	-
36		Admitted (%)	22 (6.9)	47 (8.5)	1.53 (0.82-2.87, p=0.181)	1.64 (0.77-3.51, p=0.198)
37		Dead (%)	31 (10.1)	133 (24.4)	3.00 (1.86-4.84, p<0.001)**	2.81 (1.22-6.50, p=0.017)*

Table 4- Association of covariates with RT-PCR status for SARS-CoV 2, following propensity score matching and binomial logistic regression; SD- Standard deviation; IQR- Interquartile Range; *p<0.05; **p<0.001

		X-ray report		OR (univariable)	OR with XR as binary outcome (multivariable)	OR with XR as ordinal variable (multivariable)	
		Other X-ray Findings	Classical COVID-19				
6	n	430	430				
7	RT-PCR for SARS-CoV 2	Negative (%)	187 (43.4)	125 (29.1)	-	-	
8		Positive (%)	243 (56.6)	305 (70.9)	1.85 (1.36-2.56, p<0.001)**	1.79 (1.25-2.56, p<0.002)*	1.94 (1.37-2.76, p<0.001)**
9	Age	Mean (SD)	65.0 (18.9)	65.3 (16.9)	1.00 (0.99-1.01, p=0.849)	0.99 (0.98-1.00, p=0.164)	1.00 (0.99-1.01, p=0.542)
10	Sex	Female (%)	176 (40.9)	175 (40.6)	-	-	-
11		Male (%)	254 (59.1)	255 (59.3)	1.01 (0.75-1.37, p=0.940)	0.87 (0.63-1.20, p=0.400)	1.02 (0.49-2.09, p=0.967)
12	Ethnicity	Other Asian (%)	49 (11.4)	48 (11.2)	-	-	-
13		South Asian (%)	29 (6.7)	29 (6.7)	1.04 (0.52-2.04, p=0.912)	1.02 (0.47-2.17, p=0.965)	1.02 (0.49-2.09, p=0.967)
14		Black (%)	61 (14.2)	61 (14.2)	1.02 (0.55-1.85, p=0.957)	0.88 (0.46-1.69, p=0.719)	0.92 (0.52-1.65, p=0.789)
15		Mixed (%)	5 (1.2)	5 (1.2)	0.92 (0.21-4.00, p=0.911)	0.86 (0.18-4.17, p=0.853)	0.85 (0.17-4.30, p=0.838)
16		Other (%)	70 (16.3)	70 (16.3)	1.02 (0.58-1.79, p=0.943)	0.98 (0.52-1.82, p=0.942)	0.93 (0.53-1.64, p=0.810)
17	White (%)	216 (50.2)	217 (50.5)	1.03 (0.63-1.67, p=0.913)	0.97 (0.57-1.67, p=0.926)	0.90 (0.55-1.47, p=0.666)	
18	Comorbidity	No (%)	82 (19.1)	78 (18.1)	-	-	-
19		Yes (%)	348 (80.9)	352 (81.9)	0.95 (0.66-1.36, p=0.777)	0.93 (0.59-1.49, p=0.782)	0.88 (0.57-1.37, p=0.592)
20	Dyspnoea	No (%)	191 (29.3)	103 (24.0)	-	-	-
21		Yes (%)	304 (70.7)	327 (76.0)	1.31 (0.92-1.88, p=0.123)	1.20 (0.80-1.82, p=0.380)	1.22 (0.83-1.80, p=0.301)
22	Oxygen Saturation	Median (IQR)	95 (7)	93 (7)	0.94 (0.91-0.96, p<0.001)**	0.94 (0.92-0.97, p<0.001)**	0.94 (0.91-0.97, p<0.001)**
23	Respiratory rate	Median (IQR)	24 (10)	24 (10)	1.01 (0.99-1.02, p=0.570)	0.97 (0.94-1.00, p=0.063)	0.98 (0.96-1.01, p=0.157)
24	Glasgow Coma Scale	Median (IQR)	15 (0)	15 (0)	1.04 (0.92-1.19, p=0.524)	1.05 (0.90-1.23, p=0.503)	1.05 (0.92-1.21, p=0.464)
25	Temperature	Mean (SD)	37.6 (1.1)	37.5 (1.3)	0.93 (0.83-1.06, p=0.297)	0.79 (0.67-0.93, p=0.006)*	0.85 (0.73-0.99, p=0.031)*
26	Heart Rate	Mean (SD)	95.7 (21.4)	95.5 (21.0)	1.00 (0.99-1.01, p=0.888)	1.00 (0.99-1.01, p=0.864)	1.00 (0.99-1.01, p=0.872)
27	Systolic Blood Pressure	Mean (SD)	133.8 (25.0)	134.0 (25.6)	1.00 (0.99-1.01, p=0.907)	1.00 (0.99-1.01, p=0.335)	1.00 (1.00-1.01, p=0.478)
28	Neutrophils	Median (IQR)	5.44 (4.54)	5.67 (4.03)	1.00 (0.97-1.04, p=0.892)	0.96 (0.92-1.01, p=0.143)	0.96 (0.92-1.01, p=0.115)
29	D-Dimer	Median (IQR)	1119 (2221)	1119 (1850)	1.00 (1.00-1.00, p=0.513)	1.00 (1.00-1.00, p=0.568)	1.00 (1.00-1.00, p=0.385)
30	C-Reactive Protein	Median (IQR)	46 (93)	88 (110)	1.00 (0.99-1.00, p<0.001)**	1.00 (1.00-1.01, p<0.001)**	1.00 (1.00-1.01, p<0.001)**
31	Troponin	Median (IQR)	23 (54)	20 (46)	1.00 (1.00-1.00, p=0.231)	1.00 (1.00-1.00, p=0.277)	1.00 (1.00-1.00, p=0.059)
32	Albumin	Median (IQR)	39 (7)	37 (6)	0.93 (0.90-0.96, p<0.001)**	0.93 (0.90-0.97, p=0.001)*	0.94 (0.91-0.97, p=0.001)*
33	Creatine Kinase	Median (IQR)	110 (183)	134 (239)	1.00 (1.00-1.00, p=0.535)	1.00 (1.00-1.00, p=0.242)	1.00 (1.00-1.00, p=0.186)
34	Admitted from ED	Admitted (%)	315 (73.3)	373 (86.7)	2.37 (1.63-3.46, p<0.001)**	2.30 (1.46-3.63, p<0.001)**	2.22 (1.47-3.33, p<0.001)**
35		Discharged (%)	115 (26.7)	57 (13.3)	-	-	-
36	Admitted to ITU from ED	No (%)	423 (98.4)	416 (96.7)	-	-	-
37		Yes (%)	7 (1.6)	14 (3.3)	2.17 (0.69-6.67, p=0.181)	1.27 (0.32-5.00, p=0.732)	1.34 (0.36-5.00, p=0.653)
38	30 Day Follow Up Status	Discharged (%)	316 (73.5)	311 (72.3)	-	-	-
39		Admitted (%)	34 (7.9)	34 (7.9)	1.31 (0.81-2.13, p=0.282)	1.32 (0.69-2.53, p=0.392)	1.43 (0.78-2.63, p=0.653)
40		Dead (%)	80 (18.6)	85 (19.8)	1.03 (0.73-1.45, p=0.886)	1.38 (0.80-2.37, p=0.247)	1.41 (0.87-2.27, p=0.157)

Table 5- Association of covariates with CXR report following propensity score matching and either binomial or ordinal logistic regression; SD- Standard deviation; IQR- Interquartile Range; *p<0.05; **p<0.001

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

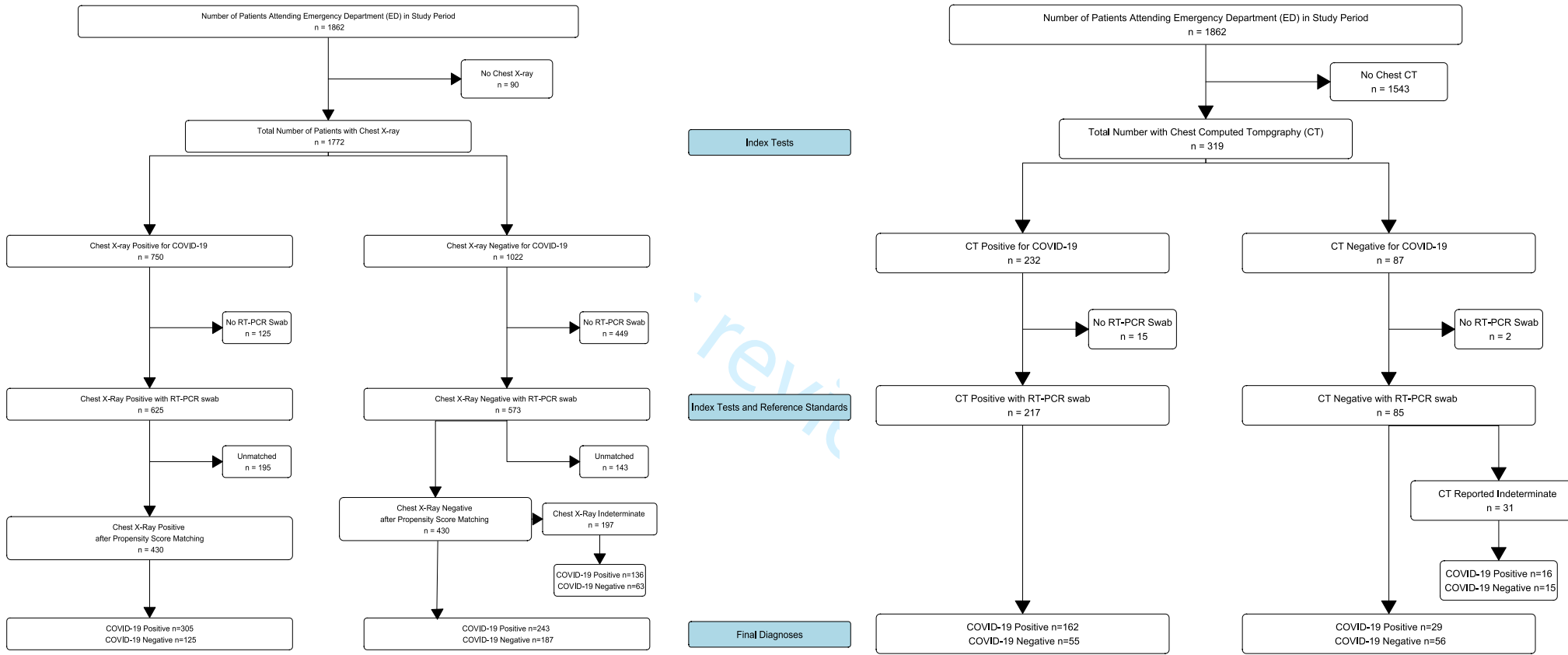


Figure 1- Inclusion and exclusion of patients during study period with test results

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Odds Ratio of Positivity for SARS-CoV 2 by RT-PCR

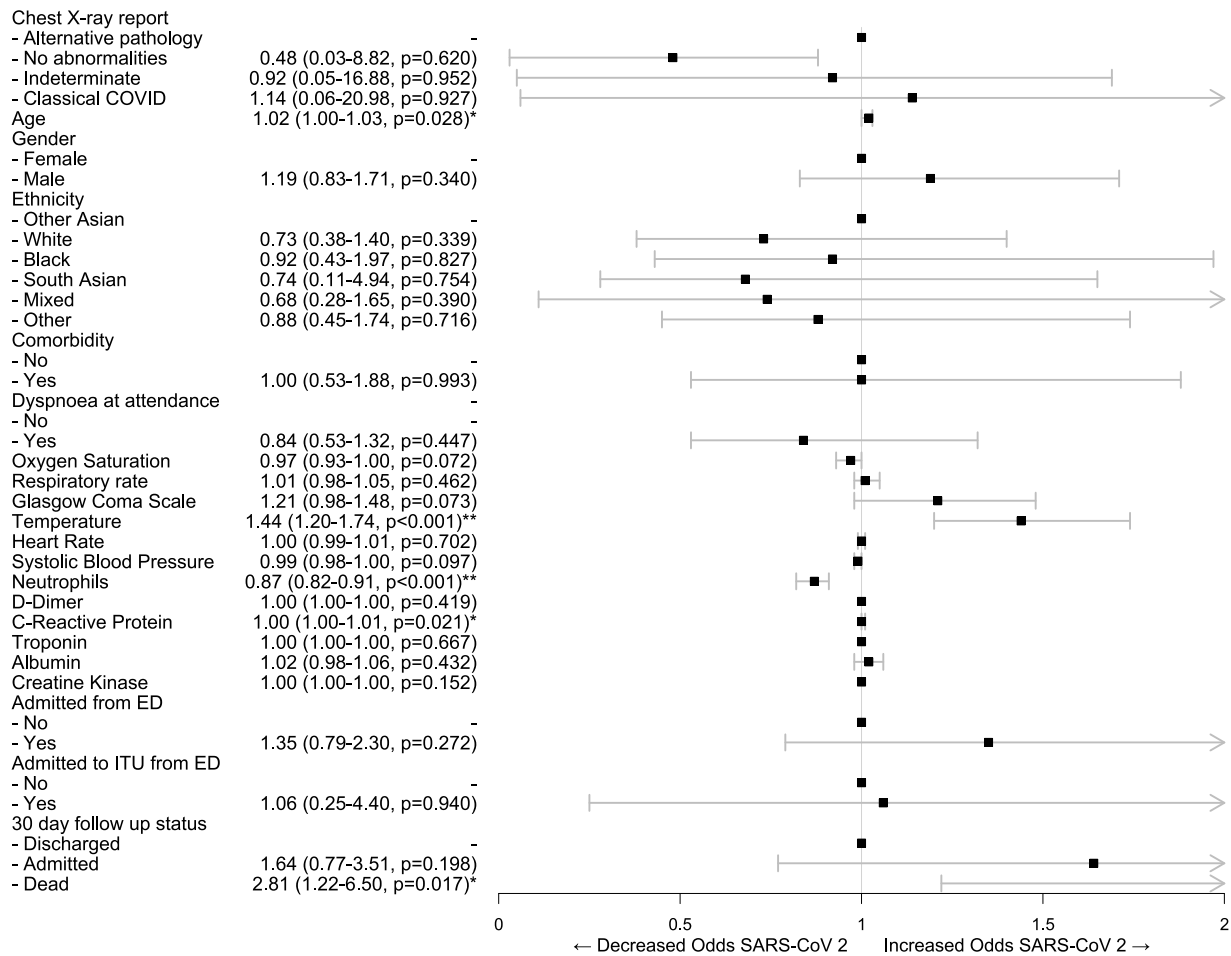


Figure 2- Forest plot of odds ratios of variables associated with RT-PCR positivity for SARS-CoV 2, following multiple imputation, propensity score matching and binomial logistic regression; *significant difference at the <0.05 level; **significant difference at the <0.001 level

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Odds Ratio of Classical COVID-19 Findings on Chest X-Ray

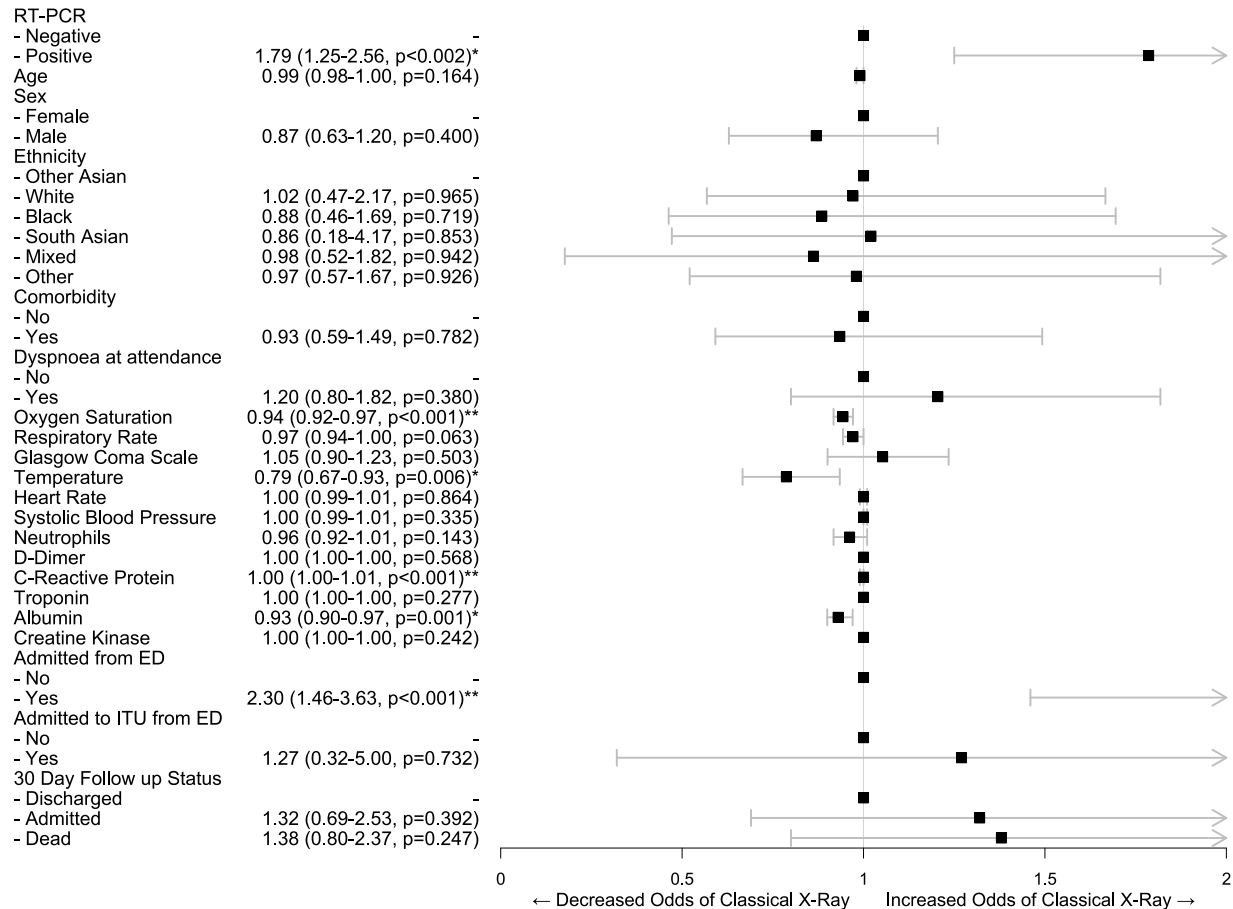
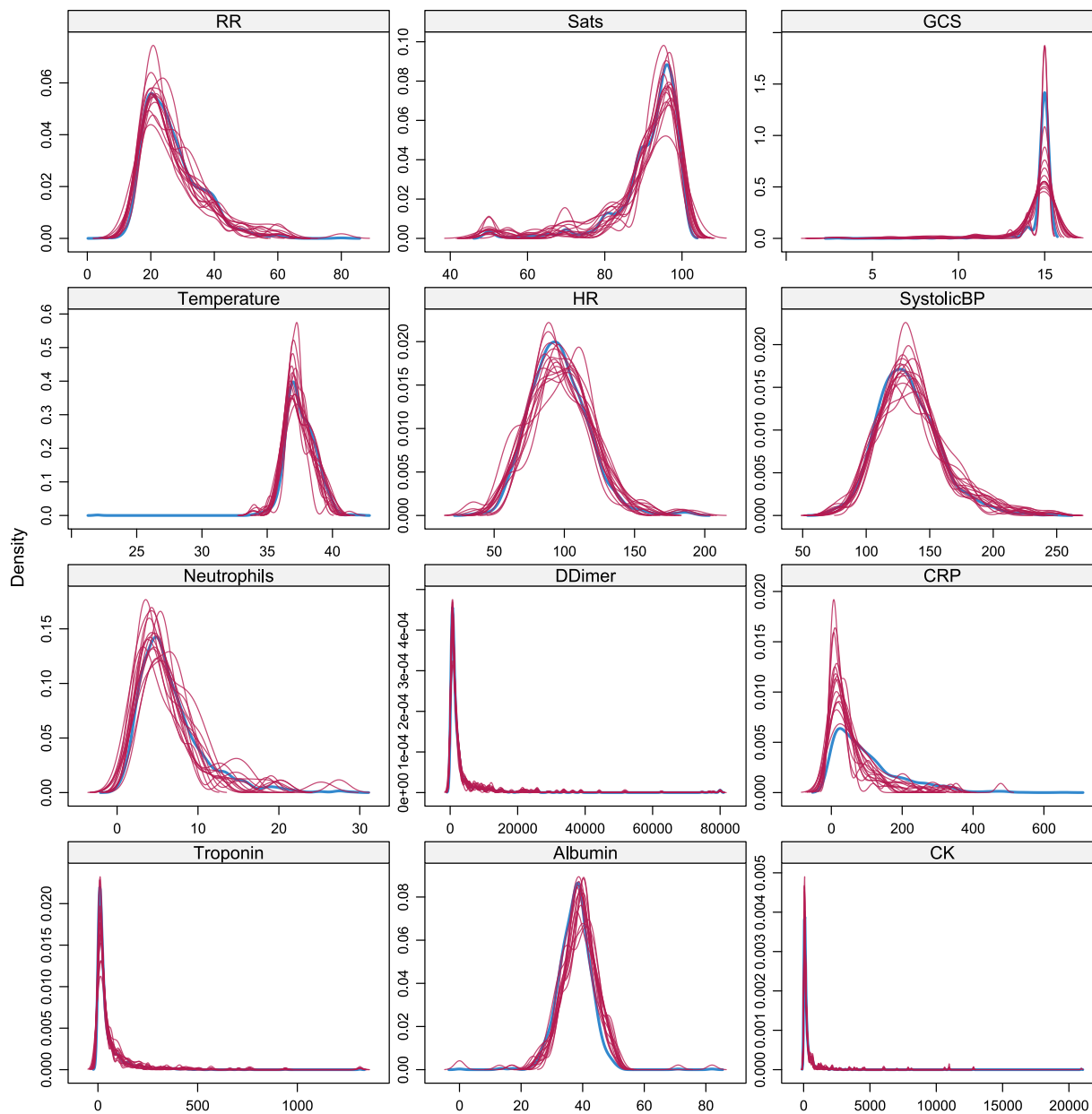


Figure 3- Forest plot of odds ratios of variables associated with classical Chest X-ray features COVID-19 following propensity score matching and binomial logistic regression; *significant difference at the <0.05 level; **significant difference at the <0.001 level

Supplementary file 1



Supplementary figure 1- Density plots of imputed datasets; Blue represents original dataset; other colours are individual imputed datasets (n=15)

Covariate:	Means Treated	Means Control	Standard Deviation Control	Mean Difference
Overall Propensity Score	0.422997940	0.53935303	0.1449627	-0.1163550897
Female	36.3782051	45.026178	0.4979547	-8.64797288
Male	63.6217949	54.973822	0.4979547	8.64797288
Age	63.796474359	66.19022688	18.5893357	-23.937525171
Comorbidity- Yes	76.1217949	84.467714	0.3625287	-8.34591892
Ethnicity- South Asian	6.5705128	6.631763	0.2490539	-0.06124983
Ethnicity- Black	16.1858974	11.518325	0.3195219	4.66757283
Ethnicity- Mixed	0.9615385	1.396161	0.1174340	-0.43462210
Ethnicity- Other	18.9102564	13.263525	0.3394765	5.64673110
Ethnicity- White	46.6346154	57.766143	0.4943635	-11.13152772
Respiratory Rate	29.214743590	24.01745201	7.2639816	5.1972915828

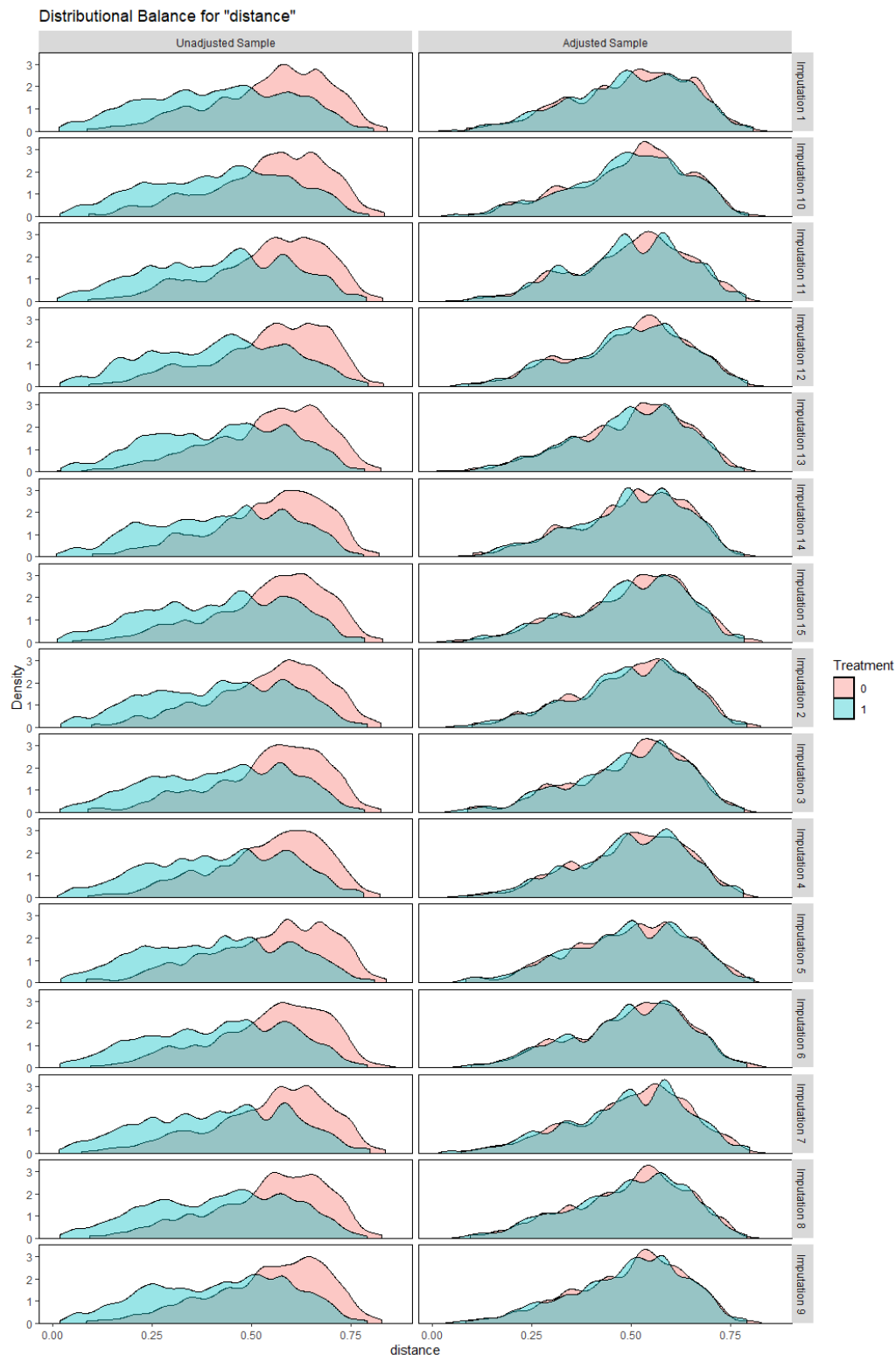
Supplementary table 1- Means of data before multiple imputation and propensity score matching

	Type	Minimum Difference Adjusted	Mean Difference Adjusted	Maximum Difference Adjusted
Distance	Distance	0.016988	0.027107	0.040963
Sex = Male	Binary	-0.03917	-0.0028	0.015982
Age	Contin.	-0.04586	-0.01371	0.027589
Comorbidity = Yes	Binary	-0.02331	-0.00778	0.004598
Ethnicity = Other Asian	Binary	-0.01392	0.002362	0.016471
Ethnicity = South Asian	Binary	-0.01399	-0.00136	0.011905
Ethnicity = Black	Binary	-0.01852	0.000443	0.015982
Ethnicity = Mixed	Binary	-0.00464	0.001403	0.007042
Ethnicity = Other	Binary	-0.01152	4.30E-06	0.00939
Ethnicity = White	Binary	-0.02353	-0.00285	0.018433
Respiratory Rate	Contin.	-0.06157	-0.03478	-0.00442

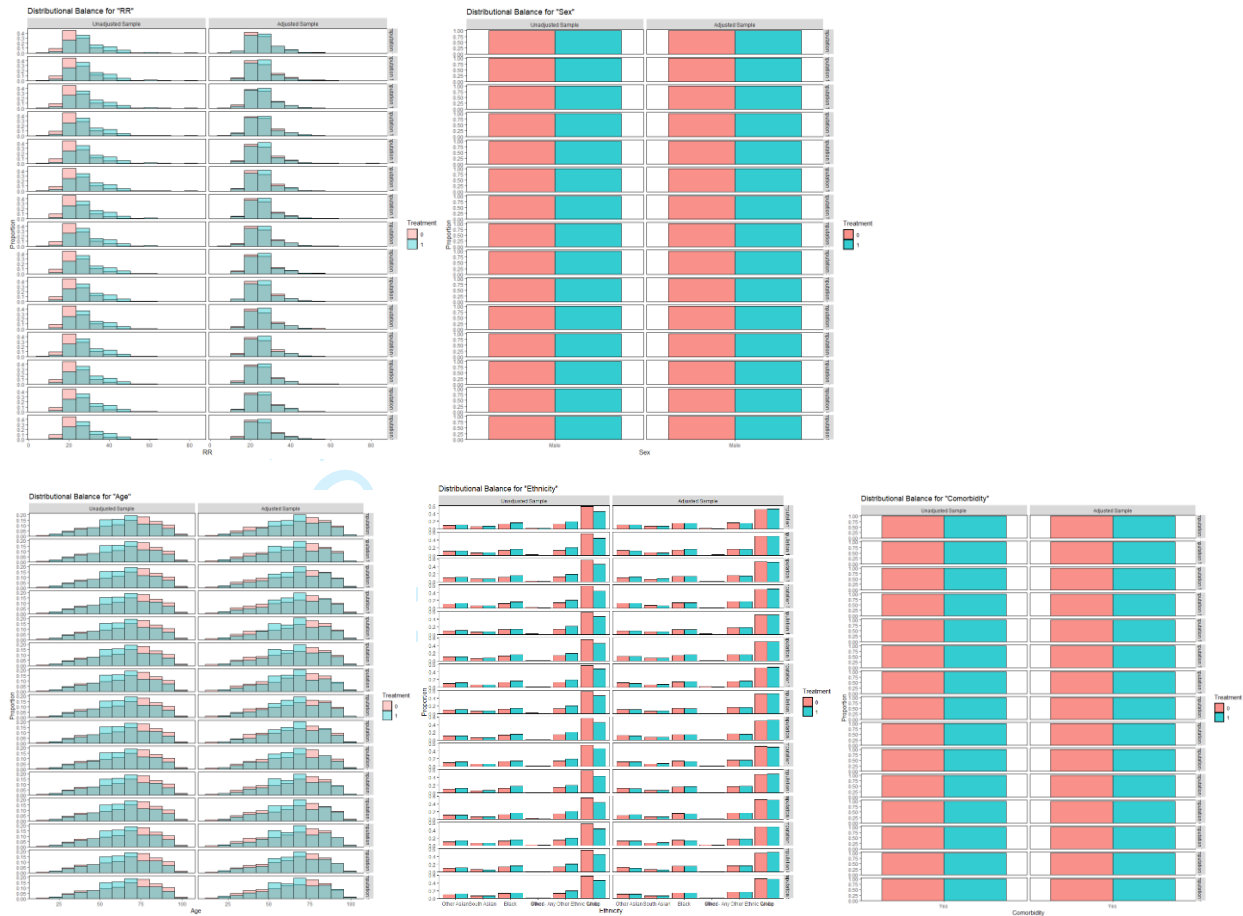
Supplementary table 2- Balance summary across imputations

	XR- Negative	XR- Positive	Total
All	573	625	1,198
Matched	430	430	860
Unmatched	143	195	338
Discarded	0	0	0

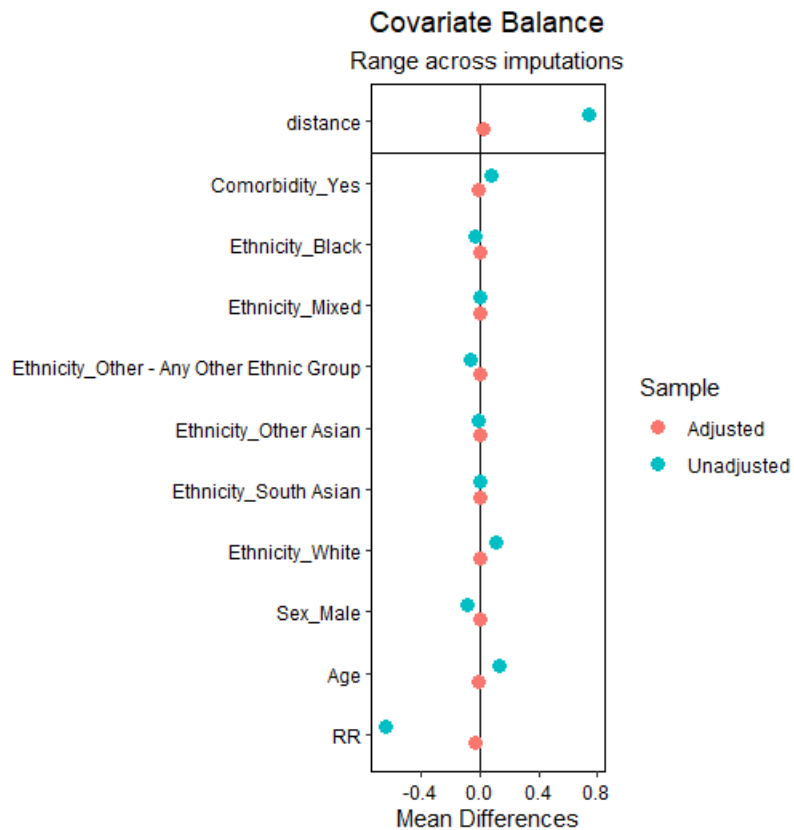
Supplementary table 3- Average Sample sizes pre- and post- matching across imputed data sets



Supplementary figure 2- Density plot of propensity scores pre- and post- matching in each imputed dataset; treatment units represent a positive X-ray for COVID-19, whereas a control unit represents a negative X-ray



Supplementary figure 3- Histogram of distributions for each matching covariate pre- and post- matching in each imputed dataset; treatment units represent a positive X-ray for COVID-19, whereas a control unit represents a negative X-ray



Supplementary figure 4- Love plot of pooled balances across imputed datasets in matching covariates after matching

CXR in COVID Analysis

Dr Aditya Borakati

Royal Free Hospital, Pond Street, London, NW3 2QG
a.borakati@doctors.org.uk

2020-10-06

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1 Software Environment and Packages	7
1.1 Load Packages and Data	9
1.1.1 Load Packages:	9
1.2 Power Calculation	9
2 Load Data:	11
3 Data Cleaning	13
3.0.1 Follow Up Swabs + Initial Swabs Positive: .	14
3.0.2 Paired XR and RT-PCR data	14
4 Demographic table of raw data	17
5 Imputation	21
6 Propensity Score Matching	23
6.1 Match Balance Diagnostics	23
7 Matched Demographics Table:	25
8 Diagnostic Accuracy	27
8.1 CT Data and Accuracy	28
8.2 CT and XR accuracy comparison	30
8.2.1 Sensitivity	30
8.3 Intermodality Agreement	34
8.3.1 Diagnostic Accuracy Analysis when Indeterminate Reports of CXR and CT are taken as positive	35
9 Pooled Regression after Multiple Imputation and Propensity Score Matching	39
9.0.1 Pooled Univariate Odds Ratios for OverallPos as dependent variable	39
9.0.2 Binomial Logistic Regression with Positive Chest X-ray Report as Dependent Variable	40
9.0.3 Univariate XRPositive as dependent	40
9.0.4 Multivariate XRPositive as dependent	40
9.0.5 Pooled Ordinal Logistic Regression with XRPositive as dependent	41

iv

9.1 Forest Plots	41
9.2 Correlation Matrix	43
9.3 STARD Flow Diagram	46

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1 Software Environment and Packages

```

R version 4.0.0 (2020-04-24)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 19041)
Matrix products: default
locale:
 LC_COLLATE=English_United Kingdom.1252 LC_CTYPE=English_United Kingdom.1252
 LC_MONETARY=English_United Kingdom.1252 LC_NUMERIC=C
 LC_TIME=English_United Kingdom.1252
attached base packages:
 stats graphics grDevices utils datasets methods base
other attached packages:
 corrplot 0.84
   Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of
   a Correlation Matrix (Version 0.84). Available from
   https://github.com/taiyun/corrplot
MKmisc 1.6
   Kohl M (2019). MKmisc: Miscellaneous functions from M. Kohl. R package
   version 1.6, http://www.stamats.de
epiR 1.0-14
   Mark Stevenson with contributions from Telmo Nunes, Cord Heuer, Jonathon
   Marshall, Javier Sanchez, Ron Thornton, Jenő Reiczigel, Jim Robison-Cox,
   Paola Sebastiani, Peter Solymos, Kazuki Yoshida, Geoff Jones, Sarah
   Pirikahu, Simon Firestone, Ryan Kyle, Johann Popp, Mathew Jay and Charles
   Reynard. (2020). epiR: Tools for the Analysis of Epidemiological Data. R
   package version 1.0-14. https://CRAN.R-project.org/package=epiR
Matching 4.9-7
   Jasjeet S. Sekhon (2011). Multivariate and Propensity Score Matching
   Software with Automated Balance Optimization: The Matching Package for R.
   Journal of Statistical Software, 42(7), 1-52. URL
   http://www.jstatsoft.org/v42/i07/.
MASS 7.3-51.5
   Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S.
   Fourth Edition. Springer, New York. ISBN 0-387-95457-0
Ordinal 2019.12-10
   Christensen, R. H. B. (2019). ordinal - Regression Models for Ordinal Data. R
   package version 2019.12-10. https://CRAN.R-
   project.org/package=ordinal.
Hmisc 4.4-0
   Frank E Harrell Jr, with contributions from Charles Dupont and many
   others. (2020). Hmisc: Harrell Miscellaneous. R package version 4.4-0.
   https://CRAN.R-project.org/package=Hmisc
Formula 1.2-3
   Achim Zeileis, Yves Croissant (2010). Extended Model Formulas in R:
   Multiple Parts and Multiple Responses. Journal of Statistical Software
   34(1), 1-13. doi:10.18637/jss.v034.i01
lattice 0.20-41
   Sarkar, Deepayan (2008) Lattice: Multivariate Data Visualization with R.
   Springer, New York. ISBN 978-0-387-75968-5

```

8 1 Software Environment and P...

```
mice 3.8.0
  Stef van Buuren, Karin Groothuis-Oudshoorn (2011). mice: Multivariate
  Imputation by Chained Equations in R. Journal of Statistical Software,
  45(3), 1-67. URL https://www.jstatsoft.org/v45/i03/.
readxl 1.3.1
  Hadley Wickham and Jennifer Bryan (2019). readxl: Read Excel Files. R
  package version 1.3.1. https://CRAN.R-project.org/package=readxl
finalfit 1.0.1
  Ewen Harrison, Tom Drake and Riinu Ots (2020). finalfit: Quickly Create
  Elegant Regression Results Tables and Plots when Modelling. R package
  version 1.0.1. https://CRAN.R-project.org/package=finalfit
MatchIt 3.0.2
  Daniel E. Ho, Kosuke Imai, Gary King, Elizabeth A. Stuart (2011). MatchIt:
  Nonparametric Preprocessing for Parametric Causal Inference. Journal of
  Statistical Software, Vol. 42, No. 8, pp. 1-28. URL
  http://www.jstatsoft.org/v42/i08/
tableone 0.11.1
  Kazuki Yoshida (2020). tableone: Create 'Table 1' to Describe Baseline
  Characteristics. R package version 0.11.1.
  https://CRAN.R-project.org/package=tableone
forcats 0.5.0
  Hadley Wickham (2020). forcats: Tools for Working with Categorical
  Variables (Factors). R package version 0.5.0.
  https://CRAN.R-project.org/package=forcats
stringr 1.4.0
  Hadley Wickham (2019). stringr: Simple, Consistent Wrappers for Common
  String Operations. R package version 1.4.0.
  https://CRAN.R-project.org/package=stringr
dplyr 0.8.5
  Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2020).
  dplyr: A Grammar of Data Manipulation. R package version 0.8.5.
  https://CRAN.R-project.org/package=dplyr
purrr 0.3.4
  Lionel Henry and Hadley Wickham (2020). purrr: Functional Programming
  Tools. R package version 0.3.4. https://CRAN.R-project.org/package=purrr
readr 1.3.1
  Hadley Wickham, Jim Hester and Romain François (2018). readr: Read
  Rectangular Text Data. R package version 1.3.1.
  https://CRAN.R-project.org/package=readr
tidyr 1.0.2
  Hadley Wickham and Lionel Henry (2020). tidyr: Tidy Messy Data. R package
  version 1.0.2. https://CRAN.R-project.org/package=tidyr
tibble 3.0.0
  Hadley Wickham and Lionel Henry (2020). tidyr: Tidy Messy Data. R package
  version 1.0.2. https://CRAN.R-project.org/package=tidyr
ggplot2 3.3.0
  H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag
  New York, 2016.
tidyverse 1.3.0
  Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source
  Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
forestplot 1.9
  Max Gordon and Thomas Lumley (2019). forestplot: Advanced Forest Plot Using
  'grid' Graphics. R package version 1.9. https://CRAN.R-project.org/package=forestplot
MatchThem 0.9.3
  Farhad Pishgar and Noah Greifer (2020). MatchThem: Matching and Weighting
  Multiply Imputed Datasets. R package version 0.9.3. https://CRAN.R-project.org/package=MatchThem
```

1.1 Load Packages and Data

9

```
miceadds 3.9-14
  Robitzsch, A., & Grund, S. (2020). miceadds: Some Additional Multiple
  Imputation Functions, Especially for 'mice'. R package version 3.9-14.
  https://CRAN.R-project.org/package=miceadds
cobalt 4.2.2
  Noah Greifer (2020). cobalt: Covariate Balance Tables and Plots. R package
  version 4.2.2. https://CRAN.R-project.org/package=cobalt
```

1.1 Load Packages and Data

1.1.1 Load Packages:

```
library(MKmisc)
library(tidyverse)
library(tableone)
library(MatchIt)
library(finalfit)
library(readxl)
library(cobalt)
library(mice)
library(miceadds)
library(Hmisc)
library(epiR)
library(MatchThem)
library(ordinal)
library(forestplot)
```

1.2 Power Calculation

1.2.0.0.1 This code calculates the sample size (positive and negative by gold standard test) needed to evaluate a diagnostic test with 56% sensitivity at 80% power with alpha 0.05. The 56% value is the lower confidence reported by Wong et al. and lower sensitivities typically require higher sample sizes, the result is the same whether specificity or sensitivities are passed as arguments, the previously published specificities are higher than sensitivities so for a generous estimate, the sensitivity was used.

```
power <- power.diagnostic.test(sens = 0.56,
  sig.level = 0.05, delta = 0.1, power = 0.8) %>%
  print()
```

10

1 Software Environment and P...

```
Diagnostic test exact power calculation
```

```
    sens = 0.56  
      n  = 165  
     n1 = 165  
    delta = 0.1  
sig.level = 0.05  
    power = 0.8  
     prev = NULL
```

```
NOTE: n is number of cases, n1 is number of controls
```

2 Load Data:

```
data <- read_csv("FullDataWithCT.csv", col_types = cols(Age = col_integer(),
  Albumin = col_number(), CK = col_number(),
  CT = col_character(), CRP = col_number(),
  DDimer = col_number(), DateOfDeath = col_date(format = "%d/%m/%Y"),
  DateOfDischarge = col_date(format = "%d/%m/%Y"),
  DateOfVisit = col_date(format = "%d/%m/%Y"),
  DateOfSymptomOnset = col_date(format = "%d/%m/%Y"),
  DiastolicBP = col_number(), FiO2 = col_skip(),
  GCS = col_number(), HR = col_number(),
  MRN = col_skip(), NEWS = col_number(),
  `NEWS2(noFiO2)` = col_skip(), Neutrophils = col_number(),
  RR = col_number(), Sats = col_number(),
  `Supplemental Oxygen` = col_skip(), SystolicBP = col_number(),
  Temperature = col_number(), Troponin = col_number(),
  CTBSTI = col_integer()))
```


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

3 Data Cleaning

3.0.0.0.1 Format data into factors/ differences between dates:

```
data <- mutate_if(data, is.character, as.factor)
data$DayOfSymptoms <- difftime(data$DateOfVisit,
  data$DateOfSymptomOnset, units = "days")
data$TimeToDeath <- abs(difftime(data$DateOfDeath,
  data$DateOfVisit, units = "days"))
data$DayOfSymptoms <- as.numeric(data$DayOfSymptoms)
data$TimeToDeath <- as.numeric(data$TimeToDeath)
```

3.0.0.1 Recode ethnicities as too many options:

3.0.0.1.0.1 This code collapses the ethnicity categories into 'White', 'Black', 'South Asian', 'Other Asian', 'Mixed' or 'Other';

```
data$Ethnicity <- fct_collapse(data$Ethnicity,
  White = c("White - British", "White - Irish",
    "White - Any Other White Background"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  Black = c("Black - Any Other Black Background",
    "Black or Black British - African",
    "Black or Black British - African",
    "Black or Black British - Caribbean"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  `South Asian` = c("Asian or Asian British - Bangladeshi",
    "Asian or Asian British - Indian",
    "Asian or Asian British - Pakistani"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  `Other Asian` = c("Asian - Any Other Asian Background",
    "Other - Chinese"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  Mixed = c("mixed - Any Other mixed Background",
    "Mixed - Any Other Mixed Background",
    "Mixed - White and Asian", "Mixed - White and Black African",
    "mixed - White and Black Caribbean",
    "Mixed - White and Black Caribbean"))
```

3.0.0.1.0.2 New XR positive column for “Classic Covid” or not:

```
data$XRPositive <- ifelse(data$XR Chest ==  
  "Classic COVID", "Positive", "Negative")  
data$XRPositive <- as.factor(data$XRPositive)
```

3.0.1 Follow Up Swabs + Initial Swabs Positive:

3.0.1.0.0.1 Creates new column ‘OverallPos’ which includes initial RT-PCR swab and follow-up swabs in 30 days of attendance, if any are positive the value will be positive in this column

```
data$OverallPos <- case_when(data$RTPCR ==  
  "Positive" | data$FollowUpPos == "Positive" ~  
  "Positive")  
data$OverallPos <- replace_na(data$OverallPos,  
  "Negative")
```

3.0.1.0.0.2 Create new vector with all variable names (i.e. the column headers)

```
explanatory <- names(data)
```

3.0.2 Paired XR and RT-PCR data

3.0.2.1 Creates new variable ‘completedata’ which contains only patients who had both CXR and RT-PCR in ED

```
completedata <- filter(data, !is.na(data$XRPositive) &  
  !is.na(data$RTPCR))
```

3.0.2.1.1 Remove missing data variable

```
completedata <- completedata[-c(31)]
```

3.0.2.2 Format complete data variables

```
completedata$OverallPos <- as.factor(completedata$OverallPos)

completedata$ThirtyDayFU <- as.factor(completedata$ThirtyDayFU)
completedata$TimeToDeath <- abs(difftime(completedata$DateOfDeath,
completedata$DateOfVisit, units = "days"))

completedata$TimeToDeath <- as.numeric(completedata$TimeToDeath)
```

3.0.2.2.0.1 Set 'XRchest' as ordinal variable on scale of 'Alternative pathology' as lowest value and 'Classical COVID' as highest

```
completedata$XRchest <- ordered(completedata$XRchest,
levels = c("Alternative pathology", "No abnormalities",
"Indeterminate", "Classic COVID"))
```

3.0.2.2.0.2 Convert CT BSTI grade column into factor:

```
completedata$CTBSTI <- as.factor(completedata$CTBSTI)
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

4 Demographic table of raw data

4.0.0.0.1 This code creates an unformatted demographic table (table 2 in manuscript), for the raw data, stratified by RT-PCR status, significance testing between RT-PCR +ve and -ve groups is carried out automatically using chi squared, t-tests, ANOVA etc.; there is also a column for the proportion of missing data

```

CreateTableOne(vars = explanatory,
               strata = 'OverallPos',
               data = completedata) -> demogtable

#### List nonnormal factors for summarisation as median / IQR and non
      parametric statistical test

explanatoryynormal<-c("Sats", "RR", "GCS", "SystolicBP", "Temperature", "HR",
                    "Neutrophils",
                    "DDimer", "Albumin", "CRP", "CK", "Troponin")
as.data.frame(print(demogtable, nonnormal = explanatoryynormal, missing =
                    TRUE))->demogtable

write.csv(demogtable, file = "Demogtable.csv")

```

Age (mean (SD))	62.74 (17.72)	66.18 (17.58)
0.001		
Ethnicity (%)		
0.097		
Other Asian	29 (8.0)	72 (11.8)
South Asian	27 (7.5)	38 (6.2)
Black	41 (11.4)	91 (14.9)
Mixed	6 (1.7)	6 (1.0)
Other - Any Other Ethnic Group	56 (15.5)	105 (17.2)
White	202 (56.0)	297 (48.8)
Sex = Male (%)	233 (53.6)	480 (62.9)
0.002		
Sats (median [IQR])	95.00 [92.00, 98.00]	93.00 [88.00,
96.00]	<0.001 nonnorm	
RR (median [IQR])	22.00 [20.00, 28.00]	26.00 [20.00,
32.00]	<0.001 nonnorm	
GCS (median [IQR])	15.00 [15.00, 15.00]	15.00 [15.00,
15.00]	0.043 nonnorm	
SystolicBP (median [IQR])	134.00 [119.00, 151.50]	130.00 [115.00,
145.00]	0.009 nonnorm	
DiastolicBP (mean (SD))	79.54 (16.40)	75.61 (14.51)
<0.001		
HR (median [IQR])	96.00 [83.00, 110.00]	94.00 [81.00,
108.00]	0.092 nonnorm	

18

4 Demographic table of raw data

Temperature (median [IQR]) 38.40] <0.001 nonnorm	37.10 [36.60, 38.00]	37.70 [37.00,
XR Chest (%) <0.001		
Alternative pathology	4 (0.9)	3 (0.4)
No abnormalities	178 (40.9)	136 (17.8)
Indeterminate	83 (19.1)	169 (22.1)
Classic COVID	170 (39.1)	455 (59.6)
CTPA = PE (%) 0.127	16 (30.2)	28 (45.9)
Comorbidity = Yes (%) 0.669	297 (79.0)	482 (80.3)
Dyspnoea = Yes (%) 0.034	274 (69.4)	497 (75.5)
Neutrophils (median [IQR]) 7.61] <0.001 nonnorm	6.42 [4.55, 9.11]	5.25 [3.69,
DDimer (median [IQR]) 2428.50] 0.204 nonnorm	1250.00 [619.00, 3059.00]	1105.00 [626.00,
Albumin (median [IQR]) 40.00] <0.001 nonnorm	39.00 [35.00, 42.00]	37.00 [34.00,
CRP (median [IQR]) 158.00] <0.001 nonnorm	51.00 [13.00, 117.00]	83.00 [42.00,
CK (median [IQR]) 342.75] <0.001 nonnorm	91.00 [54.00, 169.00]	146.50 [78.00,
Troponin (median [IQR]) 53.00] 0.278 nonnorm	19.00 [7.00, 53.00]	20.00 [9.00,
Admitted = Discharged (%) 0.003	104 (24.0)	128 (16.8)
AdmittedToITU = Yes (%) 0.005	5 (1.3)	32 (4.8)
RTPCR = Positive (%) <0.001	0 (0.0)	738 (96.7)
CT = 1 (%) 0.011	37 (57.8)	26 (86.7)
NEWS (mean (SD)) 0.032	4.36 (3.06)	5.48 (2.71)
ThirtyDayFU (%) <0.001		
1	219 (78.2)	367 (58.3)
2	14 (5.0)	49 (7.8)
3	18 (6.4)	60 (9.5)
4	29 (10.4)	154 (24.4)
CTBSTI (%) <0.001		
0	23 (22.1)	6 (3.3)
1	52 (50.0)	157 (85.8)
2	14 (13.5)	14 (7.7)
3	15 (14.4)	6 (3.3)
DayOfSymptoms (mean (SD)) 0.368	9.84 (9.63)	8.56 (15.80)
TimeToDeath (mean (SD)) 0.618	50.33 (77.93)	57.76 (70.02)
XRPositive = Positive (%) <0.001	170 (39.1)	455 (59.6)
OverallPos = Positive (%)	0 (0.0)	763 (100.0)

1
2
3
4
5
6 **4.0.0.0.2 Limited dataset comprising relevant data and those without**
7 **significant missingness:**
8

```
9  
10 limcompletedata <- dplyr::select(completedata,  
11   c("Age", "XRchest", "Ethnicity", "Sex",  
12     "RR", "Sats", "GCS", "Temperature",  
13     "HR", "SystolicBP", "DiastolicBP",  
14     "Neutrophils", "DDimer", "CRP", "Troponin",  
15     "Albumin", "CK", "OverallPos", "Admitted",  
16     "AdmittedToITU", "ThirtyDayFU", "Dyspnoea",  
17     "Comorbidity", "XRPositive"))  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

5 Imputation

5.0.0.0.0.1 This code generates 15 imputed datasets using the permuted mean matching method, based on the 'limcompletedata' dataset which has filtered the most relevant fields, with minimal missing data initially

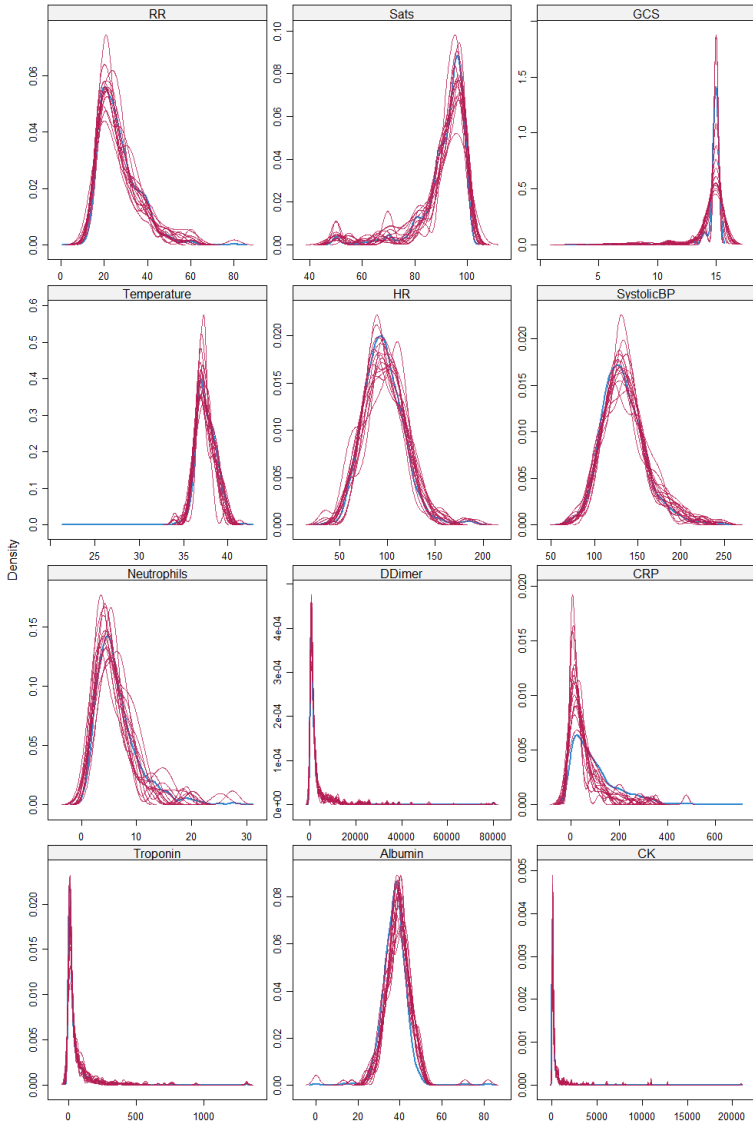
```
imputed <- mice(limcompletedata, m = 15,  
method = "pmm")
```

5.0.0.0.0.2 Imputation Diagnostics Density plot, this corresponds to supplementary figure 1:

```
densityplot(imputed)
```

22

5 Imputation



6 Propensity Score Matching

6.0.0.0.1 This code matches data in the imputed datasets on whether the XR was reported classical COVID or not, the matching is done based on the covariates Sex, Age, Comorbidity, Ethnicity and Respiratory Rate

```

library(MatchThem)
#### MatchThem package requires dependent variable to be coded as 0 or 1
imputed[["data"]][["XRPositive"]] %>% recode_factor("Positive" = "1",
"Negative" = "0") ->imputed[["data"]][["XRPositive"]]
matchthem(
  XRPositive ~ Sex + Age + Comorbidity + Ethnicity + RR,
  data = imputed,
  method = 'nearest',
  verbose = FALSE,
  replace = FALSE,
  ratio = 1,
  caliper = 0.2,
  m.order = "random",) -> matchedtest
### Set XRChest to unordered for binomial analyses
matchedtest[["datasets"]]c(1:15)[["XRChest"]] %>% factor(ordered = FALSE) ->
  matched2[["datasets"]]c(1:15)[["XRChest"]]

```

6.1 Match Balance Diagnostics

6.1.0.0.1 Creates plots and table with mean difference and distribution of values in covariates between XR +ve and -ve groups after matching across all imputed datasets:

```

#### Supplementary tables 1,2 and 3:
bal.tab(matchedtest)
#### Supplementary figure 2
bal.plot(matchedtest)
#### Supplementary figure 3:
bal.plot(matchedtest, var.name = "Age", type = "histogram",
  which = "both")
bal.plot(matchedtest, var.name = "Sex", type = "histogram",
  which = "both")
bal.plot(matchedtest, var.name = "Ethnicity",

```

```
1     type = "histogram", which = "both")
2
3
4     bal.plot(matchedtest, var.name = "RR", type = "histogram",
5             which = "both")
6     bal.plot(matchedtest, var.name = "Comorbidity",
7             type = "histogram", which = "both")
8     ### Supplementary figure 4:
9     love.plot(matchedtest)
```

7 Matched Demographics Table:

7.0.0.0.1 Stack matched imputed datasets into one large dataset and split into COVID +ve and -ve groups:

```
### 'all=FALSE' gets matched data only
stacked <- MatchThem::complete(matchedtest,
  n = c(1:15), all = FALSE)
stacked <- stacked %>% filter(.imp > 0)
```

7.0.0.0.2 Creates demographics table as above, but on propensity matched imputed datasets, corresponds to Table 4:

```
table4 <- CreateTableOne(strata = "OverallPos",
  data = stacked)
#### Means and SD kept as is, mean counts
#### calculated after dividing by 15 (as 15
#### imputed datasets)
```

7.0.0.0.3 Creates demographic table stratified by XR Positive or Negative on matched imputed datasets, corresponds to Table 5:

```
table5 <- CreateTableOne(strata = "XRPositive",
  data = stacked)
#### Means and SD kept as is, mean counts
#### calculated after dividing by 15 (as 15
#### imputed datasets)
```

7.0.0.0.4 Summary statistics for pooled data:

```
### Normal means sd
explanatorynorm <- c("Age", "Temperature",
  "HR", "SystolicBP")
summarynormalOverallPos <- stacked %>% group_by(OverallPos) %>%
```

26

7 Matched Demographics Table:

```
summarise_at(vars(explanatorynorm), list(mean.default,
sd))
summarynormalXRPositive <- stacked %>% group_by(XRPositive) %>%
summarise_at(vars(explanatorynorm), list(mean.default,
sd))

### Non normal medians and IQR
summarynormalOverallPos <- stacked %>% group_by(OverallPos) %>%
summarise_at(vars(explanatorynormal),
list(median, IQR))
summarynormalXRPositive <- stacked %>% group_by(XRPositive) %>%
summarise_at(vars(explanatorynormal),
list(median, IQR))
```

8 Diagnostic Accuracy

8.0.0.1 This section generates the diagnostic accuracy statistics (e.g. sensitivity, specificity) for CXR and CT with RT-PCR as the reference standard using the matched imputed datasets

8.0.0.2 This code creates a contingency table of False/ True Positives and Negatives for Chest X-ray taken from the demographic tables above:

```
contingxr <- matrix(c(305, 243, 125, 187),
  nrow = 2, ncol = 2)

colnames(contingxr) <- c("PCR+", "PCR-")

rownames(contingxr) <- c("XR+", "XR-")
```

8.0.0.2.1 This function calculates diagnostic accuracy test statistics:

```
xraccuracy <- epi.tests(contingxr, conf.level = 0.95)
```

8.0.0.3 Giving the diagnostic accuracy output for CXR in table 3:

```
xraccuracy
  Outcome +   Outcome -   Total
Test +       305         125     430
Test -       243         187     430
Total        548         312     860

Point estimates and 95 % CIs:
-----
Apparent prevalence           0.50 (0.47, 0.53)
True prevalence              0.64 (0.60, 0.67)
```


Sensitivity	0.56 (0.51, 0.60)
Specificity	0.60 (0.54, 0.65)
Positive predictive value	0.71 (0.66, 0.75)
Negative predictive value	0.43 (0.39, 0.48)
Positive likelihood ratio	1.39 (1.19, 1.62)
Negative likelihood ratio	0.74 (0.65, 0.84)

8.0.0.3.0.1 NB diagnostic accuracy values in table available in list view of xrraccuracy variable

8.1 CT Data and Accuracy

8.1.0.0.0.1 Only those with CT and RT PCR:

```
CTdata <- filter(data, is.na(data$CTBSTI) ==
  FALSE & is.na(data$RTPCR) == FALSE)
```

8.1.0.0.0.2 Select relevant variables

```
CTdata <- dplyr::select(CTdata, c("Age",
  "XRchest", "Ethnicity", "Sex", "RR",
  "Sats", "GCS", "Temperature", "HR", "SystolicBP",
  "DiastolicBP", "Neutrophils", "DDimer",
  "CRP", "Troponin", "OverallPos", "Admitted",
  "AdmittedToITU", "ThirtyDayFU", "Dyspnoea",
  "Comorbidity", "XRPositive", "OverallPos",
  "CTBSTI"))
```

8.1.0.0.0.3 Set RT-PCR as factor:

```
CTdata$OverallPos <- as.factor(CTdata$OverallPos)
```

8.1 CT Data and Accuracy

29

8.1.0.0.4 Rename 1 and 0 to Positive and Negative:

```
CTdata$CTPositive <- ifelse(CTdata$CTBSTI ==
  "1", "Positive", "Negative")
CTdata$CTPositive <- as.factor(CTdata$CTPositive)
```

8.1.0.0.5 Regression with CT as outcome variable:

```
CT <- finalfit(
  CTdata,
  "OverallPos",
  c(
    "Age",
    "Sex",
    "RR",
    "GCS",
    "CTPositive",
    "Temperature",
    "HR",
    "SystolicBP",
    "DiastolicBP",
    "Sats",
    "Dyspnoea",
    "Comorbidity"
  ),
  confint_level = 0.95
)
```

8.1.0.0.6 Contingency table of True/False Positives and Negatives for CT taken from Regression table:

```
contingct <- matrix(c(CT[7, 4], CT[6, 4],
  CT[7, 3], CT[6, 3]), nrow = 2, ncol = 2)
colnames(contingct) <- c("PCR+", "PCR-")
rownames(contingct) <- c("CT+", "CT-")
contingct <- substr(contingct, start = 1,
  stop = 3)
contingct <- sapply(contingct, as.numeric)
contingct <- matrix(contingct, nrow = 2,
  ncol = 2)
colnames(contingct) <- c("PCR+", "PCR-")
rownames(contingct) <- c("CT+", "CT-")
```

30

8 Diagnostic Accuracy

8.1.0.0.7 Diagnostic accuracy statistics for CT

```

epi.tests(contingct, conf.level = 0.95) -> ctaccuracy
      Outcome +   Outcome -   Total
Test +         162         55     217
Test -          29         56      85
Total          191        111     302

```

Point estimates and 95 % CIs:

```

-----
Apparent prevalence           0.72 (0.66, 0.77)
True prevalence               0.63 (0.58, 0.69)
Sensitivity                   0.85 (0.79, 0.90)
Specificity                   0.50 (0.41, 0.60)
Positive predictive value     0.75 (0.68, 0.80)
Negative predictive value     0.66 (0.55, 0.76)
Positive likelihood ratio     1.71 (1.41, 2.08)
Negative likelihood ratio     0.30 (0.21, 0.44)
-----

```

8.1.0.0.8 NB Diagnostic accuracy values found in list view rather than output

8.2 CT and XR accuracy comparison

8.2.0.1 In this section mean differences of diagnostic accuracy statistics between CT and Chest X-ray with confidence intervals and p-values are calculated

8.2.1 Sensitivity

8.2 CT and XR accuracy comp...

31

8.2.1.0.0.1 Upper confidence limit for difference in sensitivity

```
ubsens <- (ctaccuracy[["elements"]][["se.up"]] -  
xraccuracy[["elements"]][["se.low"]])
```

8.2.1.0.0.2 Lower confidence limit for difference in sensitivity

```
lbsens <- (ctaccuracy[["elements"]][["se.low"]] -  
xraccuracy[["elements"]][["se.up"]])
```

8.2.1.0.0.3 Mean difference in sensitivity

```
meansens <- ctaccuracy[["elements"]][["se"]] -  
xraccuracy[["elements"]][["se"]]
```

8.2.1.0.0.4 Standard error for sensitivity

```
sesens <- (ubsens - lbsens)/(2 * 1.96)
```

8.2.1.0.0.5 value for difference in sensitivity

```
zsens <- meansens/sesens
```

8.2.1.0.0.6 P-value for difference in sensitivity

```
psens <- exp(-0.717 * zsens - 0.416 * zsens^2)
```

32

8 Diagnostic Accuracy

8.2.1.0.0.7 Format values into 'mean difference (95% CI) p-value' rounded to 2 d.p.

```
diffsens <- sprintf("%s (%s-%s)", round(meansens,
  digits = 2), round(lbsens, digits = 2),
  round(ubsens, digits = 2))
diffsensp <- c(diffsens, psens)
```

8.2.1.0.0.8 Subsequent analyses in this section follow the code above

```
## Specificity
ubspec <- (ctaccuracy[["elements"]][["sp.up"]] -
  xraccuracy[["elements"]][["sp.low"]])
lbspec <- (ctaccuracy[["elements"]][["sp.low"]] -
  xraccuracy[["elements"]][["sp.up"]])
meanspec <- ctaccuracy[["elements"]][["sp"]] -
  xraccuracy[["elements"]][["sp"]]
sespec <- (ubspec - lbspec)/(2 * 1.96)
zspec <- meanspec/sespec
pspec <- exp(-0.717 * zspec - 0.416 * zspec^2)
diffspec <- sprintf("%s (%s-%s)", round(meanspec,
  digits = 2), round(lbspec, digits = 2),
  round(ubspec, digits = 2))
diffspecp <- c(diffspec, pspec)

ubda <- (ctaccuracy[["elements"]][["da.up"]] -
  xraccuracy[["elements"]][["da.low"]])
lbda <- (ctaccuracy[["elements"]][["da.low"]] -
  xraccuracy[["elements"]][["da.up"]])
meanda <- ctaccuracy[["elements"]][["da"]] -
  xraccuracy[["elements"]][["da"]]
seda <- (ubda - lbda)/(2 * 1.96)
zda <- meanda/seda
pda <- exp(-0.717 * zda - 0.416 * zda^2)
diffda <- sprintf("%s (%s-%s)", round(meanda,
  digits = 2), round(lbda, digits = 2),
  round(ubda, digits = 2))
diffdap <- c(diffda, pda)

## Positive Likelihood Ratio
ublrpos <- (ctaccuracy[["elements"]][["lrpos.up"]] -
  xraccuracy[["elements"]][["lrpos.low"]])
lblrpos <- (ctaccuracy[["elements"]][["lrpos.low"]] -
  xraccuracy[["elements"]][["lrpos.up"]])
meanlrpos <- ctaccuracy[["elements"]][["lrpos"]] -
  xraccuracy[["elements"]][["lrpos"]]
selrpos <- (ublrpos - lblrpos)/(2 * 1.96)
zlrpos <- meanlrpos/selrpos
plrpos <- exp(-0.717 * zlrpos - 0.416 * zlrpos^2)
diffplrpos <- sprintf("%s (%s-%s)", round(meanlrpos,
  digits = 2), round(lblrpos, digits = 2),
  round(ublrpos, digits = 2))
```

8.2 CT and XR accuracy comp...

33

```

1   round(ublrrpos, digits = 2))
2   diffplrposp <- c(diffplrpos, plrpos)
3   ## Negative Likelihood Ratios
4   ublrneg <- (ctaccuracy[["elements"]][["lrneg.up"]] -
5     xraccuracy[["elements"]][["lrneg.low"]])
6   lblrneg <- (ctaccuracy[["elements"]][["lrneg.low"]] -
7     xraccuracy[["elements"]][["lrneg.up"]])
8   meanlrneg <- ctaccuracy[["elements"]][["lrneg"]] -
9     xraccuracy[["elements"]][["lrneg"]]
10  selrneg <- (ublrneg - lblrneg)/(2 * 1.96)
11  zlrneg <- meanlrneg/selrneg
12  plrneg <- exp(-0.717 * zlrneg - 0.416 * zlrneg^2)
13  diffflrneg <- sprintf("%s (%s-%s)", round(meanlrneg,
14    digits = 2), round(lblrneg, digits = 2),
15    round(ublrneg, digits = 2))
16  diffflrnegp <- c(diffflrneg, plrneg)
17
18  ## Positive Predictive Value
19  ppv <- (ctaccuracy[["elements"]][["ppv.low"]] -
20    xraccuracy[["elements"]][["ppv.up"]])
21  meanppv <- ctaccuracy[["elements"]][["ppv"]] -
22    xraccuracy[["elements"]][["ppv"]]
23  seppv <- (ubppv - lbppv)/(2 * 1.96)
24  zppv <- meanppv/seppv
25  pppv <- exp(-0.717 * zppv - 0.416 * zppv^2)
26  diffppv <- sprintf("%s (%s-%s)", round(meanppv,
27    digits = 2), round(lbppv, digits = 2),
28    round(ubppv, digits = 2))
29  diffppvp <- c(diffppv, pppv)
30
31  ## Negative Predictive Value
32  npv <- (ctaccuracy[["elements"]][["npv.low"]] -
33    xraccuracy[["elements"]][["npv.up"]])
34  meannpv <- ctaccuracy[["elements"]][["npv"]] -
35    xraccuracy[["elements"]][["npv"]]
36  senpv <- (ubnpv - lbnpv)/(2 * 1.96)
37  znpv <- meannpv/senpv
38  pnpv <- exp(-0.717 * znpv - 0.416 * znpv^2)
39  diffnpv <- sprintf("%s (%s-%s)", round(meannpv,
40    digits = 2), round(lbnpv, digits = 2),
41    round(ubnpv, digits = 2))
42  diffnpvp <- c(diffnpv, pnpv)
43
44  ## Apparent Prevalence
45  meantp <- ctaccuracy[["elements"]][["tp"]] -
46    xraccuracy[["elements"]][["tp"]]
47  setp <- (ubtp - lbtp)/(2 * 1.96)
48  ztp <- meantp/setp
49  ptp <- exp(-0.717 * ztp - 0.416 * ztp^2)
50  diffftp <- sprintf("%s (%s-%s)", round(meantp,
51    digits = 2), round(lbtp, digits = 2),
52    round(ubtp, digits = 2))
53  diffftp <- c(diffftp, ptp)
54
55  ## True Prevalence
56  meanap <- ctaccuracy[["elements"]][["ap"]] -
57    xraccuracy[["elements"]][["ap"]]

```

```

seap <- (ubap - lbap)/(2 * 1.96)
zap <- meanap/seap
pap <- exp(-0.717 * zap - 0.416 * zap^2)
diffap <- sprintf("%s (%s-%s)", round(meanap,
  digits = 2), round(lbap, digits = 2),
  round(ubap, digits = 2))
diffapp <- c(diffap, pap)

```

8.3 Intermodality Agreement

8.3.0.0.1 This section contains code to analyse the level of agreement in the unmatched CT dataset which contains only data with CT, XR and RT-PCR

8.3.0.0.2 First- comparing CT and XR agreement

```

library(irr)
kappa2(c(CTdata$XRPositive, CTdata$CTPositive),
  weight = "squared")
d <- CTdata %>% select(c("CTPositive", "XRPositive"))
View(d)
kappa2(d, weight = "squared")

```

8.3.0.0.3 Output:

```

Cohen's Kappa for 2 Raters (Weights: squared)

Subjects = 287
Raters = 2
Kappa = 0.406

z = 7.14
p-value = 9.37e-13

```

8.3.0.0.4 The following code compares RT-PCR, CT and XR

```

d2 <- CTdata %>% select(c("CTPositive", "XRPositive",
  "OverallPos"))
View(d2)
kappam.fleiss(d2)

```

8.3 Intermodality Agreement

35

8.3.0.0.5 Output:

```
Fleiss' Kappa for m Raters

Subjects = 287
Raters = 3
Kappa = 0.361

z = 10.6
p-value = 0
```

8.3.1 Diagnostic Accuracy Analysis when Indeterminate Reports of CXR and CT are taken as positive

8.3.1.1 XR Indeterminates

8.3.1.1.0.1 New column for positive if indeterminate

```
stacked$XRIndPositive <- ifelse(stacked$XRchest ==
  "Classic COVID" | stacked$XRchest ==
  "Indeterminate", "Positive", "Negative")
stacked$XRIndPositive <- as.factor(stacked$XRIndPositive)
stackedpos <- stacked %>% filter(OverallPos ==
  "Positive")
stackedneg <- stacked %>% filter(OverallPos ==
  "Negative")
summary(stackedpos$XRIndPositive)
summary(stackedneg$XRIndPositive)

contingxrind <- matrix(c(441, 107, 186, 126),
  nrow = 2, ncol = 2)
colnames(contingxrind) <- c("PCR+", "PCR-")

rownames(contingxrind) <- c("XR+", "XR-")
xrindaccuracy <- epi.tests(contingxrind)
```


36

8 Diagnostic Accuracy

8.3.1.1.0.2 In this section mean differences of diagnostic accuracy statistics between CT (when CT indeterminates are not counted as positive) and Chest X-ray with confidence intervals and p-values are calculated, follows the same pattern as code previously

```
##### Sensitivity Upper confidence Limit for
##### difference in sensitivity

ubsens <- (ctaccuracy[["elements"]][["se.up"]] -
  xrindaccuracy[["elements"]][["se.low"]])
## Lower confidence limit for difference
## in sensitivity
lbsens <- (ctaccuracy[["elements"]][["se.low"]] -
  xrindaccuracy[["elements"]][["se.up"]])
## Mean difference in sensitivity
meansens <- ctaccuracy[["elements"]][["se"]] -
  xrindaccuracy[["elements"]][["se"]]
## Standard error for sensitivity
sesens <- (ubsens - lbsens)/(2 * 1.96)
## Z value for difference in sensitivity
zsens <- meansens/sesens
## P-value for difference in sensitivity
psens <- exp(-0.717 * zsens - 0.416 * zsens^2)
### Format values into 'mean difference
### (95% CI) p-value' rounded to 2 d.p.
diffsens <- sprintf("%s (%s-%s)", round(meansens,
  digits = 2), round(lbsens, digits = 2),
  round(ubsens, digits = 2))
diffsensp <- c(diffsens, psens)

### Subsequent analyses in this section
### follow the code above Specificity
ubspec <- (ctaccuracy[["elements"]][["sp.up"]] -
  xrindaccuracy[["elements"]][["sp.low"]])
lbspec <- (ctaccuracy[["elements"]][["sp.low"]] -
  xrindaccuracy[["elements"]][["sp.up"]])
meanspec <- ctaccuracy[["elements"]][["sp"]] -
  xrindaccuracy[["elements"]][["sp"]]
sespec <- (ubspec - lbspec)/(2 * 1.96)
zspec <- meanspec/sespec
pspec <- exp(-0.717 * zspec - 0.416 * zspec^2)
diffspec <- sprintf("%s (%s-%s)", round(meanspec,
  digits = 2), round(lbspec, digits = 2),
  round(ubspec, digits = 2))
diffspecp <- c(diffspec, pspec)

ubda <- (ctaccuracy[["elements"]][["da.up"]] -
  xrindaccuracy[["elements"]][["da.low"]])
lbda <- (ctaccuracy[["elements"]][["da.low"]] -
  xrindaccuracy[["elements"]][["da.up"]])
meanda <- ctaccuracy[["elements"]][["da"]] -
  xrindaccuracy[["elements"]][["da"]]
seda <- (ubda - lbda)/(2 * 1.96)
```

8.3 Intermodality Agreement

37

```

1  zda <- meanda/seda
2
3
4  pda <- exp(-0.717 * zda - 0.416 * zda^2)
5  diffda <- sprintf("%s (%s-%s)", round(meanda,
6    digits = 2), round(lbda, digits = 2),
7    round(ubda, digits = 2))
8  diffdap <- c(diffda, pda)
9  ## Positive Likelihood Ratio
10 ublrpos <- (ctaccuracy[["elements"]][["lrpos.up"]] -
11   xrindaccuracy[["elements"]][["lrpos.low"]])
12 lblrpos <- (ctaccuracy[["elements"]][["lrpos.low"]] -
13   xrindaccuracy[["elements"]][["lrpos.up"]])
14 meanlrpos <- ctaccuracy[["elements"]][["lrpos"]] -
15   xrindaccuracy[["elements"]][["lrpos"]]
16 selrpos <- (ublrpos - lblrpos)/(2 * 1.96)
17 zlrpos <- meanlrpos/selrpos
18 plrpos <- exp(-0.717 * zlrpos - 0.416 * zlrpos^2)
19 diffplrpos <- sprintf("%s (%s-%s)", round(meanlrpos,
20   digits = 2), round(lblrpos, digits = 2),
21   round(ublrpos, digits = 2))
22 diffplrposp <- c(diffplrpos, plrpos)
23 ## Negative Likelihood Ratios
24 ublrneg <- (ctaccuracy[["elements"]][["lrneg.up"]] -
25   xrindaccuracy[["elements"]][["lrneg.low"]])
26 lblrneg <- (ctaccuracy[["elements"]][["lrneg.low"]] -
27   xrindaccuracy[["elements"]][["lrneg.up"]])
28 meanlrneg <- ctaccuracy[["elements"]][["lrneg"]] -
29   xrindaccuracy[["elements"]][["lrneg"]]
30 selrneg <- (ublrneg - lblrneg)/(2 * 1.96)
31 zlrneg <- meanlrneg/selrneg
32 plrneg <- exp(-0.717 * zlrneg - 0.416 * zlrneg^2)
33 diffplrneg <- sprintf("%s (%s-%s)", round(meanlrneg,
34   digits = 2), round(lblrneg, digits = 2),
35   round(ublrneg, digits = 2))
36 diffplrnegp <- c(diffplrneg, plrneg)
37
38 ## Positive Predictive Value
39 ppv <- (ctaccuracy[["elements"]][["ppv.low"]] -
40   xrindaccuracy[["elements"]][["ppv.up"]])
41 meanppv <- ctaccuracy[["elements"]][["ppv"]] -
42   xrindaccuracy[["elements"]][["ppv"]]
43 seppv <- (ubppv - lbppv)/(2 * 1.96)
44 zppv <- meanppv/seppv
45 pppv <- exp(-0.717 * zppv - 0.416 * zppv^2)
46 diffppv <- sprintf("%s (%s-%s)", round(meanppv,
47   digits = 2), round(lbppv, digits = 2),
48   round(ubppv, digits = 2))
49 diffppvp <- c(diffppv, pppv)
50
51 ## Negative Predictive Value
52 npv <- (ctaccuracy[["elements"]][["npv.low"]] -
53   xrindaccuracy[["elements"]][["npv.up"]])
54 meannpv <- ctaccuracy[["elements"]][["npv"]] -
55   xrindaccuracy[["elements"]][["npv"]]
56 senpv <- (ubnpv - lbnpv)/(2 * 1.96)
57 znpv <- meannpv/senpv
58 pnpv <- exp(-0.717 * znpv - 0.416 * znpv^2)
59 diffnpv <- sprintf("%s (%s-%s)", round(meannpv,
60   digits = 2), round(lbnpv, digits = 2),

```

```

4     round(ubnpv, digits = 2))
5     diffnpvp <- c(diffnpv, pnpv)

6
7     ## True Prevalence
8     meantp <- ctaccuracy[["elements"]][["tp"]] -
9       xrindaccuracy[["elements"]][["tp"]]
10    setp <- (ubtp - lbtp)/(2 * 1.96)
11    ztp <- meantp/setp
12    ptp <- exp(-0.717 * ztp - 0.416 * ztp^2)
13    diffftp <- sprintf("%s (%s-%s)", round(meantp,
14      digits = 2), round(lbtp, digits = 2),
15      round(ubtp, digits = 2))
16    diffftpp <- c(diffftp, ptp)

17    ## Apparent Prevalence
18    meanap <- ctaccuracy[["elements"]][["ap"]] -
19      xrindaccuracy[["elements"]][["ap"]]
20    seap <- (ubap - lbap)/(2 * 1.96)
21    zap <- meanap/seap
22    pap <- exp(-0.717 * zap - 0.416 * zap^2)
23    diffpap <- sprintf("%s (%s-%s)", round(meanap,
24      digits = 2), round(lbap, digits = 2),
25      round(ubap, digits = 2))
26    diffpapp <- c(diffpap, pap)

```

8.3.1.2 CT Indeterminates

8.3.1.2.0.1 New column for positive if indeterminate

```

27
28
29     CTdata$CTIndPositive <- ifelse(CTdata$CTBSTI ==
30       "1" | CTdata$CTBSTI == "2", "Positive",
31       "Negative")
32     CTdata$CTIndPositive <- as.factor(CTdata$CTIndPositive)
33     valuesctind <- CTdata %>% group_by(OverallPos,
34       CTIndPositive) %>% summarise(n = n())
35     ctcontingind <- matrix(data = c(178, 13,
36       70, 41), nrow = 2, ncol = 2)
37
38     colnames(ctcontingind) <- c("PCR+ve", "PCR-ve")
39     rownames(ctcontingind) <- c("CT+ve", "CT-ve")
40     ctindaccuracy <- epi.tests(ctcontingind)

```

9 Pooled Regression after Multiple Imputation and Propensity Score Matching

9.0.0.0.0.1 Binomnal Logistic regression with RT-PCR as dependent variable

```

overallposmatchimp <- matchedtest %>% with(glm(formula(ff_formula(dependent =
  "OverallPos",
  explanatory = c("Age", "Ethnicity", "Sex",
    "RR", "GCS", "Temperature", "HR",
    "SystolicBP", "Neutrophils", "DDimer",
    "CRP", "Troponin", "Albumin", "CK",
    "Sats", "Admitted", "AdmittedToITU",
    "ThirtyDayFUTwo", "Dyspnoea", "Comorbidity",
    "XRchest"))), family = "binomial"),
  all = FALSE)
P <- overallposmatchimp %>% pool()
multivarpooleddoverallpos = P %>% fit2df(estimate_name = "OR (multiple
  imputation)",
  exp = TRUE)

```

9.0.0.0.0.2 'multivarpooleddoverallpos' produces multivariate odds ratios for each explanatory variable, corresponding to Table 4

9.0.1 Pooled Univariate Odds Ratios for OverallPos as dependent variable

9.0.1.0.0.1 This code is run with each of the explanatory variables in table 4 as arguments to produce their respective odds Ratios in table 4

```

overallposmatchimpunivar <- matchedtest %>%
  with(glm(formula(ff_formula(dependent = "OverallPos",

```

40

9 Pooled Regression after Multi...

```

    explanatory = "XRchest")), family = "binomial"))
P <- overallposmatchimpunivar %>% pool()
univarpooledoverallpos = univaroverallpos <- P %>%
  fit2df(estimate_name = "OR (univariate)",
        exp = TRUE)
univaroverallpos

```

9.0.2 Binomial Logistic Regression with Positive Chest X-ray Report as Dependent Variable

9.0.2.0.0.1 This code follows the format above to produce univariate and multivariate odds ratios for each explanatory variable for having a positive XR report

9.0.3 Univariate XRPositive as dependent

9.0.3.0.0.1 (different explanatory variables passed into function to produce Odds ratios for each)

```

XRchestmatchimp <- matchedtest %>% with(glm(formula(ff_formula(dependent =
  "XRPositive",
  explanatory = "Comorbidity")), family = "binomial"))
P <- XRchestmatchimp %>% pool()
multivarpooledXRchest = univarXRchest <- P %>%
  fit2df(estimate_name = "OR (univariate)",
        exp = TRUE)
univarXRchest

```

9.0.4 Multivariate XRPositive as dependent

```

XRchestmatchimp <- matchedtest %>% with(glm(formula(ff_formula(dependent =
  "XRPositive",
  explanatory = c("Age", "OverallPos",
  "Ethnicity", "Sex", "RR", "GCS",
  "Temperature", "HR", "SystolicBP",
  "Neutrophils", "DDimer", "CRP", "Troponin",
  "Albumin", "CK", "Sats", "Admitted",
  "AdmittedToITU", "ThirtyDayFUTwo",
  "Dyspnoea", "Comorbidity"))), family = "binomial"))
P <- XRchestmatchimp %>% pool()
multivarpooledXRchest = multivarXRchest <- P %>%
  fit2df(estimate_name = "OR (multivariate)",

```

9.1 Forest Plots

41

```

exp = TRUE)
multivarXRchest

```

9.0.5 Pooled Ordinal Logistic Regression with XRPositive as dependent

9.0.5.0.1 This code also produces multivariate odds ratios for table 5, however, uses ordinal linear regression after the CXR report variable is converted to an ordered categorical variable, with alternative pathology as the lowest and classic covid as the highest value (see table 3)

```

XRchestmatchimpord <- matchedtest %>% with(clm(formula = XRchest ~
  Age + OverallPos + Ethnicity + Sex +
  RR + GCS + Temperature + HR + SystolicBP +
  Neutrophils + DDimer + CRP + Troponin +
  Sats + Admitted + AdmittedToITU +
  ThirtyDayFUTwo + Dyspnoea + Comorbidity))
P <- pool(object = XRchestmatchimpord[["analyses"]])
multivarXRchestord = multivarXRchestord <- P %>%
  fit2df(estimate_name = "OR (multivariate)",
        exp = TRUE)
multivarXRchestord

```

9.1 Forest Plots

9.1.0.0.1 Creates forest plots for post matched regression tables above:

```

Figure1Forest <- read_excel("Figure1Forest.xlsx",
  col_types = c("text", "numeric", "numeric",
  "numeric", "text", "text"))

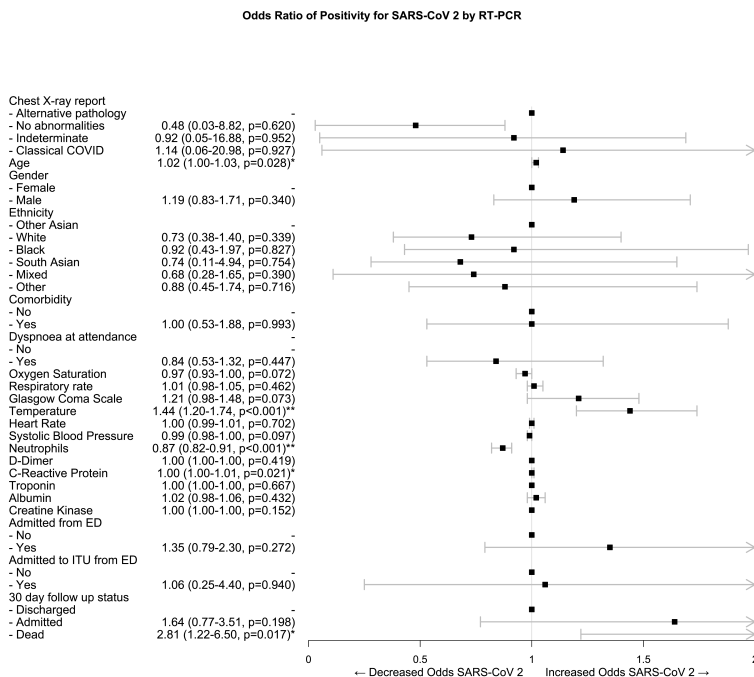
tabletext1 <- cbind(Figure1Forest$explanatory,
  Figure1Forest$summary)
forestplot(tabletext1, Figure1Forest$Mean,
  Figure1Forest$Lower, Figure1Forest$Upper,
  is.summary = FALSE, clip = c(0, 2), xlab = "<U+2190> Decreased Odds SARS-
  CoV 2 Increased Odds SARS-CoV 2 <U+2192>",
  zero = 1, cex = 0.9, lineheight = unit(6,
  "mm"), boxsize = 0.4, colgap = unit(6,
  "mm"), lwd.ci = 2, ci.vertices = TRUE,
  ci.vertices.height = 0.4, title = "Odds Ratio of Positivity for SARS-CoV 2
  by RT-PCR",
  txt_gp = fpTxtGp(label = gpar(cex = 1.25),
  ticks = gpar(cex = 1.1), xlab = gpar(cex = 1.2),

```

42 9 Pooled Regression after Multi...

```
title = gpar(cex = 1.2)), graphwidth = unit(200,
"mm"))
```

9.1.0.0.2 Figure 2:



9.1.0.0.3 Figure 3 (XR dependent):

```
Figure2Forest <- read_excel("Figure2Forest.xlsx",
col_types = c("text", "numeric", "numeric",
"numeric", "text", "text"))

tabletext2<-cbind(Figure2Forest$explanatory,Figure2Forest$summary)
forestplot (tabletext2, Figure2Forest$Mean,
Figure2Forest$Lower, Figure2Forest$Upper, is.summary = FALSE,
clip = c(0, 2),
```

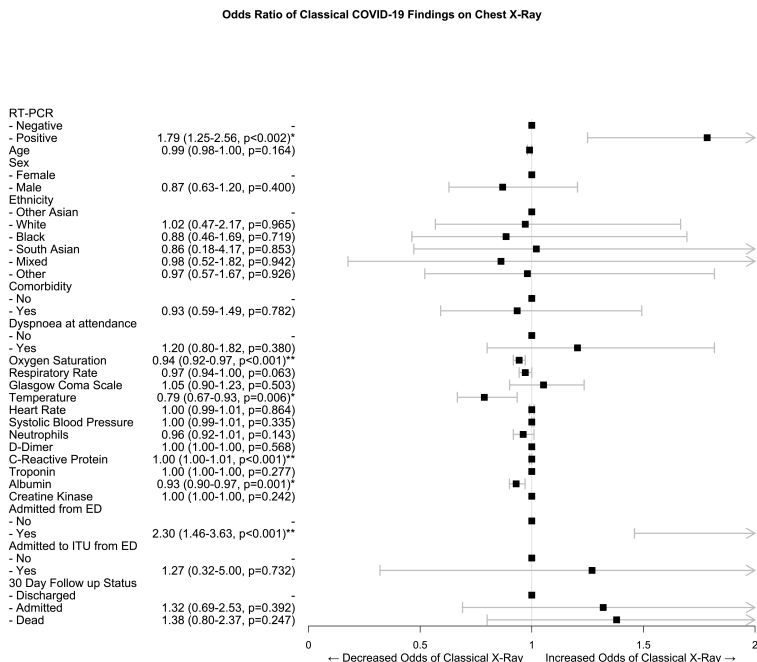
9.1 Forest Plots

43

```

xlab="\u2190 Decreased Odds of Classical X-Ray      Increased Odds
of Classical X-Ray \u2192",
zero=1, cex=0.9, lineheight = unit(6,"mm"), boxsize=0.5,
colgap=unit(6,"mm"),
lwd.ci=2, ci.vertices=TRUE, ci.vertices.height = 0.4,
title="Odds Ratio of Classical COVID-19 Findings on Chest X-Ray",
txt_gp=fpTxtGp(label=gpar(cex=1.25),
ticks=gpar(cex=1.1),
xlab=gpar(cex = 1.2),
title=gpar(cex = 1.2)),
graphwidth = unit(200,"mm")
)

```



9.2 Correlation Matrix

9.2.0.0.1 This section creates a plot of correlation between all the variables in the raw data

```
library(corrplot)
library(Hmisc)
```

9.2.0.0.2 Relevel factors so relevant value is first

```
data$XRPositive <- relevel(data$XRPositive,
  "Negative")

data$Admitted <- relevel(data$Admitted, "Discharged")
data$AdmittedToITU <- relevel(data$AdmittedToITU,
  "No")
```

9.2.0.0.3 New variable for correlation matrix

```
cor <- data
```

9.2.0.0.4 Remove variables with high missings/ data which won't work e.g. date, RT-PCR removed as it only represents initial ED swab, OverallPos used instead as this includes subsequent swabs in 30 days

```
cor <- subset(data, select = -c(CT, DateOfDeath, DateOfDischarge, RTPCR,
  DateOfVisit, DateOfSymptomOnset, FollowUpPos, TimeToDeath, NEWS))'
```

9.2.0.0.5 Format and re-name values

```
cor$CTPositive <- ifelse(cor$CTBSTI == "1",
  "Positive", "Negative")
cor$CTPositive <- as.factor(cor$CTPositive)
cor$CTPositive <- relevel(cor$CTPositive,
```

9.2 Correlation Matrix

45

```
"Negative")
cor$Death <- as.factor(ifelse(cor$ThirtyDayFU ==
  "4", "Dead", "Alive"))
cor$Death <- relevel(cor$Death, "Alive")
cor$OverallPos <- as.factor(cor$OverallPos)
cor <- sapply(cor, as.numeric)
```

9.2.0.0.6 Create new numerical correlation matrix

```
cormatrixall <- cor(cor, method = "spearman",
  use = "pairwise.complete.obs")
```

9.2.0.0.7 This variable also contains p-values so identification of only significant correlations is possible:

```
cormatrixall2 <- rcorr(as.matrix(cor), type = "spearman")
```

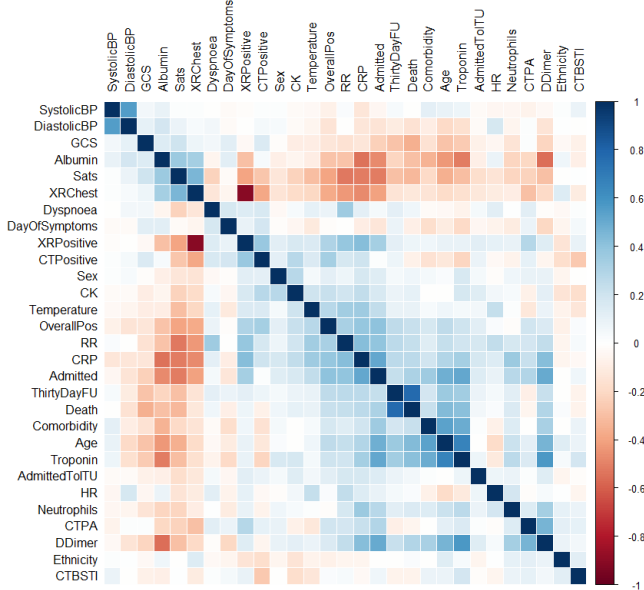
9.2.0.0.8 Function to create and format correlation matrix plot

```
corrplot(cormatrixall2$r, method = "color",
  type = "full", order = "hclust", p.mat = cormatrixall2$p,
  sig.level = 0.05, insig = "blank", tl.col = "black",
  outline = "white", title = "Correlation Matrix of Explanatory and Outcome
  Variables",
  line = -1, cex.main = 2, adj.main = 0.5)
```

46

9 Pooled Regression after Multi...

Correlation Matrix of Explanatory and Outcome Variables



9.3 STARD Flow Diagram

9.3.0.0.1 See instructions from <https://www.r-bloggers.com/flow-charts-in-r/>

9.3.0.0.2 Produces flow charts in Figure 1, (images need to be stretched out, output as svgs)

```
library(grid)
library(Gmisc)

grid.newpage()
# set some parameters to use repeatedly
leftx <- 0.25
```

9.3 STARD Flow Diagram

47

```

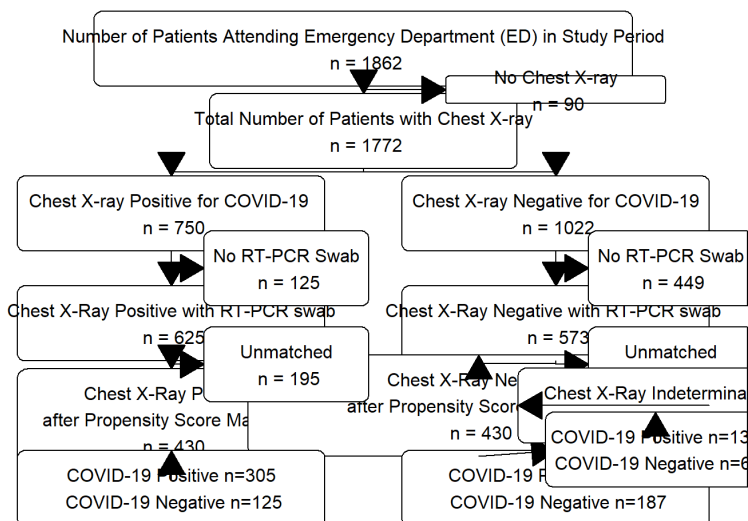
1 midx <- 0.5
2 rightx <- 0.75
3 width <- 0.4
4 gp <- gpar(fill = "white")
5 # create boxes
6 (totalattendance <- boxGrob("Number of Patients Attending Emergency Department
7 (ED) in Study Period\n n = 1862",
8 x = midx, y = 0.9, box_gp = gp, width = 0.7))
9
10 (numberwithxr <- boxGrob("Total Number of Patients with Chest X-ray\n n =
11 1772",
12 x = midx, y = 0.75, box_gp = gp, width = width))
13 # connect boxes like this
14 connectGrob(totalattendance, numberwithxr,
15 "v")
16
17 (numberwithoutxr <- boxGrob("No Chest X-ray\n n = 90",
18 x = rightx, y = 0.825, box_gp = gp, width = unit(2,
19 "inch"), height = 0.05))
20
21 connectGrob(totalattendance, numberwithoutxr,
22 "-")
23
24 (XRPos <- boxGrob("Chest X-ray Positive for COVID-19 \n n = 750",
25 x = leftx, y = 0.6, box_gp = gp, width = width))
26 (XRNeg <- boxGrob("Chest X-ray Negative for COVID-19\n n = 1022",
27 x = rightx, y = 0.6, box_gp = gp, width = width))
28
29 connectGrob(numberwithxr, XRPos, "N")
30 connectGrob(numberwithxr, XRNeg, "N")
31
32 (RTPCRXRPos <- boxGrob("Chest X-Ray Positive with RT-PCR swab\n n = 625",
33 x = leftx, y = 0.4, box_gp = gp, width = width))
34 (RTPCRXRNeg <- boxGrob("Chest X-Ray Negative with RT-PCR swab \n n = 573",
35 x = rightx, y = 0.4, box_gp = gp, width = width))
36
37 connectGrob(XRPos, RTPCRXRPos, "N")
38 connectGrob(XRNeg, RTPCRXRNeg, "N")
39
40 (NoRTPCRXRPos <- boxGrob("No RT-PCR Swab\n n = 125",
41 x = 0.4, y = 0.5, box_gp = gp, width = unit(1.5,
42 "inch")))
43 (NoRTPCRXRNeg <- boxGrob("No RT-PCR Swab\n n = 449",
44 x = 0.9, y = 0.5, box_gp = gp, width = unit(1.5,
45 "inch")))
46
47 connectGrob(XRPos, NoRTPCRXRPos, "-")
48 connectGrob(XRNeg, NoRTPCRXRNeg, "-")
49
50 (MatchedXRPos <- boxGrob("Chest X-Ray Positive \n after Propensity Score
51 Matching\n n = 430",
52 x = leftx, y = 0.225, box_gp = gp, width = width))
53 (MatchedXRNeg <- boxGrob("Chest X-Ray Negative \n after Propensity Score
54 Matching \n n = 430",
55 x = 0.65, y = 0.25, box_gp = gp, width = unit(4.2,
56 "inch")))
57
58 connectGrob(RTPCRXRPos, MatchedXRPos, "N")
59 connectGrob(RTPCRXRNeg, MatchedXRNeg, "N")

```

```
(UnmatchedXRPos <- boxGrob("Unmatched\n n = 195",  
  x = 0.4, y = 0.325, box_gp = gp, width = unit(1.5,  
  "inch")))  
(UnmatchedXRNeg <- boxGrob("Unmatched\n n = 143",  
  x = 0.9, y = 0.325, box_gp = gp, width = unit(1.5,  
  "inch")))  
  
connectGrob(RTPCRXRPos, UnmatchedXRPos, "-")  
connectGrob(RTPCRXRNeg, UnmatchedXRNeg, "L")  
  
(DiagXRPositive <- boxGrob("COVID-19 Positive n=305\n COVID-19 Negative n=125",  
  x = leftx, y = 0.1, box_gp = gp, width = width))  
(DiagXRNegative <- boxGrob("COVID-19 Positive n=243 \n COVID-19 Negative  
  n=187",  
  x = rightx, y = 0.1, box_gp = gp, width = width))  
  
connectGrob(MatchedXRPos, DiagXRPositive,  
  "N")  
connectGrob(MatchedXRNeg, DiagXRNegative,  
  "vertical")  
  
(XRInd <- boxGrob("Chest X-Ray Indeterminate \n n = 197",  
  x = 0.88, y = 0.25, box_gp = gp, width = unit(2.5,  
  "inch")))  
  
connectGrob(MatchedXRNeg, XRInd, "horizontal")  
  
(DiagXRInd <- boxGrob("COVID-19 Positive n=136\n COVID-19 Negative n=63",  
  x = 0.88, y = 0.17, box_gp = gp, width = unit(2,  
  "inch")))  
connectGrob(XRInd, DiagXRInd, "vertical")
```

9.3 STARD Flow Diagram

49



```
##### CT Flow Chart#####
```

```
grid.newpage()
(totalattendance <- boxGrob("Number of Patients Attending Emergency Department
(ED) in Study Period\n n = 1862",
  x = midx, y = 0.9, box_gp = gp, width = 0.7))

(numberwithCT <- boxGrob("Total Number with Chest Computed Tomography (CT)\n n
= 319",
  x = midx, y = 0.75, box_gp = gp, width = width))
connectGrob(totalattendance, numberwithCT,
  "vertical")

(numberwithoutCT <- boxGrob("No Chest CT\n n = 1543",
  x = rightx, y = 0.825, box_gp = gp, width = unit(2,
  "inch"), height = 0.05))

connectGrob(totalattendance, numberwithoutCT,
  "-")

(CTPos <- boxGrob("CT Positive for COVID-19 \n n = 232",
  x = leftx, y = 0.6, box_gp = gp, width = width))
(CTNeg <- boxGrob("CT Negative for COVID-19\n n = 87",
  x = rightx, y = 0.6, box_gp = gp, width = width))

connectGrob(numberwithCT, CTPos, "N")
connectGrob(numberwithCT, CTNeg, "N")

(RTPCRCTPos <- boxGrob("CT Positive with RT-PCR swab\n n = 217",
  x = leftx, y = 0.4, box_gp = gp, width = width))
```

50

9 Pooled Regression after Multi...

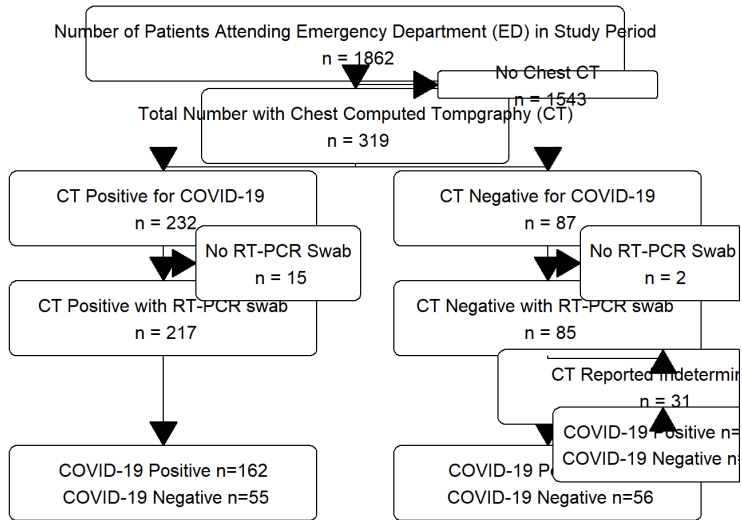
```

1 (RTPCRCTNeg <- boxGrob("CT Negative with RT-PCR swab \n n = 85",
2   x = rightx, y = 0.4, box_gp = gp, width = width))
3
4 connectGrob(CTPos, RTPCRCTPos, "N")
5 connectGrob(CTNeg, RTPCRCTNeg, "N")
6
7 (NoRTPCRCTPos <- boxGrob("No RT-PCR Swab\n n = 15",
8   x = 0.4, y = 0.5, box_gp = gp, width = unit(1.5,
9     "inch")))
10 (NoRTPCRCTNeg <- boxGrob("No RT-PCR Swab\n n = 2",
11   x = 0.9, y = 0.5, box_gp = gp, width = unit(1.5,
12     "inch")))
13 connectGrob(CTPos, NoRTPCRCTPos, "-")
14 connectGrob(CTNeg, NoRTPCRCTNeg, "-")
15
16 (DiagCTPositive <- boxGrob("COVID-19 Positive n=162\n COVID-19 Negative n=55",
17   x = leftx, y = 0.1, box_gp = gp, width = width))
18 (DiagCTNegative <- boxGrob("COVID-19 Positive n=29\n COVID-19 Negative n=56",
19   x = rightx, y = 0.1, box_gp = gp, width = width))
20 connectGrob(RTPCRCTPos, DiagCTPositive, "N")
21 connectGrob(RTPCRCTNeg, DiagCTNegative, "N")
22
23 (CTInd <- boxGrob("CT Reported Indeterminate \n n = 31",
24   x = 0.9, y = 0.275, box_gp = gp, width = unit(3,
25     "inch")))
26 connectGrob(RTPCRCTNeg, CTInd, "N")
27
28 (DiagCTInd <- boxGrob("COVID-19 Positive n=16\n COVID-19 Negative n=15",
29   x = 0.9, y = 0.17, box_gp = gp, width = unit(2,
30     "inch")))
31 connectGrob(CTInd, DiagCTInd, "vertical")
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

```

9.3 STARD Flow Diagram

51



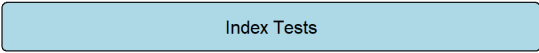
```

24  ### Labels###
25  grid.newpage()
26  (indextest <- boxGrob("Index Tests", x = midx,
27    y = 0.9, box_gp = gpar(fill = "light blue"),
28    width = 0.7))
29  (reftest <- boxGrob("Index Tests and Reference Standards",
30    x = midx, y = 0.4, box_gp = gpar(fill = "light blue"),
31    width = 0.7))
32  (finaldiag <- boxGrob("Final Diagnoses",
33    x = midx, y = 0.1, box_gp = gpar(fill = "light blue"),
34    width = 0.7))
  
```

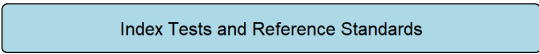

52

9 Pooled Regression after Multi...

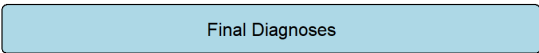
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50



Index Tests



Index Tests and Reference Standards

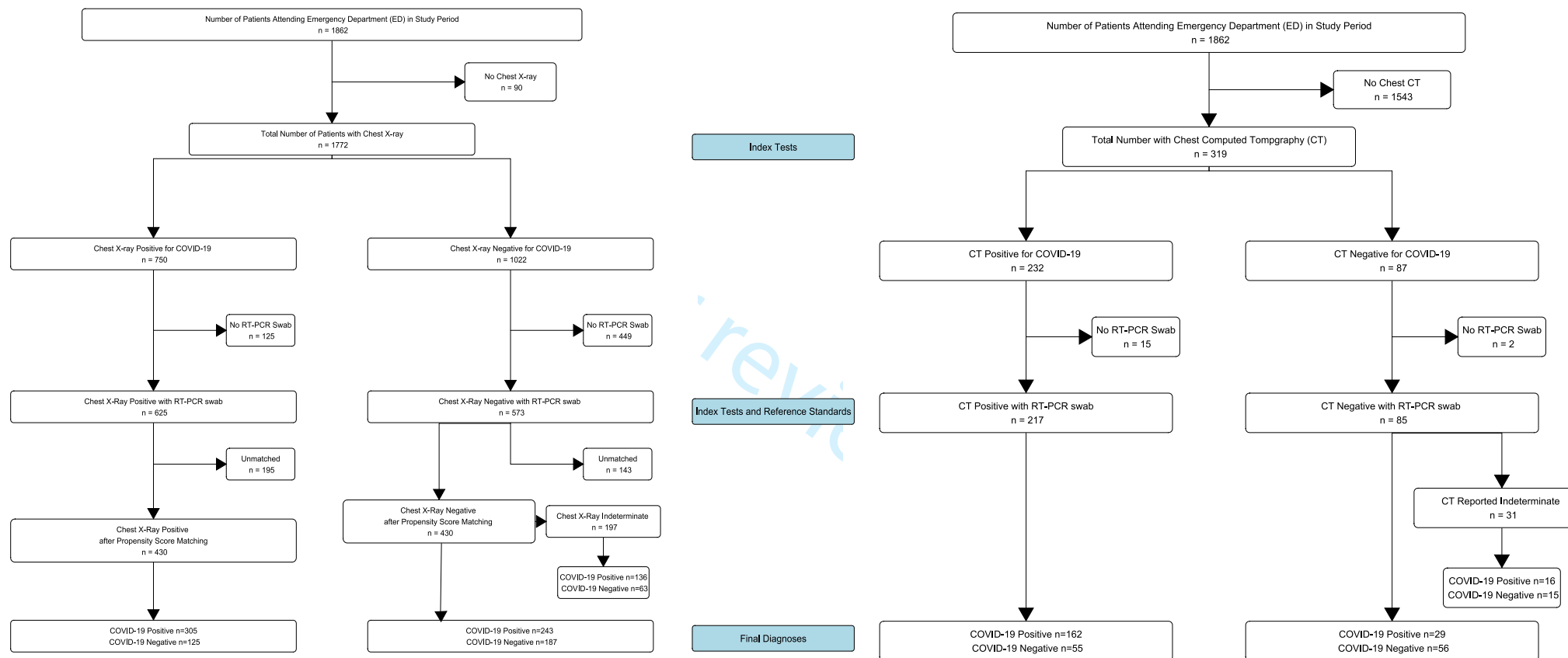


Final Diagnoses

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	5
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
<i>Participants</i>	6	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	5,20
	11	Rationale for choosing the reference standard (if alternatives exist)	N/A
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	5
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	20
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	12
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	6,7
	15	How indeterminate index test or reference standard results were handled	5
	16	How missing data on the index test and reference standard were handled	N/A, excluded
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	N/A
	18	Intended sample size and how it was determined	7
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	22, diagram below
	20	Baseline demographic and clinical characteristics of participants	21
	21a	Distribution of severity of disease in those with the target condition	21
	21b	Distribution of alternative diagnoses in those without the target condition	N/A
	22	Time interval and any clinical interventions between index test and reference standard	N/A
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	22
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	22
	25	Any adverse events from performing the index test or the reference standard	N/A
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12
	27	Implications for practice, including the intended use and clinical role of the index test	14
OTHER INFORMATION			
	28	Registration number and name of registry	N/A
	29	Where the full study protocol can be accessed	N/A
	30	Sources of funding and other support; role of funders	N/A



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Supplementary Figure- STARD Flow Diagram

For peer review only

BMJ Open

Diagnostic Accuracy of X-ray versus CT in COVID-19: A Propensity Matched Database Study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-042946.R3
Article Type:	Original research
Date Submitted by the Author:	07-Oct-2020
Complete List of Authors:	Borakati, Aditya; Royal Free Hospital, Emergency Department; University College London, Division of Surgery and Interventional Science Perera, Adrian; Royal Free Hospital, Emergency Department Johnson, James; Royal Free Hospital, Emergency Department Sood, Tara; Royal Free Hospital, Emergency Department
Primary Subject Heading:	Emergency medicine
Secondary Subject Heading:	Radiology and imaging, Medical management, Infectious diseases, Respiratory medicine, Diagnostics
Keywords:	COVID-19, Chest imaging < RADIOLOGY & IMAGING, ACCIDENT & EMERGENCY MEDICINE, GENERAL MEDICINE (see Internal Medicine), Diagnostic radiology < RADIOLOGY & IMAGING

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Diagnostic Accuracy of X-ray versus CT in COVID-19: A Propensity Matched Database Study

Authors: Aditya Borakati (0000-0003-0457-4944), Adrian Perera (0000-0002-1753-2586), James Johnson (0000-0001-7796-659X), Tara Sood (0000-0003-1620-916X)

Emergency Department, Royal Free Hospital, Pond Street, London, UK NW3 2QG

Aditya Borakati, Academic Foundation Doctor, Division of Surgery and Interventional Science, Royal Free Campus, University College London

Adrian Perera, Emergency Medicine and Intensive Care Medicine Specialist Registrar, Emergency Department, Royal Free Hospital

James Johnson, Foundation Doctor, Emergency Department, Royal Free Hospital

Tara Sood, Consultant in Emergency Medicine, Emergency Department, Royal Free Hospital

Correspondence:

Dr Aditya Borakati

Academic Foundation Doctor

Emergency Department, Royal Free Hospital, London, UK NW3 2QG

a.borakati@doctors.org.uk

Author contribution (CRediT) statement:

Aditya Borakati: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project Administration

Adrian Perera: Conceptualization, Methodology, Investigation, Writing- Review & Editing, Supervision, Project Administration

James Johnson: Investigation

Tara Sood: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project Administration

Aditya Borakati is the overall guarantor of this work.

Word count: 4236

Abstract

Objectives: To identify the diagnostic accuracy of common imaging modalities, chest X-ray (CXR) and computed tomography (CT) for diagnosis of COVID-19 in the general emergency population in the UK and to find the association between imaging features and outcomes in these patients.

Design: Retrospective analysis of electronic patient records

Setting: Tertiary academic health science centre and designated centre for high consequence infectious diseases in London, UK.

Participants: 1,198 patients who attended the emergency department with paired RT-PCR swabs for SARS-CoV 2 and CXR between 16th March and 16th April 2020

Main outcome measures: Sensitivity and specificity of CXR and CT for diagnosis of COVID-19 using the British Society of Thoracic Imaging reporting templates. Reference standard was any reverse transcriptase polymerase chain reaction (RT-PCR) positive naso-oropharyngeal swab within 30 days of attendance. Odds ratios of CXR in association with vital signs, laboratory values and 30-day outcomes were calculated.

Results: Sensitivity and specificity of CXR for COVID-19 diagnosis were 0.56 (95% CI 0.51-0.60) and 0.60 (95% CI 0.54-0.65), respectively. For CT scans these were 0.85 (95% CI 0.79-0.90) and 0.50 (95% CI 0.41-0.60), respectively. This gave a statistically significant mean increase in sensitivity with CT compared with CXR, of 29% (95% CI 19%-38%, $p < 0.0001$). Specificity was not significantly different between the two modalities.

Chest X-ray findings were not statistically significantly or clinically meaningfully associated with vital signs, laboratory parameters or 30-day outcomes.

Conclusions: Computed tomography has substantially improved diagnostic performance over CXR in COVID-19. CT should be strongly considered in the initial assessment for suspected COVID-19. This gives potential for increased sensitivity and considerably faster turnaround time, where capacity allows and balanced against excess radiation exposure risk.

Strengths and limitations

-Large, appropriately powered, study population consisting of all patients attending the emergency department rather than those solely with confirmed COVID-19; this allowed assessment of specificity for the imaging modalities and applicability to the general population who may attend medical personnel with other complaints, but have underlying SARS-CoV 2 infection

1
2
3
4 -Comprehensive statistical analyses were conducted to address confounding in
5 reporting of X-rays including propensity score matching and logistic regression to give a
6 'doubly robust' model
7
8

9
10 -Low amount of missing data and for secondary covariates only; multiple imputation was
11 performed with a good fit, however, observed data would be preferable to imputed data
12
13

14 -Single centre, retrospective study; potential for inter-reporter and inter-centre variability
15 in reporting
16
17

18
19 -Large proportion of patients excluded due to not having an RT-PCR swab,
20 predominantly, those with imaging reported as negative, this may bias the results
21 towards increased sensitivity and specificity
22
23
24

25 **Key words:** X-Rays, Computed Tomography, COVID-19, severe acute respiratory
26 syndrome coronavirus 2, Emergency Medicine, Diagnostic Imaging
27

28 **Statistical review:** The statistical methods in this manuscript and associated code have
29 been reviewed by Dr Federico Ricciardi of the Department of Statistical Science at
30 University College London and confirmed as robust and accurate.
31

32 **Ethical approval:** This study was registered with the local institutional review board
33 as a service evaluation using anonymised data only. No formal ethics committee
34 review was required.
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Declarations of Interests:** The authors have no relevant conflicts of interest to declare.
4 All authors have completed the [Unified Competing Interest form](#) (available on request
5 from the corresponding author) and declare: no support from any organisation for the
6 submitted work; no financial relationships with any organisations that might have an
7 interest in the submitted work in the previous three years, no other relationships or
8 activities that could appear to have influenced the submitted work.
9
10

11 **Transparency declaration:** The lead author (AB) affirms that this manuscript is an
12 honest, accurate, and transparent account of the study being reported; that no
13 important aspects of the study have been omitted; and that any discrepancies from the
14 study as planned (and, if relevant, registered) have been explained.
15
16

17 **Funding:** No funding was received for this study.
18

19 **Copyright:** The Corresponding Author has the right to grant on behalf of all authors
20 and does grant on behalf of all authors, a worldwide licence to the Publishers and its
21 licensees in perpetuity, in all forms, formats and media (whether known now or created
22 in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii)
23 translate the Contribution into other languages, create adaptations, reprints, include
24 within collections and create summaries, extracts and/or, abstracts of the Contribution,
25 iii) create any other derivative work(s) based on the Contribution, iv) to exploit all
26 subsidiary rights in the Contribution, v) the inclusion of electronic links from the
27 Contribution to third party material where-ever it may be located; and, vi) licence any
28 third party to do any or all of the above
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

SARS-CoV 2 and its resulting disease, COVID-19, have propagated exponentially worldwide, with over 10 million cases in 188 countries at the time of writing [1,2].

The gold standard for diagnosis of the virus is the detection of viral RNA through reverse transcriptase polymerase chain reaction (RT-PCR) of respiratory tract samples. However, this method has several limitations including: (1) low sensitivity at 59-71% [3,4], (2) relatively slow turnaround times ranging from a few hours to several days [5], (3) high expense and (4) limited capacity for testing in many countries.

Computed tomography (CT) has been shown to be more sensitive than RT-PCR for diagnosis of COVID-19 [3,4], while being significantly faster and cheaper. This comes with a large radiation dose and capacity is still lacking in many countries.

Plain film chest X-ray (CXR) is ubiquitous worldwide, with a 30-70x lower dose of radiation[6] and is commonly performed as an initial investigation in COVID-19.

Studies have so far only evaluated imaging in those with confirmed infection, it is therefore, not possible to calculate the specificity of these modalities. In the context of the global pandemic, infection may be widespread in the community, often with subclinical infection [7,8]. A reliable and rapid method to detect infection in the general population, who may present to medical personnel with other complaints, is needed.

Despite its extensive use, the specificity and sensitivity of CXR in the general emergency population for diagnosis of COVID-19 is unknown, nor how imaging features correlate with severity.

This study evaluated the performance of CXR in diagnosing COVID-19 in the emergency department (ED) of a tertiary care hospital.

Methods

This study was conducted at the Royal Free Hospital, London, UK, an academic health science centre and nationally designated centre for High Consequence Infectious Diseases [9].

All individuals attending the emergency department who had paired posterior-anterior chest radiographs and RT-PCR nasopharyngeal swabs for COVID-19 at the time of initial attendance between 16th March 2020 and 16th April 2020 were included.

All chest radiographs were reported by a Consultant Radiologist and rated on an ordinal scale for probability of COVID-19: Alternative pathology identified, not COVID-19; Clear chest, unlikely COVID; Indeterminate findings for COVID-19; Classical findings of COVID-19, based on the British Society of Thoracic Imaging's (BSTI) reporting templates (table 1) [10]. These were reported prior to RT-PCR results being available.

RT-PCR of swabs were performed in laboratories either at our centre or at a public health laboratory (PHE Collindale, UK), according to published national standard operating procedures [11]. Subsequent RT-PCR swabs taken within 30 days of initial ED attendance were also included.

CT scans performed within 30 days of attendance were retrieved. These were also reported according to the BSTI template. CT pulmonary angiogram was performed in the ED if the D-dimer was >5000 to exclude pulmonary emboli as per the locally agreed protocol. Subsequent CT chest imaging (whether pulmonary angiogram, contrast or non-contrast) was performed on the basis of clinical suspicion.

Prospectively recorded data was extracted from the Cerner Millennium electronic patient record system (Cerner Corp., Kansas City, MO).

Primary Outcome

The primary outcome is sensitivity and specificity of initial CXR, where it is reported as having classic COVID-19 features in the ED. This is compared with RT-PCR swab as the reference standard for diagnosis of COVID-19.

In the event of multiple RT-PCR swabs during one attendance, a single positive swab was taken as an overall positive test during one admission.

Secondary Outcomes

In those patients who also had CT scans of the thorax, the diagnostic accuracy was compared with CXR, with RT-PCR again as the reference standard. Sensitivity and specificity of CXR when X-rays reported as indeterminate or atypical for COVID-19 were classed as positive was also calculated.

Chest x-ray findings were correlated with vital signs at attendance and blood results, including: neutrophil counts, D-dimer and C-reactive protein, which have been associated with poor prognosis in COVID-19 [12]. Hazard ratios for clinical outcomes including direct admission to the intensive treatment unit (ITU) from ED and 30-day mortality rates were also calculated for CXR reporting categories.

Statistical Analysis

In the event of missing data, multiple imputation was conducted using a Predictive Mean Matching algorithm, via the MICE R package, as described previously [13]. Briefly, this uses a linear regression model (or logistic regression model for categorical data), to find a random value based on already observed data, to replace missing fields [14]. Variables without missing data fields were not modified. The number of imputed datasets was similar in number to the percentage of missing data as suggested by White and colleagues [15]. Balance diagnostics with density plots are available in supplementary file 1, adequate balance was assessed via visual inspection of imputed distributions with respect to the original dataset.

The propensity for a CXR being reported as positive or negative for COVID-19 was calculated for several plausible covariates that may influence image characteristics such as Age, Gender, Ethnicity, pre-existing morbidities and the respiratory rate of the patient using a generalised linear model [16]. X-ray positive and negative groups were then matched in each imputed dataset using the nearest neighbour algorithm, with a calliper of 0.2 of the propensity score standard deviation, without replacement and in random sequential order to obtain a 1:1 match as described elsewhere [17].

The balance of the match data was assessed quantitatively with mean differences of covariates in each of the X-ray groups pre- and post-matching, with a difference of less than 0.1% considered a good match (supplementary figures 1, 2). Visual inspection of matches was also conducted to ensure balance (supplementary figures 2, 3 and 4).

After matching, outcome data were adjusted for covariates including age, gender, ethnicity and presence of co-morbidities as well as C-reactive protein, D-dimer, troponin and vital signs. This was achieved by generalised linear regression for continuous outcome data, binomial logistic regression for binary categorical outcomes, or ordinal logistic regression in the case of CXR where it is the outcome variable.

These regression models were run on each imputed dataset and outcomes were pooled together across each imputed data set according to Rubin's rules [18] to give an overall estimate.

Diagnostic Accuracy Statistics

Chest X-rays reported as classical for COVID-19 as per the BSTI guidelines were considered a positive test in the primary analysis. In a secondary analysis X-rays reported as 'Indeterminate' or 'Atypical' for COVID-19 were also considered positive. All other reports were classified as a negative test. These were compared to nasopharyngeal aspirate RT-PCR results, which were taken as the gold standard for diagnosis of COVID-19. Where more than one swab was taken during the study period (up to 30 days after initial attendance), a single positive result was taken as a positive result for calculation of diagnostic accuracy statistics.

Sensitivity, specificity, predictive values and diagnostic accuracy were calculated using the propensity matched data after imputation and pooled across imputed datasets with 95% confidence intervals. Apparent and true prevalence based on this dataset are also given for interpretation of the predictive values.

Chest CTs were also reported according to the BSTI guidelines as with X-ray. Diagnostic statistics were calculated on raw, unmatched and non-imputed data (due to a low volume of

1
2
3 data for imputation and matching) in the same manner as X-ray. Mean differences and 95%
4 confidence intervals between CT and X-ray for each of the diagnostic statistics are given, with a
5 p-value calculated from the confidence intervals.
6

7 Agreement between the modalities was assessed on the unmatched dataset, in the sample
8 where CT, CXR and RT-PCR were all available using Cohen's (for two group agreement) and
9 Fleiss' Kappa (when all 3 are compared).
10

11 **Data Presentation**

12
13 Descriptive statistics are given as means and standard deviations for normally distributed data
14 and as medians and interquartile ranges for non-normally distributed data, before and after
15 matching and multiple imputation (for the latter these statistics are pooled across imputations).
16

17 Association of explanatory variables with SARS-CoV 2 and Chest X-ray findings are given as
18 odds ratios in uni- and multi-variate configurations.
19

20 Data was considered statistically significant if $p < 0.05$. Given the large number of analyses in
21 this paper, data is separately highlighted if $p < 0.001$ as a secondary threshold to address the
22 potential for false positives with multiple testing.
23

24 Analyses were conducted using R 4.0.0 (R Foundation for Statistical Computing, Vienna,
25 Austria) and code for the analyses is given in supplementary file 2.
26

27 **Sample size calculation**

28
29 In this study, the lower confidence interval for sensitivity of CXR as reported by Wong et al.[19]
30 (56%) was used as an estimate of likely sensitivity for COVID-19 . A power of 80% at an alpha
31 of 0.05 was used to calculate the sample size for sensitivities and specificities of 56%. This
32 gave an estimated sample size of 165 in each of the COVID-19 negative and positive groups
33 by RT-PCR (total 330).
34

35 **Ethical approval**

36
37 This study was registered with the local institutional review board as a service evaluation using
38 anonymised data only. No formal ethics committee review was required.
39

40 **Reporting Guidelines**

41 This study is reported according to the STARD guidelines [20] for diagnostic accuracy studies.
42
43

44 **Patient and Public Involvement**

45
46
47 Patients and the public were not involved in the design, conduct or dissemination of this study.
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

7

For peer review only

Results

1,198 eligible patients with both CXR and RT-PCR were identified in the study period (figure 1). Their characteristics, stratified by positivity for SARS-CoV 2 infection by RT-PCR is summarized in table 2. This showed that those with confirmed SARS-CoV 2 infection were more likely to be male, older (mean age 66.2 vs 62.7), have lower saturations, higher respiratory rates, whilst being more likely to be admitted and die within 30 days. There was a significant association with X-ray images and SARS-CoV 2 at baseline, with 59.6% having classic imaging features of COVID-19 in those with positive swabs versus 39.1% in those with negative swabs. There was 8.6% missing data overall in the dataset when variables with >50% missing data were removed and 15 imputations were performed on these remaining variables only.

After multiple imputation for missing data and pooled propensity score matching for plausible covariates that may affect CXR reporting, there were 430 patients in each of the X-ray positive and X-ray negative groups, for a total of 860 patients. Adequate balance was achieved for relevant covariates with a mean difference of <0.1 between groups (supplementary file 1, table 2).

Computed tomography (CT) was performed in 302 patients with paired RT-PCR during the same time period, with a median serial interval of 4.5 days (inter quartile range 0-17) after the initial attendance in ED and of these 30.1% were within one day of attendance.

Diagnostic Accuracy

The pooled sensitivity and specificity of CXR was 0.56 (95% CI 0.51-0.60) and 0.60 (95% CI 0.54-0.65), respectively (table 3). This gave an overall diagnostic accuracy of 0.57 (95% CI 0.54-0.61) for CXR.

In comparison, sensitivity and specificity for CT was 0.85 (95% CI 0.79-0.90) and 0.50 (95% CI 0.41-0.60), respectively. This gave a statistically significant mean increase in sensitivity with CT compared with CXR by 29% (95% CI 19%-38%, $p<0.0001$). Specificity was not significantly different between the two modalities. Diagnostic accuracy and negative predictive values were also significantly increased with CT at 0.15 and 0.22, respectively, while the negative likelihood ratio was significantly decreased at -0.44. This shows that the post-test odds of being negative for SARS-CoV 2 by RT-PCR with a negative CT is significantly lower.

Taking X-rays reported as indeterminate as positive increased the sensitivity of CXR to 0.80 (95% CI 0.77-0.84), however reduced specificity to 0.40 (95% CI 0.35-0.46). When CT scans reported as indeterminate are also considered positive the sensitivity of CT increased to 0.93 (95% CI 0.89-0.96), whilst mean specificity reduced to 0.37 (95% CI 0.28-0.47), although this was not statistically different from when indeterminate CTs are considered negative. Sensitivity of CT remained significantly higher than CXR (when indeterminates are considered positive for both) by 0.13 (95% CI 0.05-0.19, $p<0.001$), specificity was not significantly different between the two.

When comparing only the unimputed, unmatched subset of data where CT, RT-PCR and CXR were all performed ($n=287$), the agreement between CT and CXR was poor (Cohen's kappa 0.406). Agreement between all three modalities was also poor (Fleiss' kappa 0.361).

Association of CXR with Markers of Severity and Outcomes

Association of covariates with RT-PCR results is shown in table 4 and figure 2. Those who tested positive for SARS-CoV 2 by RT-PCR were significantly more likely to have a classical X-ray (OR 1.79 95% CI 1.25-2.56, $p < 0.002$) as would be expected by the diagnostic accuracy statistics (table 4). When the CXR report is considered as an ordered scale, worsening grades of report were associated more strongly with RT-PCR positivity, with a 1.94 x increase in odds for each grade.

Positive chest X-rays for COVID-19 were significantly associated with lower oxygen saturations (OR 0.94 95% CI 0.92-0.97, $p < 0.001$) and temperatures (2.30 95% CI 1.46-3.63, $p < 0.001$) in the ED following propensity score matching and multivariate regression (table 5 and figure 3).

They also had higher rates of admission to a general ward from the ED (OR 2.30 95% CI 1.46-3.63, $p < 0.001$) but no significant association with 30 day outcomes. There was a statistically significant increase in C-reactive protein with a positive X-ray, however, this is unlikely to be clinically meaningful due to the minimal association (OR 1.00 95% CI 1.00-1.01).

Discussion

This study is the first to report the diagnostic accuracy of CXR and CT in the general emergency population during the COVID-19 pandemic.

We show that CXR has poor sensitivity and specificity for diagnosis of COVID-19, whilst CT has 29% higher sensitivity. Many international radiological guidelines advise against CT scanning for the initial assessment of COVID-19 [21–23] or where there are equivocal CXRs, whilst in other countries CT scanning is performed as a routine first line investigation. Our results suggest that CT should be considered in the initial assessment of COVID-19 and that CXR findings poorly correlate with CT findings in this setting. We also show that indeterminate and non-classical features of COVID-19 significantly increase the sensitivity of these imaging modalities, without a significant decrease in specificity. Further, we demonstrate the limited prognostic value of CXR in COVID-19.

These findings mirror what has previously been reported in the literature on individuals with confirmed COVID-19. Wong et al. [19] showed a sensitivity of 59% for initial X-ray in confirmed COVID-19 infection, similarly initial case series in China also reported a sensitivity of 59.1% [12].

A recent in press article from Italy reported a much higher sensitivity of 89% for CXR in a smaller general emergency population (n=535) without confirmed COVID-19 at attendance [24]. However, this used telephone follow up for clinical symptoms of COVID-19 as a reference standard in individuals with an initial negative RT-PCR swab and appeared to classify any abnormal X-ray as positive, which may inflate this figure. When indeterminate CXRs are counted as positive in this study, the sensitivity would be in line with this Italian data. In the US, a study of patients attending an urgent care centre with confirmed COVID-19, showed a much lower sensitivity at 41.7% for CXR where any abnormality was found on the images [25]. In this study 97/636 reports were re-classified from 'possible pneumonia' to 'normal' on second reading from a radiologist, highlighting the importance of inter-rater agreement and possibly explaining this low estimate.

Computed tomography has been reported in previous studies as being up to 98% sensitive for the diagnosis of COVID-19 in confirmed patients, when RT-PCR is used as the reference standard in confirmed patients [3,4]. These studies used any potential features of COVID-19 (e.g. ground glass opacification, crazy paving) as a positive scan, regardless of spatial distribution or features more characteristic of alternate pathology, unlike the BSTI guidelines used in this study. When we classified indeterminate CTs as positive like these latter studies, our estimates match their sensitivity values.

Consequently, a much lower specificity of 25% was found with initial RT-PCR in previous literature; however, it is reported that 10 out of 15 (67%) of these negatives subsequently tested positive. This would give an adjusted specificity of 75%, considering subsequent swabs as a reference standard, which combined with the wider CIs in these smaller studies, would bring estimates in line with the specificity in this paper. More recent meta-analyses have placed the pooled sensitivity of CT in populations with confirmed COVID-19 only, at 89.76% (95% CI 84.42%-93.84%) [26], in line with the estimates identified here.

There is limited coverage in the literature on association of X-ray findings with clinical and laboratory parameters and outcomes in the COVID-19 pandemic. This study demonstrates that classic appearances of COVID-19 were associated with initial lower saturations and lower

1
2
3 temperature. Volume opacification of the lung fields were not quantified as a surrogate of
4 severity; however, the use of the BSTI grading templates does this somewhat. When the X-ray
5 report is considered as a graded scale from low likelihood of COVID-19 and severity to high
6 likelihood and severity of disease there was no significant difference in association with vital
7 signs or laboratory parameters compared with when the X-ray report is merely considered as
8 a binary positive and negative outcome for COVID-19.
9

10
11 Borghesi and colleagues have devised a X-ray grading system, the Brixia score, for severity in
12 admitted patients with confirmed SARS-CoV 2 infection [27]. They further found a significant
13 increase in the severity of CXR by this scoring system in those who were discharged versus
14 those who died [28,29].
15

16 Here, there were no relevant associations between CXR and laboratory values. This analysis
17 also found no association with positive X-rays and 30 day outcomes after multivariate analyses,
18 unlike Borghese et al. This is also in contrast to Guan et al. who found higher rates of ITU
19 admission and death in those with positive imaging findings. However, these studies analysed
20 only those with confirmed SARS-CoV 2 infection. The divergence observed in this study may be
21 due to classifying those with 'Alternate pathology/ Indeterminate' or 'CVXC3/ CVXC2' as per the
22 BSTI templates, negative for COVID-19 in these analyses. Other studies classified X-rays with
23 any abnormality as a positive for COVID-19. These alternate distributions may still be reflective
24 of underlying COVID-19 and we show significantly higher sensitivity for both CT and CXR when
25 these are classed as positive. It may be that correlating indeterminate X-rays (in addition to
26 classical images) with vitals, laboratory markers and 30 day outcomes would yield significant
27 associations. However this may be unlikely, Xu and Zhang et al. found that those with classical
28 bilateral and diffuse involvement in upper and lower lobes had more severe disease than those
29 without [30,31].
30
31

32 There were a total of 70 confirmed pulmonary emboli (PEs) in our dataset out of 114 CT
33 pulmonary angiograms (61.0%, 5.84% of all patients attending) performed in the emergency
34 department. The incidence of venous thromboembolism is reported as ranging from 20-30% in
35 admitted confirmed SARS-CoV 2 positive patients [32]. Although we have not focused on this
36 cohort of patients in this paper for the sake of brevity and simplicity, this high incidence
37 represents a further advantage for CT over CXR.
38
39

40 CT, even with the absence of contrast has been shown to have strong accuracy in the
41 diagnosis of pulmonary emboli and many imaging features correlate with the presence of
42 pulmonary emboli. Sensitivities of non-contrast CT for diagnosis of PE have been reported at
43 96.9% and specificity at 71.9% [33,34].
44

45 We therefore see the advantages of CT scanning in COVID-19 as threefold over other
46 diagnostic techniques: 1) The rapid turnaround; 2) Increased sensitivity and 3) The possibility to
47 identify pulmonary emboli in COVID-19, which are a significant burden in this group.
48

49 This must be balanced against the excess radiation exposure with CT. Radiation from CT and
50 its association with carcinogenesis is difficult to quantify and no definitive epidemiological
51 studies have confirmed excess risk of cancer[35]. Modern CT scanners and software
52 reconstruction techniques continue to minimise radiation exposure and many ways of shielding
53 parts of the body from radiation also exist. Nevertheless, the excess risk of lifetime cancer is
54 estimated at 1 per 5,000 CT examinations[36].
55
56
57

Strengths and Limitations

This study is the largest conducted on imaging in the COVID-19 pandemic and one of the only studies conducted in the general population during the pandemic rather than only in confirmed patients. This enables greater applicability to the clinical setting where the diagnosis is uncertain, in addition to being able to calculate specificity, which is not possible in most studies. This study was planned to be powered to detect a sensitivity and specificity of 56% for CXR and greatly exceeded the sample size necessary for this.

Comprehensive statistical analyses were conducted to account for confounders in both factors influencing reporting of CXR and in factors affecting outcomes. The data was collected from prospectively maintained electronic records; however, the retrieval took place retrospectively with its inherent disadvantages. We were not able to collect data on several relevant covariates such as specific comorbidities or markers of severity such as lymphocytes. Furthermore, there was a significant amount of missing data that required multiple imputation to replace, although the fit of this imputed data was good, actual, observed data would be ideal.

Inter-rater reliability of imaging reports was not analysed in this paper and there was the potential for individual radiologists to have greater or lesser accuracy in the diagnosis of COVID-19. The literature has so far suggested a strong degree of agreement between radiologists in reporting of COVID-19 images [28].

The single centre nature of this study further limits generalisability and the potential for inter-hospital disagreement in imaging, in addition to inter-rater disagreement.

Finally, the median time for patients to receive a CT scan was 4.5 days following initial attendance to ED. Thus, the scans may not have been directly comparable to the initial CXR, both because of the progression of disease and because the SARS-CoV 2 status may have been confirmed at this point, biasing the reporting of these scans.

Future Research

Although this study used RT-PCR of nasopharyngeal swabs as a reference standard, newer methods exist for diagnosis of the disease. Serological assays for antibodies against SARS-CoV 2 are increasingly available and may represent a better gold standard in diagnosis for future research [37]. RT-PCR is limited by swabbing technique for nasopharyngeal samples and the fact that the virus is more avid in the lower respiratory tract [38]. However, many patients may not seroconvert prior to death limiting this test to survivors only.

Point of care lung ultrasound is a new technique for diagnosis of COVID-19 which may mitigate many of the issues noted with the modalities discussed so far. It has no radiation, is fast, cheap and may be able to detect lower respiratory tract disease unlike nasopharyngeal swab.

However, there is limited evidence beyond small case series on its diagnostic accuracy [39–41]. Further, like other ultrasound techniques accuracy will likely be operator dependent [42] and experience will need to be built up for robust results in evaluating suspected COVID-19.

Finally, much research has been conducted in the use of artificial intelligence techniques to correctly diagnose COVID-19 based on imaging [43–45]. These techniques would obviate capacity limitations in reporting imaging as well as eliminate inter-reporter variability. However, as with any supervised machine learning technique, large, generalisable datasets, with correctly

1
2
3 pre-classified positive and negative cases (which in turn will depend on a truly
4 accurate reference standard) are needed [46].
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Conclusion

Chest X-ray has poor sensitivity and specificity in diagnosing COVID-19 in the general population during the pandemic. CT scanning has demonstrated excellent sensitivity and should strongly be considered during the pandemic in the initial assessment of COVID-19. This needs to be balanced against the risk of excess radiation with CT, where capacity allows.

Acknowledgements

We would like to thank Scott Wilson from the Royal Free Hospital's clinical practice group analytics department for retrieving the data from the hospital's data warehouse.

We would like to thank Dr Federico Ricciardi of the Department of Statistical Science and PRIMENT Clinical Trials Unit at University College London for reviewing the statistical methods in this study.

1
2
3 **Data availability**
4

5 Anonymised data is available on reasonable request from the corresponding author. Analysis
6 scripts are attached as a supplementary file.
7

8 **Declarations of Interest**
9

10 The authors declare no conflicts of interest.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

References

- 1 COVID-19 Map. Johns Hopkins Coronavirus Resour. Cent. <https://coronavirus.jhu.edu/map.html> (accessed 30 Jun 2020).
- 2 Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020;**0**. doi:10/ggnsjk
- 3 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**. doi:10/ggmw6p
- 4 Fang Y, Zhang H, Xie J, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* 2020;**200432**. doi:10/ggnnkj
- 5 Konrad R, Eberle U, Dangel A, et al. Rapid establishment of laboratory diagnostics for the novel coronavirus SARS-CoV-2 in Bavaria, Germany, February 2020. *Eurosurveillance* 2020;**25**. doi:10/ggp6bw
- 6 Lin EC. Radiation Risk From Medical Imaging. *Mayo Clin Proc* 2010;**85**:1142–6. doi:10/c445mk
- 7 Mizumoto K, Kagaya K, Zarebski A, et al. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance* 2020;**25**:2000180. doi:10/ggn4bd
- 8 Arons MM, Hatfield KM, Reddy SC, et al. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. *N Engl J Med* 2020;**382**:2081–90. doi:10/ggszfg
- 9 High consequence infectious diseases (HCID). GOV.UK. <https://www.gov.uk/guidance/high-consequence-infectious-diseases-hcid> (accessed 24 May 2020).
- 10 Desai S. COVID-19 BSTI Reporting templates | The British Society of Thoracic Imaging. Br. Soc. Thorac. Imaging. 2020. <https://www.bsti.org.uk/covid-19-resources/covid-19-bsti-reporting-templates/> (accessed 29 Apr 2020).
- 11 NHS England. Guidance and Standard Operating Procedure: COVID-19 virus testing in NHS Laboratories. 2020. <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 24 May 2020).
- 12 Guan W, Ni Z, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020;**0**:null. doi:10/ggm6dh
- 13 Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. *J Stat Softw* 2011;**45**. doi:10/gdqc9c
- 14 Ginkel JR van, Linting M, Rippe RCA, et al. Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data. *J Pers Assess* 2020;**102**:297–308. doi:10/gftj5w

- 15 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;**30**:377–99. doi:10.1002/sim.4067
- 16 He H, McDermott MP. A robust method using propensity score stratification for correcting verification bias for binary tests. *Biostat Oxf Engl* 2012;**13**:32–47. doi:10/c4jzn6
- 17 Ho DE, Imai K, King G, et al. MatchIt : Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw* 2011;**42**. doi:10/gdwtnq
- 18 Marshall A, Altman DG, Holder RL, et al. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009;**9**:57. doi:10.1186/1471-2288-9-57
- 19 Wong HYF, Lam HYS, Fong AH-T, et al. Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients. *Radiology* 2020;:201160. doi:10/ggqbp4
- 20 Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016;**6**:e012799. doi:10.1136/bmjopen-2016-012799
- 21 Rubin GD, Ryerson CJ, Haramati LB, et al. The Role of Chest Imaging in Patient Management during the COVID-19 Pandemic: A Multinational Consensus Statement from the Fleischner Society. *Radiology* 2020;:201365. doi:10/ggrmg4
- 22 ACR Recommendations for the use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection. <https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection> (accessed 5 Jun 2020).
- 23 British Society of Thoracic Imaging. COVID-19: BSTI STATEMENT AND GUIDANCE. 2020;:1. https://www.bsti.org.uk/media/resources/files/COVID11.3.20_2.pdf (accessed 5 Jun 2020).
- 24 Schiaffino S, Tritella S, Cozzi A, et al. Diagnostic Performance of Chest X-Ray for COVID-19 Pneumonia During the SARS-CoV-2 Pandemic in Lombardy, Italy. *J Thorac Imaging* 2020;**Publish Ahead of Print**. doi:10/ggx268
- 25 Weinstock MB, Echenique A, Russell JW, et al. Chest X-Ray Findings in 636 Ambulatory Patients with COVID-19 Presenting to an Urgent Care Center: A Normal Chest X-Ray Is no Guarantee. ;:10.
- 26 Bao C, Liu X, Zhang H, et al. Coronavirus Disease 2019 (COVID-19) CT Findings: A Systematic Review and Meta-analysis. *J Am Coll Radiol* 2020;**17**:701–9. doi:10/ggr28p
- 27 Borghesi A, Zigliani A, Masciullo R, et al. Radiographic severity index in COVID-19 pneumonia: relationship to age and sex in 783 Italian patients. *Radiol Med (Torino)* 2020;**125**:461–4. doi:10/ggtvwp

- 1
2
3 28 Borghesi A, Maroldi R. COVID-19 outbreak in Italy: experimental chest X-ray scoring system
4 for quantifying and monitoring disease progression. *Radiol Med (Torino)* 2020;**125**:509–13.
5 doi:10/ggtvwn
6
- 7 29 Borghesi A, Zigliani A, Golemi S, et al. Chest X-ray severity index as a predictor of in-
8 hospital mortality in coronavirus disease 2019: A study of 302 patients from Italy. *Int J Infect*
9 *Dis* 2020;**96**:291–3. doi:10.1016/j.ijid.2020.05.021
10
- 11 30 Xu Y-H, Dong J-H, An W-M, et al. Clinical and computed tomographic imaging features
12 of novel coronavirus pneumonia caused by SARS-CoV-2. *J Infect* 2020;**80**:394–400.
13 doi:10/ggqwf3
14
- 15 31 Zhang J-J, Dong X, Cao Y-Y, et al. Clinical characteristics of 140 patients infected with
16 SARS-CoV-2 in Wuhan, China. *Allergy Published Online First*: 19 February 2020.
17 doi:10/ggpx6g
18
- 19 32 Lodigiani C, Iapichino G, Carenzo L, et al. Venous and arterial thromboembolic
20 complications in COVID-19 patients admitted to an academic hospital in Milan, Italy. *Thromb*
21 *Res* 2020;**191**:9–14. doi:10/ggvctf
22
- 23 33 Chien C-H, Shih F-C, Chen C-Y, et al. Unenhanced multidetector computed
24 tomography findings in acute central pulmonary embolism. *BMC Med Imaging*
25 2019;**19**:65. doi:10/ggzg85
26
- 27 34 Mohamed N, Othman MoustafaHM, Hassan L, et al. The accuracy of non-contrast chest
28 computed tomographic Scan in the detection of pulmonary thromboembolism. *J Curr Med*
29 *Res Pract* 2019;**4**:61. doi:10/ggzg83
30
- 31 35 McCollough CH, Bushberg JT, Fletcher JG, et al. Answers to Common Questions About the
32 Use and Safety of CT Scans. *Mayo Clin Proc* 2015;**90**:1380–92. doi:10/f3jgqx
33
- 34 36 Moser JB, Sheard SL, Edyvean S, et al. Radiation dose-reduction strategies in thoracic CT.
35 *Clin Radiol* 2017;**72**:407–20. doi:10/f95q7p
36
- 37 37 Long Q-X, Liu B-Z, Deng H-J, et al. Antibody responses to SARS-CoV-2 in patients with
38 COVID-19. *Nat Med* 2020;:1–4. doi:10.1038/s41591-020-0897-1
39
- 40 38 Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in Different Types of Clinical
41 Specimens. *JAMA* 2020;**323**:1843–4. doi:10/ggpp6h
42
- 43 39 Smith MJ, Hayward SA, Innes SM, et al. Point-of-care lung ultrasound in patients with
44 COVID-19 – a narrative review. *Anaesthesia*;n/a. doi:10/ggr2p7
45
- 46 40 Haaksma ME, Heldeweg MLA, Matta JEL, et al. Lung ultrasound findings in patients
47 with novel SARS-CoV2. *medRxiv* 2020;:2020.05.18.20105775.
48 doi:10.1101/2020.05.18.20105775
49
- 50 41 Benchoufi M, Bokobza J, Chauvin AA, et al. Lung injury in patients with or suspected
51 COVID-19 : a comparison between lung ultrasound and chest CT-scanner severity
52 assessments, an observational study. *medRxiv* 2020;:2020.04.24.20069633.
53 doi:10.1101/2020.04.24.20069633
54
- 55
56
57
58
59
60

- 1
2
3 42 Fine D, Perring S, Herbetko J, et al. Three-dimensional (3D) ultrasound imaging of the
4 gallbladder and dilated biliary tree: reconstruction from real-time B-scans. *Br J Radiol*
5 1991;**64**:1056–7. doi:10/fqr9mh
6
7 43 Shi F, Wang J, Shi J, et al. Review of Artificial Intelligence Techniques in Imaging Data
8 Acquisition, Segmentation and Diagnosis for COVID-19. *IEEE Rev Biomed Eng* 2020;:1–1.
9 doi:10/ggs2km
10
11 44 Li L, Qin L, Xu Z, et al. Artificial Intelligence Distinguishes COVID-19 from Community
12 Acquired Pneumonia on Chest CT. *Radiology* 2020;:200905. doi:10/ggpdgp
13
14 45 Wang L, Wong A. COVID-Net: A Tailored Deep Convolutional Neural Network Design
15 for Detection of COVID-19 Cases from Chest X-Ray Images. *ArXiv200309871 Cs Eess*
16 Published Online First: 11 May 2020. <http://arxiv.org/abs/2003.09871> (accessed 13 Jun
17 2020).
18
19 46 Kotsiantis SB. Use of machine learning techniques for educational proposes: a
20 decision support system for forecasting students' grades. *Artif Intell Rev* 2012;**37**:331–
21 44. doi:10/fmbng4
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Tables

Ordinal scale for study	BSTI grade	Features on X-ray
		Alternative pathology such as
	CVCX3- Non-COVID-19 pneumothorax with no features of COVID-19 identified	
1	CVCX0- Normal	No pathology seen
2	CVCX2- Indeterminate for COVID-19 or atypical features	Poor quality film or central/ basal consolidation
3	CVCX1- Classic findings of COVID-19	Peripheral ground glass opacities

Table 1- Ordinal scale used in this study based on the British Society of Thoracic Imaging (BSTI) Reporting Template [10]

	SARS-CoV 2 RT-PCR		p-value	Missing (%)
	Negative	Positive		
n (%)	435 (36.3)	763 (63.7)		
Number of Swabs (%)	810 (48.3)	868 (51.7)		
Age (mean (SD))	62.74 (17.72)	66.18 (17.58)	0.001*	0
Ethnicity			0.097	19
Other- Asian (%)	29 (8.0)	72 (11.8)		
South- Asian (%)	27 (7.5)	38 (6.2)		
Black (%)	41 (11.4)	91 (14.9)		
Mixed (%)	6 (1.7)	6 (1.0)		
Other (%)	56 (15.5)	105 (17.2)		
White (%)	202 (56.0)	297 (48.8)		
Sex – Male (%)	233 (53.6)	480 (62.9)	0.002*	0
Oxygen Saturation (median (IQR))	95 (6)	93 (8)	<0.001**	6.3
Respiratory Rate (median (IQR))	22 (8)	26 (12)	<0.001**	6.3
Glasgow Coma Scale (median (IQR))	15 (0)	15 (0)	0.043*	6.6
Systolic BP (median (IQR))	134 (32)	130 (30)	0.009*	15.8
Heart Rate (median (IQR))	96 (27)	94 (27)	0.092	6.4
Temperature (median (IQR))	37.1 (1.4)	37.7 (1.4)	<0.001**	6.7
Chest X-ray report			<0.001**	0
Alternative pathology (%)	4 (0.9)	3 (0.4)		
No abnormalities (%)	178 (40.9)	136 (17.8)		
Indeterminate (%)	83 (19.1)	169 (22.1)		
Classic COVID-19 (%)	170 (39.1)	455 (59.6)		
Presence of comorbidities (%)	297 (79.0)	482 (80.3)	0.669	18.5
Dyspnoea (%)	274 (69.4)	497 (75.5)	0.034	12.1
Neutrophils (median (IQR))	6.42 (4.56)	5.25 (3.92)	<0.001**	2.3
D-Dimer (median (IQR))	1250 (2440)	1105 (1803)	0.204	23.2
Albumin (median (IQR))	39 (7)	37 (6)	<0.001**	10
C-Reactive Protein (median (IQR))	91.0 (115)	146.5 (264.8)	<0.001**	3
Creatine Kinase (median (IQR))	51 (104)	145 (260)	<0.001**	23.3
Troponin (median (IQR))	19 (46)	20 (44)	0.278	19.1
Admitted (%)	331 (76.0)	635 (83.2)	0.003*	0.1
Admitted to ITU (%)	5 (1.3)	32 (4.8)	0.005*	12.4
Thirty Day Follow Up Status			<0.001**	24
Discharged (%)	219 (78.2)	367 (58.3)		
On Ambulatory Follow Up (%)	14 (5.0)	49 (7.8)		
Admitted (%)	18 (6.4)	60 (9.5)		
Died (%)	29 (10.4)	154 (24.4)		
CT report			<0.001**	0
No pathology identified (%)	23 (22.1)	6 (3.3)		
Classic COVID-19 findings (%)	52 (50.0)	157 (85.8)		
Indeterminate for COVID-19 (%)	14 (13.5)	14 (7.7)		
Alternative pathology identified (%)	15 (14.4)	6 (3.3)		
Day of Symptoms (mean (SD))	9.84 (9.63)	8.56 (15.80)	0.368	69.2

Table 2- Baseline characteristics of dataset stratified by overall SARS-CoV 2 RT-PCR status, including subsequent swabs during the study period- NB there were 480 additional swabs on 399 unique patients with a median of 2 and mean of 3.5 per patient; *significant at $p < 0.05$; **significant at $p < 0.001$

	Chest X-ray	CT Chest	Mean Difference	p-value
Total (n)	860	302		
True Positives (n)	305	162	-	-
False Positives (n)	125	55	-	-
True Negatives (n)	187	56	-	-
False Negatives (n)	243	29	-	-
Apparent prevalence (95% CI)	0.50 (0.47-0.53)	0.72 (0.66-0.77)	0.22 (0.04-0.21)	<0.0001**
True prevalence (95% CI)	0.64 (0.60-0.67)	0.63 (0.58-0.69)	-0.00 (-0.09-0.03)	0.111
Sensitivity (95% CI)	0.56 (0.51-0.60)	0.85 (0.79-0.90)	0.29 (0.19-0.38)	<0.0001**
Specificity (95% CI)	0.60 (0.54-0.65)	0.50 (0.41-0.60)	-0.10 (-0.25-0.04)	0.119
Positive Predictive Value (95% CI)	0.71 (0.66-0.75)	0.75 (0.68-0.80)	0.04 (-0.06-0.14)	0.492
Negative Predictive Value (95% CI)	0.43 (0.39-0.48)	0.66 (0.55-0.76)	0.22 (0.06-0.37)	0.005*
Positive Likelihood Ratio (95% CI)	1.39 (1.19-1.62)	1.71 (1.41- 2.08)	0.32 (-0.22-0.89)	0.258
Negative Likelihood Ratio (95% CI)	0.74 (0.64-0.84)	0.30 (0.21-0.44)	-0.44 (-0.64 - -0.21)	0.022*
Diagnostic Accuracy (95% CI)	0.57 (0.54-0.61)	0.72 (0.66-0.77)	0.15 (0.06-0.23)	<0.0001**

Table 3- Diagnostic Accuracy Metrics for CXR and CT Chest with RT-PCR for SARS-CoV 2, as the reference standard; *significant difference at the <0.05 level; **significant difference at the <0.0001 level

		SARS-CoV 2 RT-PCR		OR (univariable)	OR (multivariable)	
		Negative	Positive			
1	n	312	548			
2	Chest X-ray report	Alternative pathology (%)	3 (0.8)	3 (0.5)	-	-
3		No abnormalities (%)	123 (39.6)	104 (19.1)	0.76 (0.08-6.82, p=0.801)	0.48 (0.03-8.82, p=0.620)
4		Indeterminate/ atypical findings (%)	61 (19.5)	136 (4.8)	1.99 (0.22-17.81, p=0.535)	0.92 (0.05-16.88, p=0.952)
5		Classic COVID (%)	125 (40.1)	305 (55.6)	2.17 (0.24-19.19, p=0.484)	1.14 (0.06-20.98, p=0.927)
6	Age	Mean (SD)	61.8 (17.9)	67.0 (17.7)	1.02 (1.01-1.02, p<0.001)**	1.02 (1.00-1.03, p=0.028)*
7	Sex	Female (%)	138 (44.3)	212 (38.7)	-	-
8		Male (%)	174 (55.7)	336 (61.3)	1.26 (0.93-1.70, p=0.137)	1.19 (0.83-1.71, p=0.340)
9	Ethnicity	Other Asian (%)	31 (9.9)	66 (12.0)	-	-
10		White (%)	164 (52.7)	270 (49.2)	0.76 (0.44-1.31, p=0.326)	0.73 (0.38-1.40, p=0.339)
11		Black (%)	39 (12.4)	84 (15.3)	1.01 (0.52-1.98, p=0.974)	0.92 (0.43-1.97, p=0.827)
12		Mixed (%)	6 (1.8)	4 (0.8)	0.36 (0.08-1.62, p=0.184)	0.74 (0.11-4.94, p=0.754)
13		South Asian (%)	22 (7.0)	36 (6.6)	0.77 (0.34-1.76, p=0.531)	0.68 (0.28-1.65, p=0.390)
14		Other (%)	51 (16.2)	89 (16.2)	0.82 (0.43-1.55, p=0.535)	0.88 (0.45-1.74, p=0.716)
15	Comorbidity	No (%)	65 (20.8)	95 (17.4)	-	-
16		Yes (%)	247 (79.2)	453 (82.6)	1.25 (0.82-1.89, p=0.296)	1.00 (0.53-1.88, p=0.993)
17	Dyspnoea on attendance	No (%)	90 (28.8)	139 (25.4)	-	-
18		Yes (%)	222 (71.2)	409 (74.6)	1.19 (0.82-1.73, p=0.356)	0.84 (0.53-1.32, p=0.447)
19	Oxygen Saturation	Median (IQR)	96 (6)	93 (8)	0.94 (0.91-0.97, p<0.001)**	0.97 (0.93-1.00, p=0.072)
20	Respiratory rate	Median (IQR)	23 (8)	25 (8)	1.04 (1.01-1.07, p=0.002)*	1.01 (0.98-1.05, p=0.462)
21	Glasgow Coma Scale	Median (IQR)	15 (0)	15 (0)	1.02 (0.89-1.17, p=0.819)	1.21 (0.98-1.48, p=0.073)
22	Temperature	Mean (SD)	37.2 (1.4)	37.7 (1.1)	1.48 (1.26-1.73, p<0.001)**	1.44 (1.20-1.74, p<0.001)**
23	Heart Rate	Mean (SD)	96.7 (20.5)	94.9 (21.5)	1.00 (0.99-1.00, p=0.305)	1.00 (0.99-1.01, p=0.702)
24	Systolic Blood Pressure	Mean (SD)	136.2 (25.8)	132.6 (24.5)	0.99 (0.99-1.00, p=0.086)	0.99 (0.98-1.00, p=0.097)
25	Neutrophils	Median (IQR)	6.26 (4.52)	5.05 (3.93)	0.92 (0.89-0.96, p<0.001)**	0.87 (0.82-0.91, p<0.001)**
26	D-Dimer	Median (IQR)	1220 (2343)	1061 (1814)	1.00 (1.00-1.00, p=0.403)	1.00 (1.00-1.00, p=0.419)
27	C-Reactive Protein	Median (IQR)	45 (100)	77 (107)	1.00 (1.00-1.01, p<0.001)**	1.00 (1.00-1.01, p=0.021)*
28	Troponin	Median (IQR)	20 (55)	21 (46)	1.00 (1.00-1.00, p=0.890)	1.00 (1.00-1.00, p=0.667)
29	Albumin	Median (IQR)	39 (7)	37 (6)	0.97 (0.94-1.00, p=0.071)	1.02 (0.98-1.06, p=0.432)
30	Creatine Kinase	Median (IQR)	94 (131)	145 (263)	1.00 (1.00-1.00, p=0.119)	1.00 (1.00-1.00, p=0.152)
31	Admitted from ED	Admitted (%)	235 (75.2)	453 (82.7)	-	-
32		Discharged (%)	77 (24.8)	95 (17.3)	1.56 (1.06 -2.33, p=0.022)**	1.35 (0.79-2.30, p=0.272)
33	Admitted To ITU from ED	No (%)	307 (98.5)	532 (97.1)	-	-
34		Yes (%)	5 (1.5)	16 (2.9)	1.92 (0.60-6.18, p=0.274)	1.06 (0.25-4.40, p=0.940)
35	Thirty Day Follow up Status	Discharged (%)	259 (83.0)	368 (67.1)	-	-
36		Admitted (%)	22 (6.9)	47 (8.5)	1.53 (0.82-2.87, p=0.181)	1.64 (0.77-3.51, p=0.198)
37		Dead (%)	31 (10.1)	133 (24.4)	3.00 (1.86-4.84, p<0.001)**	2.81 (1.22-6.50, p=0.017)*

Table 4- Association of covariates with RT-PCR status for SARS-CoV 2, following propensity score matching and binomial logistic regression; SD- Standard deviation; IQR- Interquartile Range; *p<0.05; **p<0.001

		X-ray report		OR (univariable)	OR with XR as binary outcome (multivariable)	OR with XR as ordinal variable (multivariable)	
		Other X-ray Findings	Classical COVID-19				
6	n	430	430				
7	RT-PCR for SARS-CoV 2	Negative (%)	187 (43.4)	125 (29.1)	-	-	-
8		Positive (%)	243 (56.6)	305 (70.9)	1.85 (1.36-2.56, p<0.001)**	1.79 (1.25-2.56, p<0.002)*	1.94 (1.37-2.76, p<0.001)**
9	Age	Mean (SD)	65.0 (18.9)	65.3 (16.9)	1.00 (0.99-1.01, p=0.849)	0.99 (0.98-1.00, p=0.164)	1.00 (0.99-1.01, p=0.542)
10	Sex	Female (%)	176 (40.9)	175 (40.6)	-	-	-
11		Male (%)	254 (59.1)	255 (59.3)	1.01 (0.75-1.37, p=0.940)	0.87 (0.63-1.20, p=0.400)	1.02 (0.49-2.09, p=0.967)
12	Ethnicity	Other Asian (%)	49 (11.4)	48 (11.2)	-	-	-
13		South Asian (%)	29 (6.7)	29 (6.7)	1.04 (0.52-2.04, p=0.912)	1.02 (0.47-2.17, p=0.965)	1.02 (0.49-2.09, p=0.967)
14		Black (%)	61 (14.2)	61 (14.2)	1.02 (0.55-1.85, p=0.957)	0.88 (0.46-1.69, p=0.719)	0.92 (0.52-1.65, p=0.789)
15		Mixed (%)	5 (1.2)	5 (1.2)	0.92 (0.21-4.00, p=0.911)	0.86 (0.18-4.17, p=0.853)	0.85 (0.17-4.30, p=0.838)
16		Other (%)	70 (16.3)	70 (16.3)	1.02 (0.58-1.79, p=0.943)	0.98 (0.52-1.82, p=0.942)	0.93 (0.53-1.64, p=0.810)
17	Comorbidity	White (%)	216 (50.2)	217 (50.5)	1.03 (0.63-1.67, p=0.913)	0.97 (0.57-1.67, p=0.926)	0.90 (0.55-1.47, p=0.666)
18		No (%)	82 (19.1)	78 (18.1)	-	-	-
19		Yes (%)	348 (80.9)	352 (81.9)	0.95 (0.66-1.36, p=0.777)	0.93 (0.59-1.49, p=0.782)	0.88 (0.57-1.37, p=0.592)
20	Dyspnoea	No (%)	191 (29.3)	103 (24.0)	-	-	-
21		Yes (%)	304 (70.7)	327 (76.0)	1.31 (0.92-1.88, p=0.123)	1.20 (0.80-1.82, p=0.380)	1.22 (0.83-1.80, p=0.301)
22	Oxygen Saturation	Median (IQR)	95 (7)	93 (7)	0.94 (0.91-0.96, p<0.001)**	0.94 (0.92-0.97, p<0.001)**	0.94 (0.91-0.97, p<0.001)**
23	Respiratory rate	Median (IQR)	24 (10)	24 (10)	1.01 (0.99-1.02, p=0.570)	0.97 (0.94-1.00, p=0.063)	0.98 (0.96-1.01, p=0.157)
24	Glasgow Coma Scale	Median (IQR)	15 (0)	15 (0)	1.04 (0.92-1.19, p=0.524)	1.05 (0.90-1.23, p=0.503)	1.05 (0.92-1.21, p=0.464)
25	Temperature	Mean (SD)	37.6 (1.1)	37.5 (1.3)	0.93 (0.83-1.06, p=0.297)	0.79 (0.67-0.93, p=0.006)*	0.85 (0.73-0.99, p=0.031)*
26	Heart Rate	Mean (SD)	95.7 (21.4)	95.5 (21.0)	1.00 (0.99-1.01, p=0.888)	1.00 (0.99-1.01, p=0.864)	1.00 (0.99-1.01, p=0.872)
27	Systolic Blood Pressure	Mean (SD)	133.8 (25.0)	134.0 (25.6)	1.00 (0.99-1.01, p=0.907)	1.00 (0.99-1.01, p=0.335)	1.00 (1.00-1.01, p=0.478)
28	Neutrophils	Median (IQR)	5.44 (4.54)	5.67 (4.03)	1.00 (0.97-1.04, p=0.892)	0.96 (0.92-1.01, p=0.143)	0.96 (0.92-1.01, p=0.115)
29	D-Dimer	Median (IQR)	1119 (2221)	1119 (1850)	1.00 (1.00-1.00, p=0.513)	1.00 (1.00-1.00, p=0.568)	1.00 (1.00-1.00, p=0.385)
30	C-Reactive Protein	Median (IQR)	46 (93)	88 (110)	1.00 (0.99-1.00, p<0.001)**	1.00 (1.00-1.01, p<0.001)**	1.00 (1.00-1.01, p<0.001)**
31	Troponin	Median (IQR)	23 (54)	20 (46)	1.00 (1.00-1.00, p=0.231)	1.00 (1.00-1.00, p=0.277)	1.00 (1.00-1.00, p=0.059)
32	Albumin	Median (IQR)	39 (7)	37 (6)	0.93 (0.90-0.96, p<0.001)**	0.93 (0.90-0.97, p=0.001)*	0.94 (0.91-0.97, p=0.001)*
33	Creatine Kinase	Median (IQR)	110 (183)	134 (239)	1.00 (1.00-1.00, p=0.535)	1.00 (1.00-1.00, p=0.242)	1.00 (1.00-1.00, p=0.186)
34	Admitted from ED	Admitted (%)	315 (73.3)	373 (86.7)	2.37 (1.63-3.46, p<0.001)**	2.30 (1.46-3.63, p<0.001)**	2.22 (1.47-3.33, p<0.001)**
35		Discharged (%)	115 (26.7)	57 (13.3)	-	-	-
36	Admitted to ITU from ED	No (%)	423 (98.4)	416 (96.7)	-	-	-
37		Yes (%)	7 (1.6)	14 (3.3)	2.17 (0.69-6.67, p=0.181)	1.27 (0.32-5.00, p=0.732)	1.34 (0.36-5.00, p=0.653)
38	30 Day Follow Up Status	Discharged (%)	316 (73.5)	311 (72.3)	-	-	-
39		Admitted (%)	34 (7.9)	34 (7.9)	1.31 (0.81-2.13, p=0.282)	1.32 (0.69-2.53, p=0.392)	1.43 (0.78-2.63, p=0.653)
40		Dead (%)	80 (18.6)	85 (19.8)	1.03 (0.73-1.45, p=0.886)	1.38 (0.80-2.37, p=0.247)	1.41 (0.87-2.27, p=0.157)

Table 5- Association of covariates with CXR report following propensity score matching and either binomial or ordinal logistic regression; SD- Standard deviation; IQR- Interquartile Range; *p<0.05; **p<0.001

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

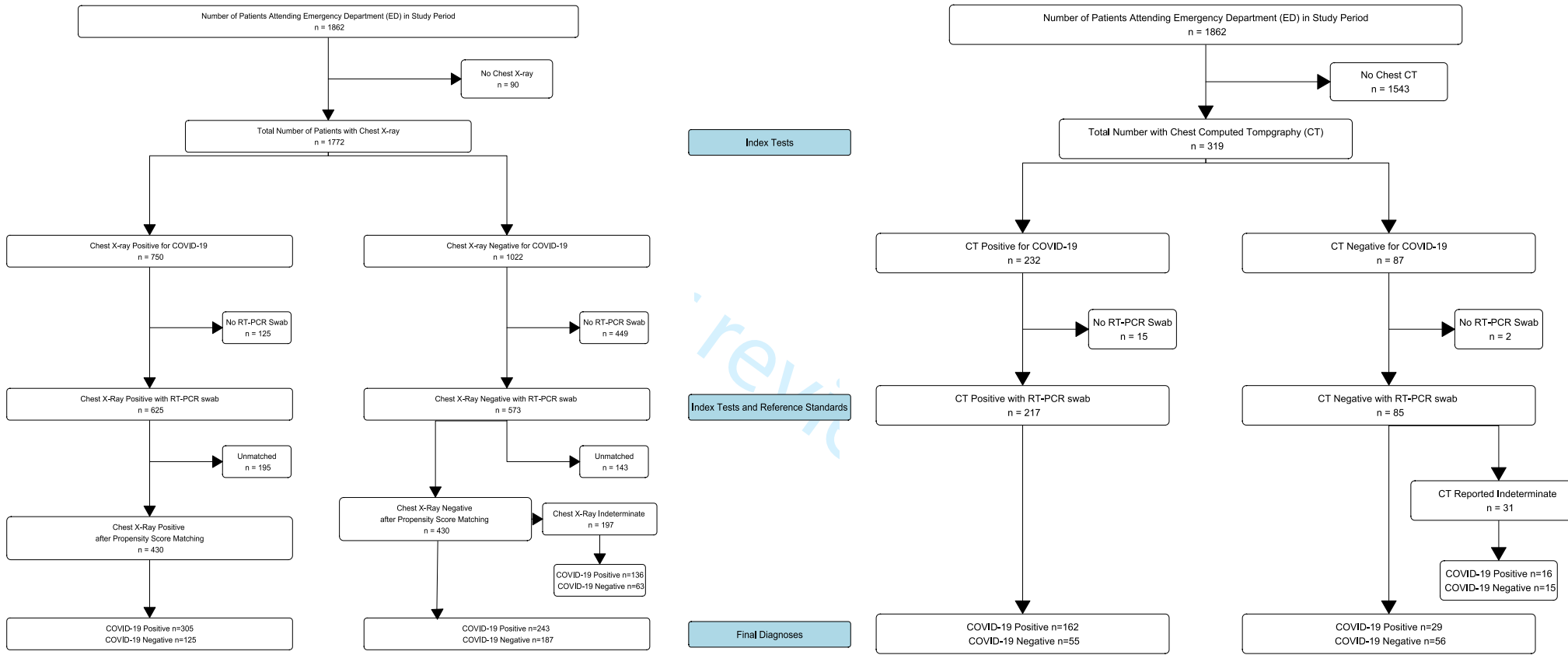


Figure 1- Inclusion and exclusion of patients during study period with test results

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Odds Ratio of Positivity for SARS-CoV 2 by RT-PCR

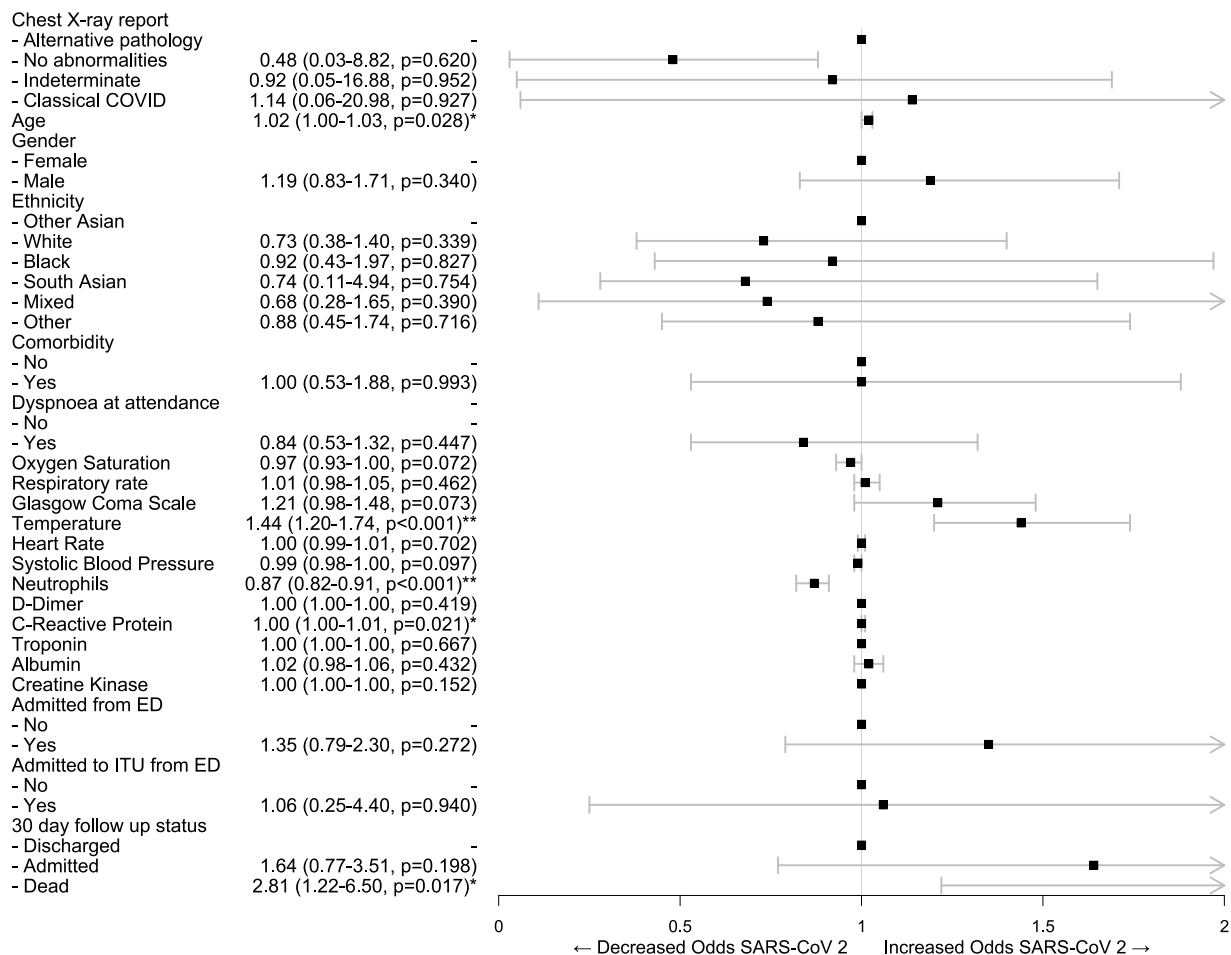


Figure 2- Forest plot of odds ratios of variables associated with RT-PCR positivity for SARS-CoV 2, following multiple imputation, propensity score matching and binomial logistic regression; *significant difference at the <0.05 level; **significant difference at the <0.001 level

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Odds Ratio of Classical COVID-19 Findings on Chest X-Ray

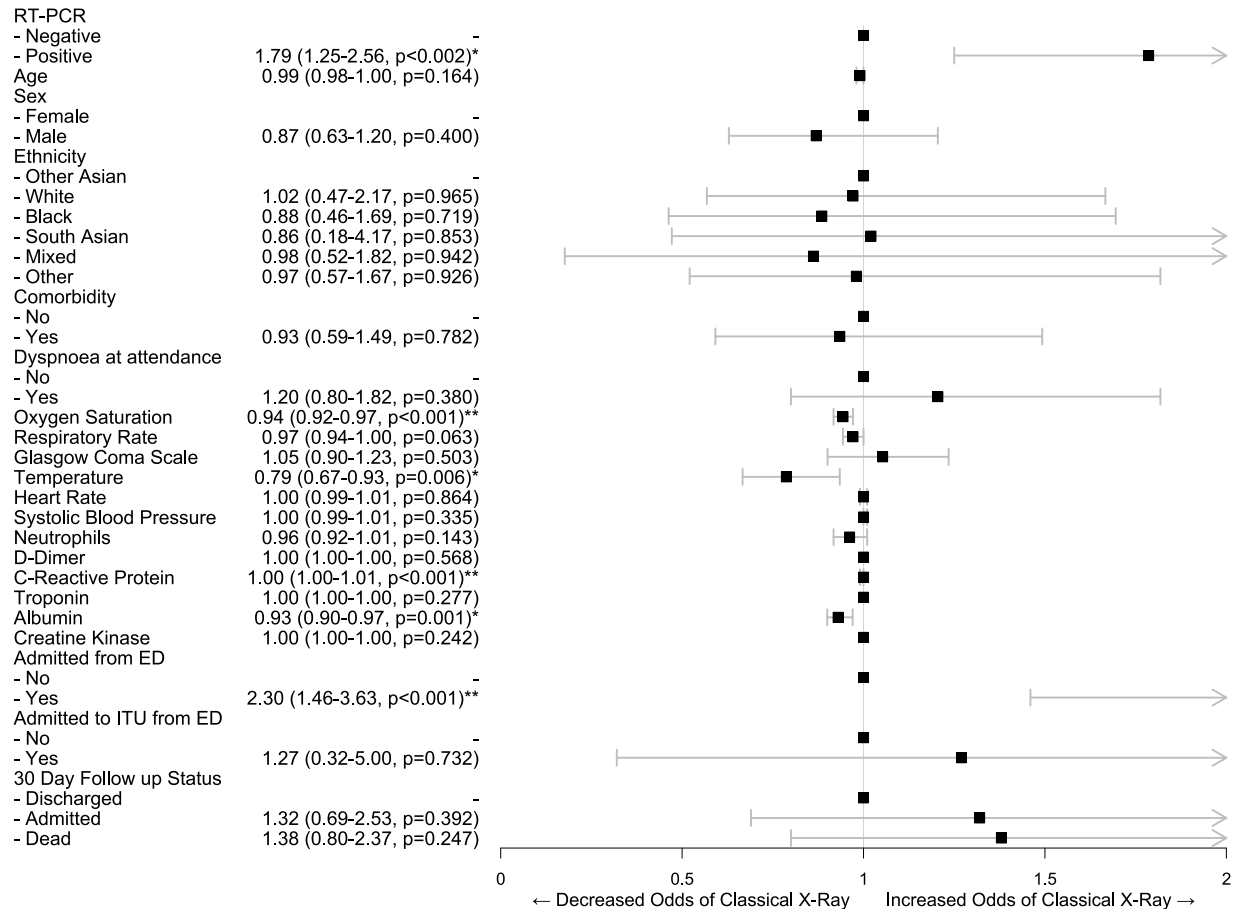
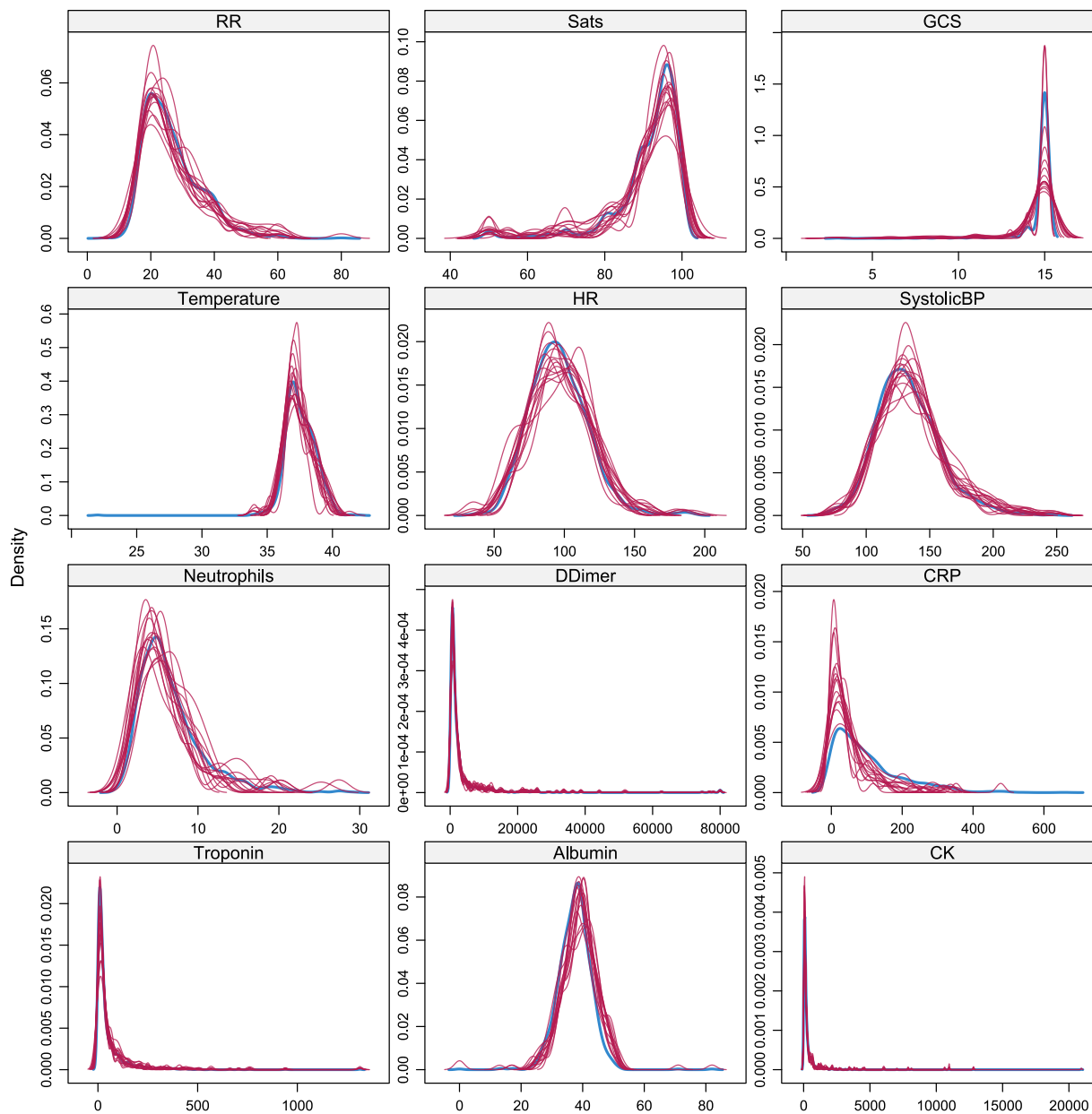


Figure 3- Forest plot of odds ratios of variables associated with classical Chest X-ray features COVID-19 following propensity score matching and binomial logistic regression; *significant difference at the <0.05 level; **significant difference at the <0.001 level

Supplementary file 1



Supplementary figure 1- Density plots of imputed datasets; Blue represents original dataset; other colours are individual imputed datasets (n=15)

Covariate:	Means Treated	Means Control	Standard Deviation Control	Mean Difference
Overall Propensity Score	0.422997940	0.53935303	0.1449627	-0.1163550897
Female	36.3782051	45.026178	0.4979547	-8.64797288
Male	63.6217949	54.973822	0.4979547	8.64797288
Age	63.796474359	66.19022688	18.5893357	-23.937525171
Comorbidity- Yes	76.1217949	84.467714	0.3625287	-8.34591892
Ethnicity- South Asian	6.5705128	6.631763	0.2490539	-0.06124983
Ethnicity- Black	16.1858974	11.518325	0.3195219	4.66757283
Ethnicity- Mixed	0.9615385	1.396161	0.1174340	-0.43462210
Ethnicity- Other	18.9102564	13.263525	0.3394765	5.64673110
Ethnicity- White	46.6346154	57.766143	0.4943635	-11.13152772
Respiratory Rate	29.214743590	24.01745201	7.2639816	5.1972915828

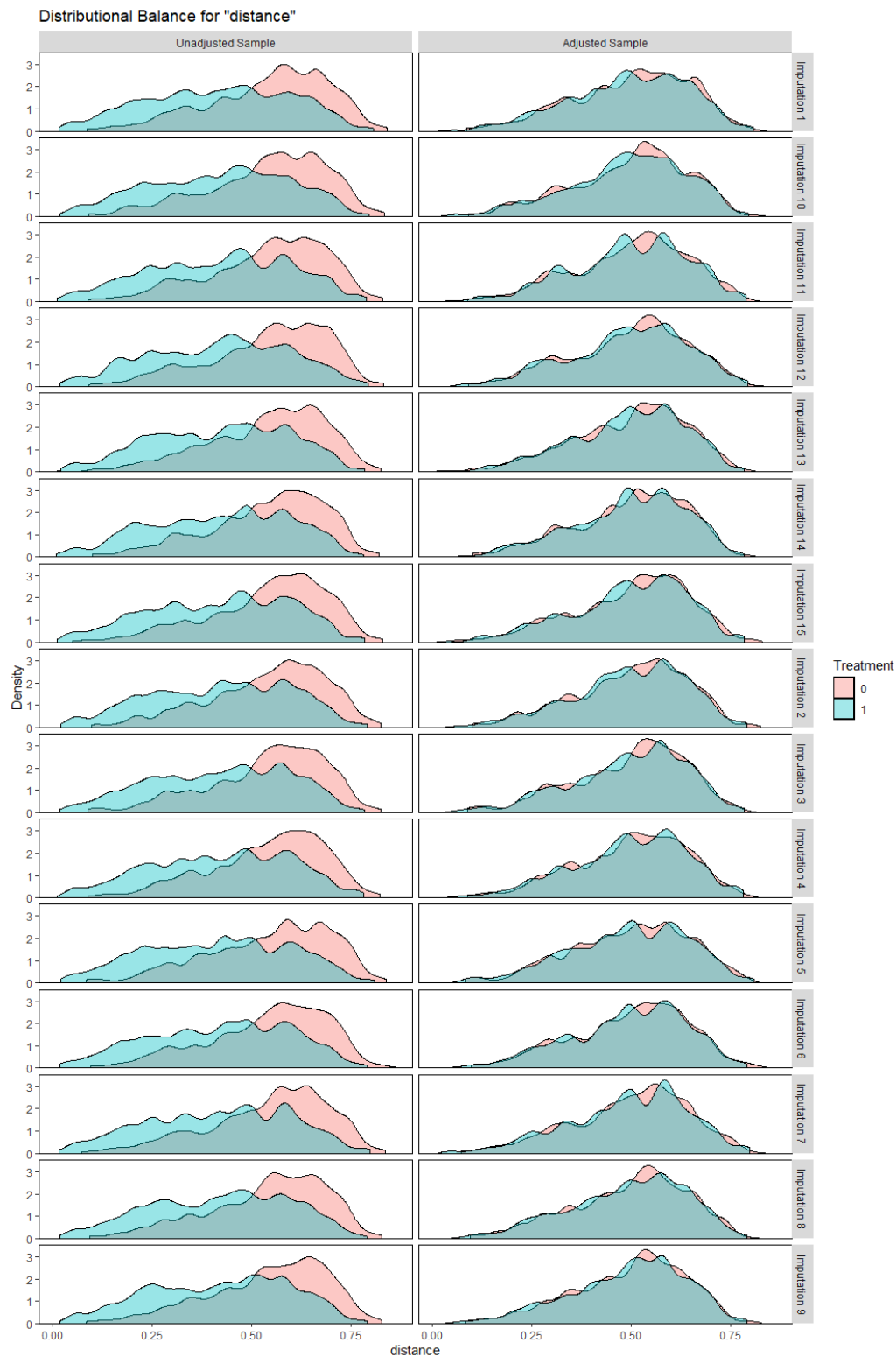
Supplementary table 1- Means of data before multiple imputation and propensity score matching

	Type	Minimum Difference Adjusted	Mean Difference Adjusted	Maximum Difference Adjusted
Distance	Distance	0.016988	0.027107	0.040963
Sex = Male	Binary	-0.03917	-0.0028	0.015982
Age	Contin.	-0.04586	-0.01371	0.027589
Comorbidity = Yes	Binary	-0.02331	-0.00778	0.004598
Ethnicity = Other Asian	Binary	-0.01392	0.002362	0.016471
Ethnicity = South Asian	Binary	-0.01399	-0.00136	0.011905
Ethnicity = Black	Binary	-0.01852	0.000443	0.015982
Ethnicity = Mixed	Binary	-0.00464	0.001403	0.007042
Ethnicity = Other	Binary	-0.01152	4.30E-06	0.00939
Ethnicity = White	Binary	-0.02353	-0.00285	0.018433
Respiratory Rate	Contin.	-0.06157	-0.03478	-0.00442

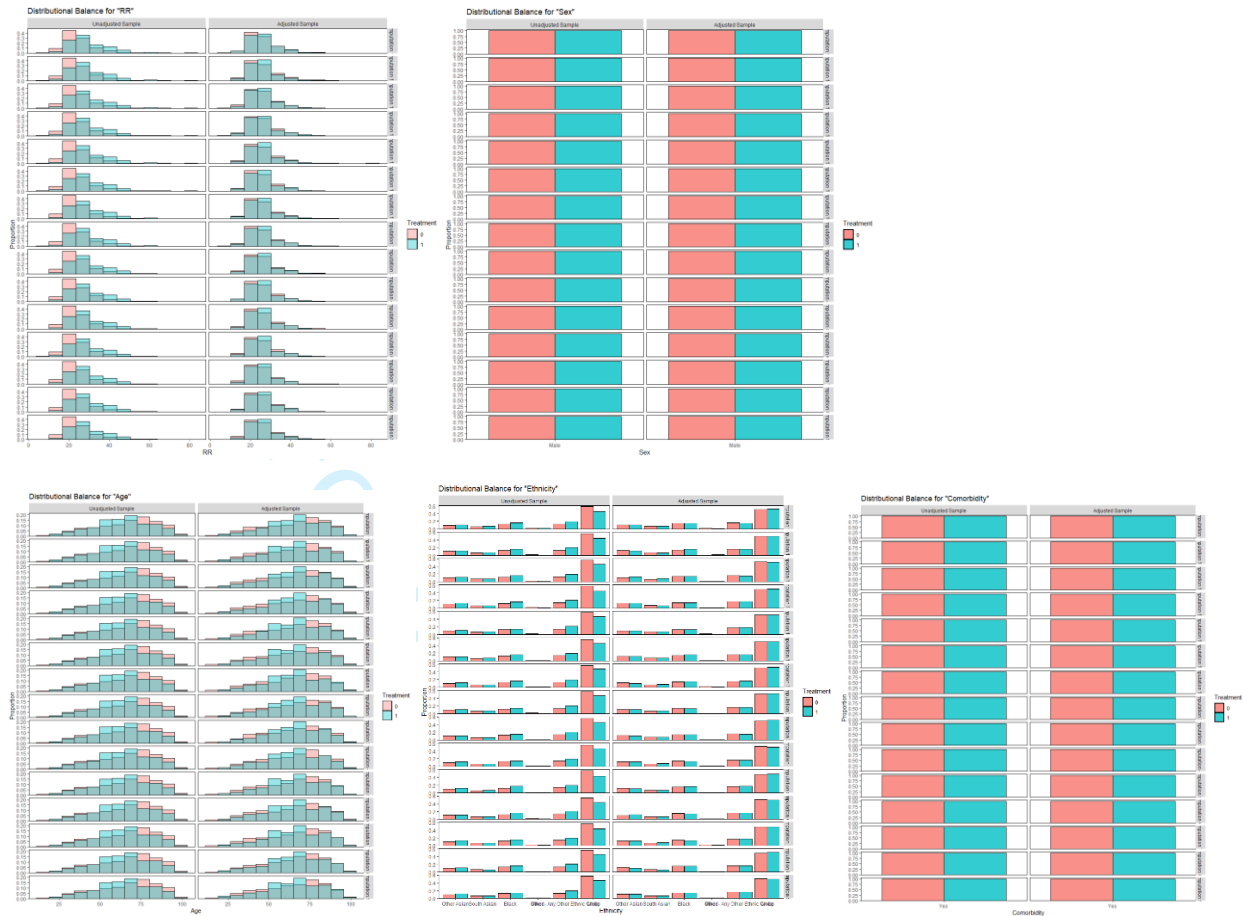
Supplementary table 2- Balance summary across imputations

	XR- Negative	XR- Positive	Total
All	573	625	1,198
Matched	430	430	860
Unmatched	143	195	338
Discarded	0	0	0

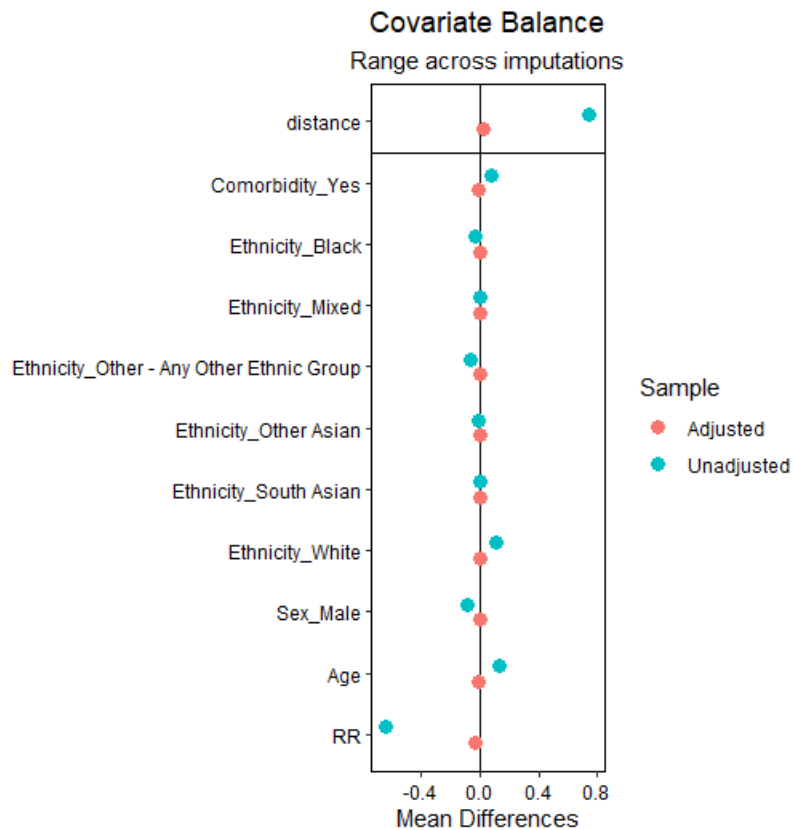
Supplementary table 3- Average Sample sizes pre- and post- matching across imputed data sets



Supplementary figure 2- Density plot of propensity scores pre- and post- matching in each imputed dataset; treatment units represent a positive X-ray for COVID-19, whereas a control unit represents a negative X-ray



Supplementary figure 3- Histogram of distributions for each matching covariate pre- and post- matching in each imputed dataset; treatment units represent a positive X-ray for COVID-19, whereas a control unit represents a negative X-ray



Supplementary figure 4- Love plot of pooled balances across imputed datasets in matching covariates after matching

CXR in COVID Analysis

Dr Aditya Borakati

Royal Free Hospital, Pond Street, London, NW3 2QG
a.borakati@doctors.org.uk

2020-10-06

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1	
2	
3	
4	
5	
6	1 Software Environment and Packages 7
7	1.1 Load Packages and Data 9
8	1.1.1 Load Packages: 9
9	1.2 Power Calculation 9
10	2 Load Data: 11
11	3 Data Cleaning 13
12	3.0.1 Follow Up Swabs + Initial Swabs Positive: . 14
13	3.0.2 Paired XR and RT-PCR data 14
14	4 Demographic table of raw data 17
15	5 Imputation 21
16	6 Propensity Score Matching 23
17	6.1 Match Balance Diagnostics 23
18	7 Matched Demographics Table: 25
19	8 Diagnostic Accuracy 27
20	8.1 CT Data and Accuracy 28
21	8.2 CT and XR accuracy comparison 30
22	8.2.1 Sensitivity 30
23	8.3 Intermodality Agreement 34
24	8.3.1 Diagnostic Accuracy Analysis when
25	Indeterminate Reports of CXR and CT are taken
26	as positive 35
27	9 Pooled Regression after Multiple Imputation and
28	Propensity Score Matching 39
29	9.0.1 Pooled Univariate Odds Ratios for OverallPos
30	as dependent variable 39
31	9.0.2 Binomial Logistic Regression with Positive
32	Chest X-ray Report as Dependent Variable 40
33	9.0.3 Univariate XRPositive as dependent 40
34	9.0.4 Multivariate XRPositive as dependent 40
35	9.0.5 Pooled Ordinal Logistic Regression with
36	XRPositive as dependent 41
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	

iv

9.1 Forest Plots	41
9.2 Correlation Matrix	43
9.3 STARD Flow Diagram	46

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1 Software Environment and Packages

```

R version 4.0.0 (2020-04-24)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 19041)
Matrix products: default
locale:
 LC_COLLATE=English_United Kingdom.1252 LC_CTYPE=English_United Kingdom.1252
 LC_MONETARY=English_United Kingdom.1252 LC_NUMERIC=C
 LC_TIME=English_United Kingdom.1252
attached base packages:
 stats graphics grDevices utils datasets methods base
other attached packages:
 corrplot 0.84
   Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of
   a Correlation Matrix (Version 0.84). Available from
   https://github.com/taiyun/corrplot
MKmisc 1.6
   Kohl M (2019). MKmisc: Miscellaneous functions from M. Kohl. R package
   version 1.6, http://www.stamats.de
epiR 1.0-14
   Mark Stevenson with contributions from Telmo Nunes, Cord Heuer, Jonathon
   Marshall, Javier Sanchez, Ron Thornton, Jenő Reiczigel, Jim Robison-Cox,
   Paola Sebastiani, Peter Solymos, Kazuki Yoshida, Geoff Jones, Sarah
   Pirikahu, Simon Firestone, Ryan Kyle, Johann Popp, Mathew Jay and Charles
   Reynard. (2020). epiR: Tools for the Analysis of Epidemiological Data. R
   package version 1.0-14. https://CRAN.R-project.org/package=epiR
Matching 4.9-7
   Jasjeet S. Sekhon (2011). Multivariate and Propensity Score Matching
   Software with Automated Balance Optimization: The Matching Package for R.
   Journal of Statistical Software, 42(7), 1-52. URL
   http://www.jstatsoft.org/v42/i07/.
MASS 7.3-51.5
   Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S.
   Fourth Edition. Springer, New York. ISBN 0-387-95457-0
Ordinal 2019.12-10
   Christensen, R. H. B. (2019). ordinal - Regression Models for Ordinal Data. R
   package version 2019.12-10. https://CRAN.R-
   project.org/package=ordinal.
Hmisc 4.4-0
   Frank E Harrell Jr, with contributions from Charles Dupont and many
   others. (2020). Hmisc: Harrell Miscellaneous. R package version 4.4-0.
   https://CRAN.R-project.org/package=Hmisc
Formula 1.2-3
   Achim Zeileis, Yves Croissant (2010). Extended Model Formulas in R:
   Multiple Parts and Multiple Responses. Journal of Statistical Software
   34(1), 1-13. doi:10.18637/jss.v034.i01
lattice 0.20-41
   Sarkar, Deepayan (2008) Lattice: Multivariate Data Visualization with R.
   Springer, New York. ISBN 978-0-387-75968-5

```

8 1 Software Environment and P...

```
mice 3.8.0
  Stef van Buuren, Karin Groothuis-Oudshoorn (2011). mice: Multivariate
  Imputation by Chained Equations in R. Journal of Statistical Software,
  45(3), 1-67. URL https://www.jstatsoft.org/v45/i03/.
readxl 1.3.1
  Hadley Wickham and Jennifer Bryan (2019). readxl: Read Excel Files. R
  package version 1.3.1. https://CRAN.R-project.org/package=readxl
finalfit 1.0.1
  Ewen Harrison, Tom Drake and Riinu Ots (2020). finalfit: Quickly Create
  Elegant Regression Results Tables and Plots when Modelling. R package
  version 1.0.1. https://CRAN.R-project.org/package=finalfit
MatchIt 3.0.2
  Daniel E. Ho, Kosuke Imai, Gary King, Elizabeth A. Stuart (2011). MatchIt:
  Nonparametric Preprocessing for Parametric Causal Inference. Journal of
  Statistical Software, Vol. 42, No. 8, pp. 1-28. URL
  http://www.jstatsoft.org/v42/i08/
tableone 0.11.1
  Kazuki Yoshida (2020). tableone: Create 'Table 1' to Describe Baseline
  Characteristics. R package version 0.11.1.
  https://CRAN.R-project.org/package=tableone
forcats 0.5.0
  Hadley Wickham (2020). forcats: Tools for Working with Categorical
  Variables (Factors). R package version 0.5.0.
  https://CRAN.R-project.org/package=forcats
stringr 1.4.0
  Hadley Wickham (2019). stringr: Simple, Consistent Wrappers for Common
  String Operations. R package version 1.4.0.
  https://CRAN.R-project.org/package=stringr
dplyr 0.8.5
  Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2020).
  dplyr: A Grammar of Data Manipulation. R package version 0.8.5.
  https://CRAN.R-project.org/package=dplyr
purrr 0.3.4
  Lionel Henry and Hadley Wickham (2020). purrr: Functional Programming
  Tools. R package version 0.3.4. https://CRAN.R-project.org/package=purrr
readr 1.3.1
  Hadley Wickham, Jim Hester and Romain François (2018). readr: Read
  Rectangular Text Data. R package version 1.3.1.
  https://CRAN.R-project.org/package=readr
tidyr 1.0.2
  Hadley Wickham and Lionel Henry (2020). tidyr: Tidy Messy Data. R package
  version 1.0.2. https://CRAN.R-project.org/package=tidyr
tibble 3.0.0
  Hadley Wickham and Lionel Henry (2020). tidyr: Tidy Messy Data. R package
  version 1.0.2. https://CRAN.R-project.org/package=tidyr
ggplot2 3.3.0
  H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag
  New York, 2016.
tidyverse 1.3.0
  Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source
  Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
forestplot 1.9
  Max Gordon and Thomas Lumley (2019). forestplot: Advanced Forest Plot Using
  'grid' Graphics. R package version 1.9. https://CRAN.R-project.org/package=forestplot
MatchThem 0.9.3
  Farhad Pishgar and Noah Greifer (2020). MatchThem: Matching and Weighting
  Multiply Imputed Datasets. R package version 0.9.3. https://CRAN.R-project.org/package=MatchThem
```

1.1 Load Packages and Data

9

```
miceadds 3.9-14
  Robitzsch, A., & Grund, S. (2020). miceadds: Some Additional Multiple
  Imputation Functions, Especially for 'mice'. R package version 3.9-14.
  https://CRAN.R-project.org/package=miceadds
cobalt 4.2.2
  Noah Greifer (2020). cobalt: Covariate Balance Tables and Plots. R package
  version 4.2.2. https://CRAN.R-project.org/package=cobalt
```

1.1 Load Packages and Data

1.1.1 Load Packages:

```
library(MKmisc)
library(tidyverse)
library(tableone)
library(MatchIt)
library(finalfit)
library(readxl)
library(cobalt)
library(mice)
library(miceadds)
library(Hmisc)
library(epiR)
library(MatchThem)
library(ordinal)
library(forestplot)
```

1.2 Power Calculation

1.2.0.0.1 This code calculates the sample size (positive and negative by gold standard test) needed to evaluate a diagnostic test with 56% sensitivity at 80% power with alpha 0.05. The 56% value is the lower confidence reported by Wong et al. and lower sensitivities typically require higher sample sizes, the result is the same whether specificity or sensitivities are passed as arguments, the previously published specificities are higher than sensitivities so for a generous estimate, the sensitivity was used.

```
power <- power.diagnostic.test(sens = 0.56,
  sig.level = 0.05, delta = 0.1, power = 0.8) %>%
  print()
```

10

1 Software Environment and P...

```
Diagnostic test exact power calculation
```

```
    sens = 0.56  
      n  = 165  
     n1 = 165  
    delta = 0.1  
sig.level = 0.05  
    power = 0.8  
     prev = NULL
```

```
NOTE: n is number of cases, n1 is number of controls
```

2 Load Data:

```
data <- read_csv("FullDataWithCT.csv", col_types = cols(Age = col_integer(),
  Albumin = col_number(), CK = col_number(),
  CT = col_character(), CRP = col_number(),
  DDimer = col_number(), DateOfDeath = col_date(format = "%d/%m/%Y"),
  DateOfDischarge = col_date(format = "%d/%m/%Y"),
  DateOfVisit = col_date(format = "%d/%m/%Y"),
  DateOfSymptomOnset = col_date(format = "%d/%m/%Y"),
  DiastolicBP = col_number(), FiO2 = col_skip(),
  GCS = col_number(), HR = col_number(),
  MRN = col_skip(), NEWS = col_number(),
  `NEWS2(noFiO2)` = col_skip(), Neutrophils = col_number(),
  RR = col_number(), Sats = col_number(),
  `Supplemental Oxygen` = col_skip(), SystolicBP = col_number(),
  Temperature = col_number(), Troponin = col_number(),
  CTBSTI = col_integer()))
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

3 Data Cleaning

3.0.0.0.1 Format data into factors/ differences between dates:

```
data <- mutate_if(data, is.character, as.factor)
data$DayOfSymptoms <- difftime(data$DateOfVisit,
  data$DateOfSymptomOnset, units = "days")
data$TimeToDeath <- abs(difftime(data$DateOfDeath,
  data$DateOfVisit, units = "days"))
data$DayOfSymptoms <- as.numeric(data$DayOfSymptoms)
data$TimeToDeath <- as.numeric(data$TimeToDeath)
```

3.0.0.1 Recode ethnicities as too many options:

3.0.0.1.0.1 This code collapses the ethnicity categories into 'White', 'Black', 'South Asian', 'Other Asian', 'Mixed' or 'Other';

```
data$Ethnicity <- fct_collapse(data$Ethnicity,
  White = c("White - British", "White - Irish",
    "White - Any Other White Background"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  Black = c("Black - Any Other Black Background",
    "Black or Black British - African",
    "Black or Black British - African",
    "Black or Black British - Caribbean"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  `South Asian` = c("Asian or Asian British - Bangladeshi",
    "Asian or Asian British - Indian",
    "Asian or Asian British - Pakistani"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  `Other Asian` = c("Asian - Any Other Asian Background",
    "Other - Chinese"))
data$Ethnicity <- fct_collapse(data$Ethnicity,
  Mixed = c("mixed - Any Other mixed Background",
    "Mixed - Any Other Mixed Background",
    "Mixed - White and Asian", "Mixed - White and Black African",
    "mixed - White and Black Caribbean",
    "Mixed - White and Black Caribbean"))
```


3.0.0.1.0.2 New XR positive column for “Classic Covid” or not:

```
data$XRPositive <- ifelse(data$XR Chest ==  
  "Classic COVID", "Positive", "Negative")  
data$XRPositive <- as.factor(data$XRPositive)
```

3.0.1 Follow Up Swabs + Initial Swabs Positive:

3.0.1.0.0.1 Creates new column ‘OverallPos’ which includes initial RT-PCR swab and follow-up swabs in 30 days of attendance, if any are positive the value will be positive in this column

```
data$OverallPos <- case_when(data$RTPCR ==  
  "Positive" | data$FollowUpPos == "Positive" ~  
  "Positive")  
data$OverallPos <- replace_na(data$OverallPos,  
  "Negative")
```

3.0.1.0.0.2 Create new vector with all variable names (i.e. the column headers)

```
explanatory <- names(data)
```

3.0.2 Paired XR and RT-PCR data

3.0.2.1 Creates new variable ‘completedata’ which contains only patients who had both CXR and RT-PCR in ED

```
completedata <- filter(data, !is.na(data$XRPositive) &  
  !is.na(data$RTPCR))
```

3.0.2.1.1 Remove missing data variable

```
completedata <- completedata[-c(31)]
```

3.0.2.2 Format complete data variables

```
completedata$OverallPos <- as.factor(completedata$OverallPos)

completedata$ThirtyDayFU <- as.factor(completedata$ThirtyDayFU)
completedata$TimeToDeath <- abs(difftime(completedata$DateOfDeath,
completedata$DateOfVisit, units = "days"))

completedata$TimeToDeath <- as.numeric(completedata$TimeToDeath)
```

3.0.2.2.0.1 Set 'XRchest' as ordinal variable on scale of 'Alternative pathology' as lowest value and 'Classical COVID' as highest

```
completedata$XRchest <- ordered(completedata$XRchest,
levels = c("Alternative pathology", "No abnormalities",
"Indeterminate", "Classic COVID"))
```

3.0.2.2.0.2 Convert CT BSTI grade column into factor:

```
completedata$CTBSTI <- as.factor(completedata$CTBSTI)
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

4 Demographic table of raw data

4.0.0.0.1 This code creates an unformatted demographic table (table 2 in manuscript), for the raw data, stratified by RT-PCR status, significance testing between RT-PCR +ve and -ve groups is carried out automatically using chi squared, t-tests, ANOVA etc.; there is also a column for the proportion of missing data

```

CreateTableOne(vars = explanatory,
               strata = 'OverallPos',
               data = completedata) -> demogtable

#### List nonnormal factors for summarisation as median / IQR and non
      parametric statistical test

explanatoryynormal<-c("Sats", "RR", "GCS", "SystolicBP", "Temperature", "HR",
                    "Neutrophils",
                    "DDimer", "Albumin", "CRP", "CK", "Troponin")
as.data.frame(print(demogtable, nonnormal = explanatoryynormal, missing =
                    TRUE))->demogtable

write.csv(demogtable, file = "Demogtable.csv")

```

Age (mean (SD))	62.74 (17.72)	66.18 (17.58)
0.001		
Ethnicity (%)		
0.097		
Other Asian	29 (8.0)	72 (11.8)
South Asian	27 (7.5)	38 (6.2)
Black	41 (11.4)	91 (14.9)
Mixed	6 (1.7)	6 (1.0)
Other - Any Other Ethnic Group	56 (15.5)	105 (17.2)
White	202 (56.0)	297 (48.8)
Sex = Male (%)	233 (53.6)	480 (62.9)
0.002		
Sats (median [IQR])	95.00 [92.00, 98.00]	93.00 [88.00,
96.00]	<0.001 nonnorm	
RR (median [IQR])	22.00 [20.00, 28.00]	26.00 [20.00,
32.00]	<0.001 nonnorm	
GCS (median [IQR])	15.00 [15.00, 15.00]	15.00 [15.00,
15.00]	0.043 nonnorm	
SystolicBP (median [IQR])	134.00 [119.00, 151.50]	130.00 [115.00,
145.00]	0.009 nonnorm	
DiastolicBP (mean (SD))	79.54 (16.40)	75.61 (14.51)
<0.001		
HR (median [IQR])	96.00 [83.00, 110.00]	94.00 [81.00,
108.00]	0.092 nonnorm	

18

4 Demographic table of raw data

Temperature (median [IQR]) 38.40] <0.001 nonnorm	37.10 [36.60, 38.00]	37.70 [37.00,
XR Chest (%) <0.001		
Alternative pathology	4 (0.9)	3 (0.4)
No abnormalities	178 (40.9)	136 (17.8)
Indeterminate	83 (19.1)	169 (22.1)
Classic COVID	170 (39.1)	455 (59.6)
CTPA = PE (%) 0.127	16 (30.2)	28 (45.9)
Comorbidity = Yes (%) 0.669	297 (79.0)	482 (80.3)
Dyspnoea = Yes (%) 0.034	274 (69.4)	497 (75.5)
Neutrophils (median [IQR]) 7.61] <0.001 nonnorm	6.42 [4.55, 9.11]	5.25 [3.69,
DDimer (median [IQR]) 2428.50] 0.204 nonnorm	1250.00 [619.00, 3059.00]	1105.00 [626.00,
Albumin (median [IQR]) 40.00] <0.001 nonnorm	39.00 [35.00, 42.00]	37.00 [34.00,
CRP (median [IQR]) 158.00] <0.001 nonnorm	51.00 [13.00, 117.00]	83.00 [42.00,
CK (median [IQR]) 342.75] <0.001 nonnorm	91.00 [54.00, 169.00]	146.50 [78.00,
Troponin (median [IQR]) 53.00] 0.278 nonnorm	19.00 [7.00, 53.00]	20.00 [9.00,
Admitted = Discharged (%) 0.003	104 (24.0)	128 (16.8)
AdmittedToITU = Yes (%) 0.005	5 (1.3)	32 (4.8)
RTPCR = Positive (%) <0.001	0 (0.0)	738 (96.7)
CT = 1 (%) 0.011	37 (57.8)	26 (86.7)
NEWS (mean (SD)) 0.032	4.36 (3.06)	5.48 (2.71)
ThirtyDayFU (%) <0.001		
1	219 (78.2)	367 (58.3)
2	14 (5.0)	49 (7.8)
3	18 (6.4)	60 (9.5)
4	29 (10.4)	154 (24.4)
CTBSTI (%) <0.001		
0	23 (22.1)	6 (3.3)
1	52 (50.0)	157 (85.8)
2	14 (13.5)	14 (7.7)
3	15 (14.4)	6 (3.3)
DayOfSymptoms (mean (SD)) 0.368	9.84 (9.63)	8.56 (15.80)
TimeToDeath (mean (SD)) 0.618	50.33 (77.93)	57.76 (70.02)
XRPositive = Positive (%) <0.001	170 (39.1)	455 (59.6)
OverallPos = Positive (%)	0 (0.0)	763 (100.0)

1
2
3
4
5
6 **4.0.0.0.2 Limited dataset comprising relevant data and those without**
7 **significant missingness:**
8

```
9  
10 limcompletedata <- dplyr::select(completedata,  
11   c("Age", "XRchest", "Ethnicity", "Sex",  
12     "RR", "Sats", "GCS", "Temperature",  
13     "HR", "SystolicBP", "DiastolicBP",  
14     "Neutrophils", "DDimer", "CRP", "Troponin",  
15     "Albumin", "CK", "OverallPos", "Admitted",  
16     "AdmittedToITU", "ThirtyDayFU", "Dyspnoea",  
17     "Comorbidity", "XRPositive"))  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

5 Imputation

5.0.0.0.1 This code generates 15 imputed datasets using the permuted mean matching method, based on the 'limcompletedata' dataset which has filtered the most relevant fields, with minimal missing data initially

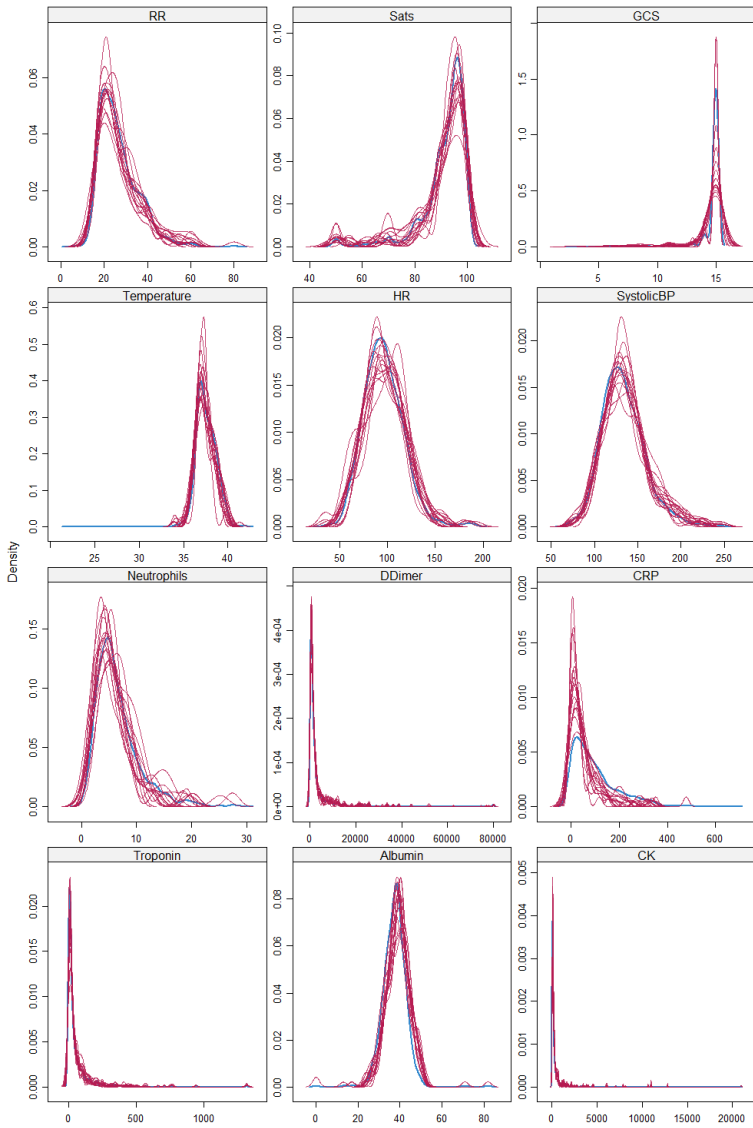
```
imputed <- mice(limcompletedata, m = 15,  
method = "pmm")
```

5.0.0.0.2 Imputation Diagnostics Density plot, this corresponds to supplementary figure 1:

```
densityplot(imputed)
```


22

5 Imputation



6 Propensity Score Matching

6.0.0.0.1 This code matches data in the imputed datasets on whether the XR was reported classical COVID or not, the matching is done based on the covariates Sex, Age, Comorbidity, Ethnicity and Respiratory Rate

```

library(MatchThem)
#### MatchThem package requires dependent variable to be coded as 0 or 1
imputed[["data"]][["XRPositive"]] %>% recode_factor("Positive" = "1",
"Negative" = "0") ->imputed[["data"]][["XRPositive"]]
matchthem(
  XRPositive ~ Sex + Age + Comorbidity + Ethnicity + RR,
  data = imputed,
  method = 'nearest',
  verbose = FALSE,
  replace = FALSE,
  ratio = 1,
  caliper = 0.2,
  m.order = "random",) -> matchedtest
### Set XRChest to unordered for binomial analyses
matchedtest[["datasets"]]c(1:15)[["XRChest"]] %>% factor(ordered = FALSE) ->
  matched2[["datasets"]]c(1:15)[["XRChest"]]

```

6.1 Match Balance Diagnostics

6.1.0.0.1 Creates plots and table with mean difference and distribution of values in covariates between XR +ve and -ve groups after matching across all imputed datasets:

```

#### Supplementary tables 1,2 and 3:
bal.tab(matchedtest)
#### Supplementary figure 2
bal.plot(matchedtest)
#### Supplementary figure 3:
bal.plot(matchedtest, var.name = "Age", type = "histogram",
  which = "both")
bal.plot(matchedtest, var.name = "Sex", type = "histogram",
  which = "both")
bal.plot(matchedtest, var.name = "Ethnicity",

```

```
1     type = "histogram", which = "both")
2
3
4     bal.plot(matchedtest, var.name = "RR", type = "histogram",
5             which = "both")
6     bal.plot(matchedtest, var.name = "Comorbidity",
7             type = "histogram", which = "both")
8     ### Supplementary figure 4:
9     love.plot(matchedtest)
```

7 Matched Demographics Table:

7.0.0.0.1 Stack matched imputed datasets into one large dataset and split into COVID +ve and -ve groups:

```
### 'all=FALSE' gets matched data only
stacked <- MatchThem::complete(matchedtest,
  n = c(1:15), all = FALSE)
stacked <- stacked %>% filter(.imp > 0)
```

7.0.0.0.2 Creates demographics table as above, but on propensity matched imputed datasets, corresponds to Table 4:

```
table4 <- CreateTableOne(strata = "OverallPos",
  data = stacked)
#### Means and SD kept as is, mean counts
#### calculated after dividing by 15 (as 15
#### imputed datasets)
```

7.0.0.0.3 Creates demographic table stratified by XR Positive or Negative on matched imputed datasets, corresponds to Table 5:

```
table5 <- CreateTableOne(strata = "XRPositive",
  data = stacked)
#### Means and SD kept as is, mean counts
#### calculated after dividing by 15 (as 15
#### imputed datasets)
```

7.0.0.0.4 Summary statistics for pooled data:

```
### Normal means sd
explanatorynorm <- c("Age", "Temperature",
  "HR", "SystolicBP")
summarynormalOverallPos <- stacked %>% group_by(OverallPos) %>%
```

26

7 Matched Demographics Table:

```
summarise_at(vars(explanatorynorm), list(mean.default,
sd))
summarynormalXRPositive <- stacked %>% group_by(XRPositive) %>%
summarise_at(vars(explanatorynorm), list(mean.default,
sd))

### Non normal medians and IQR
summarynormalOverallPos <- stacked %>% group_by(OverallPos) %>%
summarise_at(vars(explanatorynormal),
list(median, IQR))
summarynormalXRPositive <- stacked %>% group_by(XRPositive) %>%
summarise_at(vars(explanatorynormal),
list(median, IQR))
```

8 Diagnostic Accuracy

8.0.0.1 This section generates the diagnostic accuracy statistics (e.g. sensitivity, specificity) for CXR and CT with RT-PCR as the reference standard using the matched imputed datasets

8.0.0.2 This code creates a contingency table of False/ True Positives and Negatives for Chest X-ray taken from the demographic tables above:

```
contingxr <- matrix(c(305, 243, 125, 187),
  nrow = 2, ncol = 2)

colnames(contingxr) <- c("PCR+", "PCR-")

rownames(contingxr) <- c("XR+", "XR-")
```

8.0.0.2.1 This function calculates diagnostic accuracy test statistics:

```
xraccuracy <- epi.tests(contingxr, conf.level = 0.95)
```

8.0.0.3 Giving the diagnostic accuracy output for CXR in table 3:

```
xraccuracy
  Outcome +   Outcome -   Total
Test +       305         125     430
Test -       243         187     430
Total        548         312     860

Point estimates and 95 % CIs:
-----
Apparent prevalence           0.50 (0.47, 0.53)
True prevalence              0.64 (0.60, 0.67)
```

Sensitivity	0.56 (0.51, 0.60)
Specificity	0.60 (0.54, 0.65)
Positive predictive value	0.71 (0.66, 0.75)
Negative predictive value	0.43 (0.39, 0.48)
Positive likelihood ratio	1.39 (1.19, 1.62)
Negative likelihood ratio	0.74 (0.65, 0.84)

8.0.0.3.0.1 NB diagnostic accuracy values in table available in list view of xrraccuracy variable

8.1 CT Data and Accuracy

8.1.0.0.0.1 Only those with CT and RT PCR:

```
CTdata <- filter(data, is.na(data$CTBSTI) ==
  FALSE & is.na(data$RTPCR) == FALSE)
```

8.1.0.0.0.2 Select relevant variables

```
CTdata <- dplyr::select(CTdata, c("Age",
  "XRchest", "Ethnicity", "Sex", "RR",
  "Sats", "GCS", "Temperature", "HR", "SystolicBP",
  "DiastolicBP", "Neutrophils", "DDimer",
  "CRP", "Troponin", "OverallPos", "Admitted",
  "AdmittedToITU", "ThirtyDayFU", "Dyspnoea",
  "Comorbidity", "XRPositive", "OverallPos",
  "CTBSTI"))
```

8.1.0.0.0.3 Set RT-PCR as factor:

```
CTdata$OverallPos <- as.factor(CTdata$OverallPos)
```

8.1 CT Data and Accuracy

29

8.1.0.0.4 Rename 1 and 0 to Positive and Negative:

```
CTdata$CTPositive <- ifelse(CTdata$CTBSTI ==
  "1", "Positive", "Negative")
CTdata$CTPositive <- as.factor(CTdata$CTPositive)
```

8.1.0.0.5 Regression with CT as outcome variable:

```
CT <- finalfit(
  CTdata,
  "OverallPos",
  c(
    "Age",
    "Sex",
    "RR",
    "GCS",
    "CTPositive",
    "Temperature",
    "HR",
    "SystolicBP",
    "DiastolicBP",
    "Sats",
    "Dyspnoea",
    "Comorbidity"
  ),
  confint_level = 0.95
)
```

8.1.0.0.6 Contingency table of True/False Positives and Negatives for CT taken from Regression table:

```
contingct <- matrix(c(CT[7, 4], CT[6, 4],
  CT[7, 3], CT[6, 3]), nrow = 2, ncol = 2)
colnames(contingct) <- c("PCR+", "PCR-")
rownames(contingct) <- c("CT+", "CT-")
contingct <- substr(contingct, start = 1,
  stop = 3)
contingct <- sapply(contingct, as.numeric)
contingct <- matrix(contingct, nrow = 2,
  ncol = 2)
colnames(contingct) <- c("PCR+", "PCR-")
rownames(contingct) <- c("CT+", "CT-")
```


30

8 Diagnostic Accuracy

8.1.0.0.7 Diagnostic accuracy statistics for CT

```

epi.tests(contingct, conf.level = 0.95) -> ctaccuracy
      Outcome +   Outcome -   Total
Test +         162         55     217
Test -          29         56      85
Total          191        111     302

```

Point estimates and 95 % CIs:

```

-----
Apparent prevalence           0.72 (0.66, 0.77)
True prevalence               0.63 (0.58, 0.69)
Sensitivity                   0.85 (0.79, 0.90)
Specificity                   0.50 (0.41, 0.60)
Positive predictive value     0.75 (0.68, 0.80)
Negative predictive value     0.66 (0.55, 0.76)
Positive likelihood ratio     1.71 (1.41, 2.08)
Negative likelihood ratio     0.30 (0.21, 0.44)
-----

```

8.1.0.0.8 NB Diagnostic accuracy values found in list view rather than output

8.2 CT and XR accuracy comparison

8.2.0.1 In this section mean differences of diagnostic accuracy statistics between CT and Chest X-ray with confidence intervals and p-values are calculated

8.2.1 Sensitivity

8.2 CT and XR accuracy comp...

31

8.2.1.0.0.1 Upper confidence limit for difference in sensitivity

```
ubsens <- (ctaccuracy[["elements"]][["se.up"]] -  
xraccuracy[["elements"]][["se.low"]])
```

8.2.1.0.0.2 Lower confidence limit for difference in sensitivity

```
lbsens <- (ctaccuracy[["elements"]][["se.low"]] -  
xraccuracy[["elements"]][["se.up"]])
```

8.2.1.0.0.3 Mean difference in sensitivity

```
meansens <- ctaccuracy[["elements"]][["se"]] -  
xraccuracy[["elements"]][["se"]]
```

8.2.1.0.0.4 Standard error for sensitivity

```
sesens <- (ubsens - lbsens)/(2 * 1.96)
```

8.2.1.0.0.5 value for difference in sensitivity

```
zsens <- meansens/sesens
```

8.2.1.0.0.6 P-value for difference in sensitivity

```
psens <- exp(-0.717 * zsens - 0.416 * zsens^2)
```

32

8 Diagnostic Accuracy

8.2.1.0.0.7 Format values into 'mean difference (95% CI) p-value' rounded to 2 d.p.

```
diffsens <- sprintf("%s (%s-%s)", round(meansens,
  digits = 2), round(lbsens, digits = 2),
  round(ubsens, digits = 2))
diffsensp <- c(diffsens, psens)
```

8.2.1.0.0.8 Subsequent analyses in this section follow the code above

```
## Specificity
ubspec <- (ctaccuracy[["elements"]][["sp.up"]] -
  xraccuracy[["elements"]][["sp.low"]])
lbspec <- (ctaccuracy[["elements"]][["sp.low"]] -
  xraccuracy[["elements"]][["sp.up"]])
meanspec <- ctaccuracy[["elements"]][["sp"]] -
  xraccuracy[["elements"]][["sp"]]
sespec <- (ubspec - lbspec)/(2 * 1.96)
zspec <- meanspec/sespec
pspec <- exp(-0.717 * zspec - 0.416 * zspec^2)
diffspec <- sprintf("%s (%s-%s)", round(meanspec,
  digits = 2), round(lbspec, digits = 2),
  round(ubspec, digits = 2))
diffspecp <- c(diffspec, pspec)

ubda <- (ctaccuracy[["elements"]][["da.up"]] -
  xraccuracy[["elements"]][["da.low"]])
lbda <- (ctaccuracy[["elements"]][["da.low"]] -
  xraccuracy[["elements"]][["da.up"]])
meanda <- ctaccuracy[["elements"]][["da"]] -
  xraccuracy[["elements"]][["da"]]
seda <- (ubda - lbda)/(2 * 1.96)
zda <- meanda/seda
pda <- exp(-0.717 * zda - 0.416 * zda^2)
diffda <- sprintf("%s (%s-%s)", round(meanda,
  digits = 2), round(lbda, digits = 2),
  round(ubda, digits = 2))
diffdap <- c(diffda, pda)

## Positive Likelihood Ratio
ublrpos <- (ctaccuracy[["elements"]][["lrpos.up"]] -
  xraccuracy[["elements"]][["lrpos.low"]])
lblrpos <- (ctaccuracy[["elements"]][["lrpos.low"]] -
  xraccuracy[["elements"]][["lrpos.up"]])
meanlrpos <- ctaccuracy[["elements"]][["lrpos"]] -
  xraccuracy[["elements"]][["lrpos"]]
selrpos <- (ublrpos - lblrpos)/(2 * 1.96)
zlrpos <- meanlrpos/selrpos
plrpos <- exp(-0.717 * zlrpos - 0.416 * zlrpos^2)
difflrpos <- sprintf("%s (%s-%s)", round(meanlrpos,
  digits = 2), round(lblrpos, digits = 2),
```

8.2 CT and XR accuracy comp...

33

```

1   round(ublrrpos, digits = 2))
2   diffplrposp <- c(diffplrpos, plrpos)
3   ## Negative Likelihood Ratios
4   ublrneg <- (ctaccuracy[["elements"]][["lrneg.up"]] -
5     xraccuracy[["elements"]][["lrneg.low"]])
6   lblrneg <- (ctaccuracy[["elements"]][["lrneg.low"]] -
7     xraccuracy[["elements"]][["lrneg.up"]])
8   meanlrneg <- ctaccuracy[["elements"]][["lrneg"]] -
9     xraccuracy[["elements"]][["lrneg"]]
10  selrneg <- (ublrneg - lblrneg)/(2 * 1.96)
11  zlrneg <- meanlrneg/selrneg
12  plrneg <- exp(-0.717 * zlrneg - 0.416 * zlrneg^2)
13  diffflrneg <- sprintf("%s (%s-%s)", round(meanlrneg,
14    digits = 2), round(lblrneg, digits = 2),
15    round(ublrneg, digits = 2))
16  diffflrnegp <- c(diffflrneg, plrneg)
17
18  ## Positive Predictive Value
19  ppv <- (ctaccuracy[["elements"]][["ppv.low"]] -
20    xraccuracy[["elements"]][["ppv.up"]])
21  meanppv <- ctaccuracy[["elements"]][["ppv"]] -
22    xraccuracy[["elements"]][["ppv"]]
23  seppv <- (ubppv - lbppv)/(2 * 1.96)
24  zppv <- meanppv/seppv
25  pppv <- exp(-0.717 * zppv - 0.416 * zppv^2)
26  diffppv <- sprintf("%s (%s-%s)", round(meanppv,
27    digits = 2), round(lbppv, digits = 2),
28    round(ubppv, digits = 2))
29  diffppvp <- c(diffppv, pppv)
30
31  ## Negative Predictive Value
32  npv <- (ctaccuracy[["elements"]][["npv.low"]] -
33    xraccuracy[["elements"]][["npv.up"]])
34  meannpv <- ctaccuracy[["elements"]][["npv"]] -
35    xraccuracy[["elements"]][["npv"]]
36  senpv <- (ubnpv - lbnpv)/(2 * 1.96)
37  znpv <- meannpv/senpv
38  pnpv <- exp(-0.717 * znpv - 0.416 * znpv^2)
39  diffnpv <- sprintf("%s (%s-%s)", round(meannpv,
40    digits = 2), round(lbnpv, digits = 2),
41    round(ubnpv, digits = 2))
42  diffnpvp <- c(diffnpv, pnpv)
43
44  ## Apparent Prevalence
45  meantp <- ctaccuracy[["elements"]][["tp"]] -
46    xraccuracy[["elements"]][["tp"]]
47  setp <- (ubtp - lbtp)/(2 * 1.96)
48  ztp <- meantp/setp
49  ptp <- exp(-0.717 * ztp - 0.416 * ztp^2)
50  diffftp <- sprintf("%s (%s-%s)", round(meantp,
51    digits = 2), round(lbtp, digits = 2),
52    round(ubtp, digits = 2))
53  diffftp <- c(diffftp, ptp)
54
55  ## True Prevalence
56  meanap <- ctaccuracy[["elements"]][["ap"]] -
57    xraccuracy[["elements"]][["ap"]]

```

```

seap <- (ubap - lbap)/(2 * 1.96)
zap <- meanap/seap
pap <- exp(-0.717 * zap - 0.416 * zap^2)
diffap <- sprintf("%s (%s-%s)", round(meanap,
  digits = 2), round(lbap, digits = 2),
  round(ubap, digits = 2))
diffapp <- c(diffap, pap)

```

8.3 Intermodality Agreement

8.3.0.0.1 This section contains code to analyse the level of agreement in the unmatched CT dataset which contains only data with CT, XR and RT-PCR

8.3.0.0.2 First- comparing CT and XR agreement

```

library(irr)
kappa2(c(CTdata$XRPositive, CTdata$CTPositive),
  weight = "squared")
d <- CTdata %>% select(c("CTPositive", "XRPositive"))
View(d)
kappa2(d, weight = "squared")

```

8.3.0.0.3 Output:

```

Cohen's Kappa for 2 Raters (Weights: squared)

Subjects = 287
Raters = 2
Kappa = 0.406

z = 7.14
p-value = 9.37e-13

```

8.3.0.0.4 The following code compares RT-PCR, CT and XR

```

d2 <- CTdata %>% select(c("CTPositive", "XRPositive",
  "OverallPos"))
View(d2)
kappam.fleiss(d2)

```

8.3 Intermodality Agreement

35

8.3.0.0.5 Output:

```

Fleiss' Kappa for m Raters

Subjects = 287
Raters = 3
Kappa = 0.361

z = 10.6
p-value = 0

```

8.3.1 Diagnostic Accuracy Analysis when Indeterminate Reports of CXR and CT are taken as positive

8.3.1.1 XR Indeterminates

8.3.1.1.0.1 New column for positive if indeterminate

```

stacked$XRIndPositive <- ifelse(stacked$XR Chest ==
  "Classic COVID" | stacked$XR Chest ==
  "Indeterminate", "Positive", "Negative")
stacked$XRIndPositive <- as.factor(stacked$XRIndPositive)
stackedpos <- stacked %>% filter(OverallPos ==
  "Positive")
stackedneg <- stacked %>% filter(OverallPos ==
  "Negative")
summary(stackedpos$XRIndPositive)
summary(stackedneg$XRIndPositive)

contingxrind <- matrix(c(441, 107, 186, 126),
  nrow = 2, ncol = 2)
colnames(contingxrind) <- c("PCR+", "PCR-")

rownames(contingxrind) <- c("XR+", "XR-")
xrindaccuracy <- epi.tests(contingxrind)

```

8.3.1.1.0.2 In this section mean differences of diagnostic accuracy statistics between CT (when CT indeterminates are not counted as positive) and Chest X-ray with confidence intervals and p-values are calculated, follows the same pattern as code previously

```
##### Sensitivity Upper confidence Limit for
##### difference in sensitivity

ubsens <- (ctaccuracy[["elements"]][["se.up"]] -
  xrindaccuracy[["elements"]][["se.low"]])
## Lower confidence limit for difference
## in sensitivity
lbsens <- (ctaccuracy[["elements"]][["se.low"]] -
  xrindaccuracy[["elements"]][["se.up"]])
## Mean difference in sensitivity
meansens <- ctaccuracy[["elements"]][["se"]] -
  xrindaccuracy[["elements"]][["se"]]
## Standard error for sensitivity
sesens <- (ubsens - lbsens)/(2 * 1.96)
## Z value for difference in sensitivity
zsens <- meansens/sesens
## P-value for difference in sensitivity
psens <- exp(-0.717 * zsens - 0.416 * zsens^2)
### Format values into 'mean difference
### (95% CI) p-value' rounded to 2 d.p.
diffsens <- sprintf("%s (%s-%s)", round(meansens,
  digits = 2), round(lbsens, digits = 2),
  round(ubsens, digits = 2))
diffsensp <- c(diffsens, psens)

### Subsequent analyses in this section
### follow the code above Specificity
ubspec <- (ctaccuracy[["elements"]][["sp.up"]] -
  xrindaccuracy[["elements"]][["sp.low"]])
lbspec <- (ctaccuracy[["elements"]][["sp.low"]] -
  xrindaccuracy[["elements"]][["sp.up"]])
meanspec <- ctaccuracy[["elements"]][["sp"]] -
  xrindaccuracy[["elements"]][["sp"]]
sespec <- (ubspec - lbspec)/(2 * 1.96)
zspec <- meanspec/sespec
pspec <- exp(-0.717 * zspec - 0.416 * zspec^2)
diffspec <- sprintf("%s (%s-%s)", round(meanspec,
  digits = 2), round(lbspec, digits = 2),
  round(ubspec, digits = 2))
diffspecp <- c(diffspec, pspec)

ubda <- (ctaccuracy[["elements"]][["da.up"]] -
  xrindaccuracy[["elements"]][["da.low"]])
lbda <- (ctaccuracy[["elements"]][["da.low"]] -
  xrindaccuracy[["elements"]][["da.up"]])
meanda <- ctaccuracy[["elements"]][["da"]] -
  xrindaccuracy[["elements"]][["da"]]
seda <- (ubda - lbda)/(2 * 1.96)
```

8.3 Intermodality Agreement

37

```

1  zda <- meanda/seda
2
3
4  pda <- exp(-0.717 * zda - 0.416 * zda^2)
5  diffda <- sprintf("%s (%s-%s)", round(meanda,
6      digits = 2), round(lbda, digits = 2),
7      round(ubda, digits = 2))
8  diffdap <- c(diffda, pda)
9  ## Positive Likelihood Ratio
10 ublrpos <- (ctaccuracy[["elements"]][["lrpos.up"]] -
11    xrindaccuracy[["elements"]][["lrpos.low"]])
12 lblrpos <- (ctaccuracy[["elements"]][["lrpos.low"]] -
13    xrindaccuracy[["elements"]][["lrpos.up"]])
14 meanlrpos <- ctaccuracy[["elements"]][["lrpos"]] -
15    xrindaccuracy[["elements"]][["lrpos"]]
16 selrpos <- (ublrpos - lblrpos)/(2 * 1.96)
17 zlrpos <- meanlrpos/selrpos
18 plrpos <- exp(-0.717 * zlrpos - 0.416 * zlrpos^2)
19 diffplrpos <- sprintf("%s (%s-%s)", round(meanlrpos,
20    digits = 2), round(lblrpos, digits = 2),
21    round(ublrpos, digits = 2))
22 diffplrposp <- c(diffplrpos, plrpos)
23 ## Negative Likelihood Ratios
24 ublrneg <- (ctaccuracy[["elements"]][["lrneg.up"]] -
25    xrindaccuracy[["elements"]][["lrneg.low"]])
26 lblrneg <- (ctaccuracy[["elements"]][["lrneg.low"]] -
27    xrindaccuracy[["elements"]][["lrneg.up"]])
28 meanlrneg <- ctaccuracy[["elements"]][["lrneg"]] -
29    xrindaccuracy[["elements"]][["lrneg"]]
30 selrneg <- (ublrneg - lblrneg)/(2 * 1.96)
31 zlrneg <- meanlrneg/selrneg
32 plrneg <- exp(-0.717 * zlrneg - 0.416 * zlrneg^2)
33 diffplrneg <- sprintf("%s (%s-%s)", round(meanlrneg,
34    digits = 2), round(lblrneg, digits = 2),
35    round(ublrneg, digits = 2))
36 diffplrnegp <- c(diffplrneg, plrneg)
37
38 ## Positive Predictive Value
39 ppv <- (ctaccuracy[["elements"]][["ppv.low"]] -
40    xrindaccuracy[["elements"]][["ppv.up"]])
41 meanppv <- ctaccuracy[["elements"]][["ppv"]] -
42    xrindaccuracy[["elements"]][["ppv"]]
43 seppv <- (ubppv - lbppv)/(2 * 1.96)
44 zppv <- meanppv/seppv
45 pppv <- exp(-0.717 * zppv - 0.416 * zppv^2)
46 diffppv <- sprintf("%s (%s-%s)", round(meanppv,
47    digits = 2), round(lbppv, digits = 2),
48    round(ubppv, digits = 2))
49 diffppvp <- c(diffppv, pppv)
50
51 ## Negative Predictive Value
52 npv <- (ctaccuracy[["elements"]][["npv.low"]] -
53    xrindaccuracy[["elements"]][["npv.up"]])
54 meannpv <- ctaccuracy[["elements"]][["npv"]] -
55    xrindaccuracy[["elements"]][["npv"]]
56 senpv <- (ubnpv - lbnpv)/(2 * 1.96)
57 znpv <- meannpv/senpv
58 pnpv <- exp(-0.717 * znpv - 0.416 * znpv^2)
59 diffnpv <- sprintf("%s (%s-%s)", round(meannpv,
60    digits = 2), round(lbnpv, digits = 2),

```



```

4     round(ubnpv, digits = 2))
5     diffnpvp <- c(diffnpv, pnpv)

6
7     ## True Prevalence
8     meantp <- ctaccuracy[["elements"]][["tp"]] -
9       xrindaccuracy[["elements"]][["tp"]]
10    setp <- (ubtp - lbtp)/(2 * 1.96)
11    ztp <- meantp/setp
12    ptp <- exp(-0.717 * ztp - 0.416 * ztp^2)
13    diffftp <- sprintf("%s (%s-%s)", round(meantp,
14      digits = 2), round(lbtp, digits = 2),
15      round(ubtp, digits = 2))
16    diffftpp <- c(diffftp, ptp)

17    ## Apparent Prevalence
18    meanap <- ctaccuracy[["elements"]][["ap"]] -
19      xrindaccuracy[["elements"]][["ap"]]
20    seap <- (ubap - lbap)/(2 * 1.96)
21    zap <- meanap/seap
22    pap <- exp(-0.717 * zap - 0.416 * zap^2)
23    diffpap <- sprintf("%s (%s-%s)", round(meanap,
24      digits = 2), round(lbap, digits = 2),
25      round(ubap, digits = 2))
26    diffpapp <- c(diffpap, pap)

```

8.3.1.2 CT Indeterminates

8.3.1.2.0.1 New column for positive if indeterminate

```

27
28
29     CTdata$CTIndPositive <- ifelse(CTdata$CTBSTI ==
30       "1" | CTdata$CTBSTI == "2", "Positive",
31       "Negative")
32     CTdata$CTIndPositive <- as.factor(CTdata$CTIndPositive)
33     valuesctind <- CTdata %>% group_by(OverallPos,
34       CTIndPositive) %>% summarise(n = n())
35     ctcontingind <- matrix(data = c(178, 13,
36       70, 41), nrow = 2, ncol = 2)
37
38     colnames(ctcontingind) <- c("PCR+ve", "PCR-ve")
39     rownames(ctcontingind) <- c("CT+ve", "CT-ve")
40     ctindaccuracy <- epi.tests(ctcontingind)

```

9 Pooled Regression after Multiple Imputation and Propensity Score Matching

9.0.0.0.0.1 Binomnal Logistic regression with RT-PCR as dependent variable

```

overallposmatchimp <- matchedtest %>% with(glm(formula(ff_formula(dependent =
  "OverallPos",
  explanatory = c("Age", "Ethnicity", "Sex",
    "RR", "GCS", "Temperature", "HR",
    "SystolicBP", "Neutrophils", "DDimer",
    "CRP", "Troponin", "Albumin", "CK",
    "Sats", "Admitted", "AdmittedToITU",
    "ThirtyDayFUTwo", "Dyspnoea", "Comorbidity",
    "XRchest"))), family = "binomial"),
  all = FALSE)
P <- overallposmatchimp %>% pool()
multivarpooleoverallpos = P %>% fit2df(estimate_name = "OR (multiple
  imputation)",
  exp = TRUE)

```

9.0.0.0.0.2 'multivarpooleoverallpos' produces multivariate odds ratios for each explanatory variable, corresponding to Table 4

9.0.1 Pooled Univariate Odds Ratios for OverallPos as dependent variable

9.0.1.0.0.1 This code is run with each of the explanatory variables in table 4 as arguments to produce their respective odds Ratios in table 4

```

overallposmatchimpunivar <- matchedtest %>%
  with(glm(formula(ff_formula(dependent = "OverallPos",

```

40

9 Pooled Regression after Multi...

```

    explanatory = "XRchest")), family = "binomial"))
P <- overallposmatchimpunivar %>% pool()
univarpooledoverallpos = univaroverallpos <- P %>%
  fit2df(estimate_name = "OR (univariate)",
        exp = TRUE)
univaroverallpos

```

9.0.2 Binomial Logistic Regression with Positive Chest X-ray Report as Dependent Variable

9.0.2.0.0.1 This code follows the format above to produce univariate and multivariate odds ratios for each explanatory variable for having a positive XR report

9.0.3 Univariate XRPositive as dependent

9.0.3.0.0.1 (different explanatory variables passed into function to produce Odds ratios for each)

```

XRchestmatchimp <- matchedtest %>% with(glm(formula(ff_formula(dependent =
  "XRPositive",
  explanatory = "Comorbidity")), family = "binomial"))
P <- XRchestmatchimp %>% pool()
multivarpooledXRchest = univarXRchest <- P %>%
  fit2df(estimate_name = "OR (univariate)",
        exp = TRUE)
univarXRchest

```

9.0.4 Multivariate XRPositive as dependent

```

XRchestmatchimp <- matchedtest %>% with(glm(formula(ff_formula(dependent =
  "XRPositive",
  explanatory = c("Age", "OverallPos",
  "Ethnicity", "Sex", "RR", "GCS",
  "Temperature", "HR", "SystolicBP",
  "Neutrophils", "DDimer", "CRP", "Troponin",
  "Albumin", "CK", "Sats", "Admitted",
  "AdmittedToITU", "ThirtyDayFUTwo",
  "Dyspnoea", "Comorbidity"))), family = "binomial"))
P <- XRchestmatchimp %>% pool()
multivarpooledXRchest = multivarXRchest <- P %>%
  fit2df(estimate_name = "OR (multivariate)",

```

9.1 Forest Plots

41

```

exp = TRUE)
multivarXRchest

```

9.0.5 Pooled Ordinal Logistic Regression with XRPositive as dependent

9.0.5.0.1 This code also produces multivariate odds ratios for table 5, however, uses ordinal linear regression after the CXR report variable is converted to an ordered categorical variable, with alternative pathology as the lowest and classic covid as the highest value (see table 3)

```

XRchestmatchimpord <- matchedtest %>% with(clm(formula = XRchest ~
  Age + OverallPos + Ethnicity + Sex +
  RR + GCS + Temperature + HR + SystolicBP +
  Neutrophils + DDimer + CRP + Troponin +
  Sats + Admitted + AdmittedToITU +
  ThirtyDayFUTwo + Dyspnoea + Comorbidity))
P <- pool(object = XRchestmatchimpord[["analyses"]])
multivarXRchestord = multivarXRchestord <- P %>%
  fit2df(estimate_name = "OR (multivariate)",
        exp = TRUE)
multivarXRchestord

```

9.1 Forest Plots

9.1.0.0.1 Creates forest plots for post matched regression tables above:

```

Figure1Forest <- read_excel("Figure1Forest.xlsx",
  col_types = c("text", "numeric", "numeric",
  "numeric", "text", "text"))

tabletext1 <- cbind(Figure1Forest$explanatory,
  Figure1Forest$summary)
forestplot(tabletext1, Figure1Forest$Mean,
  Figure1Forest$Lower, Figure1Forest$Upper,
  is.summary = FALSE, clip = c(0, 2), xlab = "<U+2190> Decreased Odds SARS-
  CoV 2 Increased Odds SARS-CoV 2 <U+2192>",
  zero = 1, cex = 0.9, lineheight = unit(6,
  "mm"), boxsize = 0.4, colgap = unit(6,
  "mm"), lwd.ci = 2, ci.vertices = TRUE,
  ci.vertices.height = 0.4, title = "Odds Ratio of Positivity for SARS-CoV 2
  by RT-PCR",
  txt_gp = fpTxtGp(label = gpar(cex = 1.25),
  ticks = gpar(cex = 1.1), xlab = gpar(cex = 1.2),

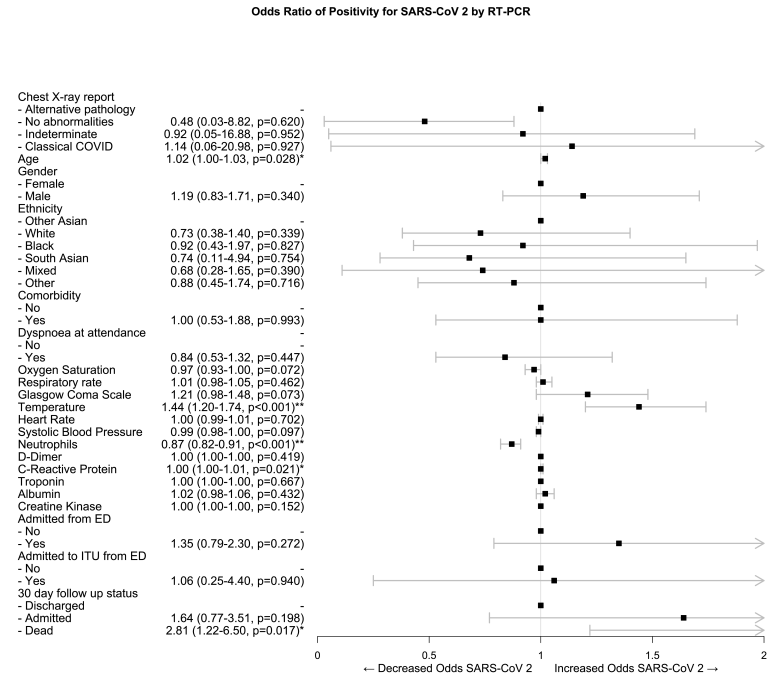
```

42

9 Pooled Regression after Multi...

```
title = gpar(cex = 1.2)), graphwidth = unit(200,
"mm"))
```

9.1.0.0.2 Figure 2:



9.1.0.0.3 Figure 3 (XR dependent):

```
Figure2Forest <- read_excel("Figure2Forest.xlsx",
col_types = c("text", "numeric", "numeric",
"numeric", "text", "text"))

tabletext2<-cbind(Figure2Forest$explanatory,Figure2Forest$summary)
forestplot (tabletext2, Figure2Forest$Mean,
Figure2Forest$Lower, Figure2Forest$Upper, is.summary = FALSE,
clip = c(0, 2),
```

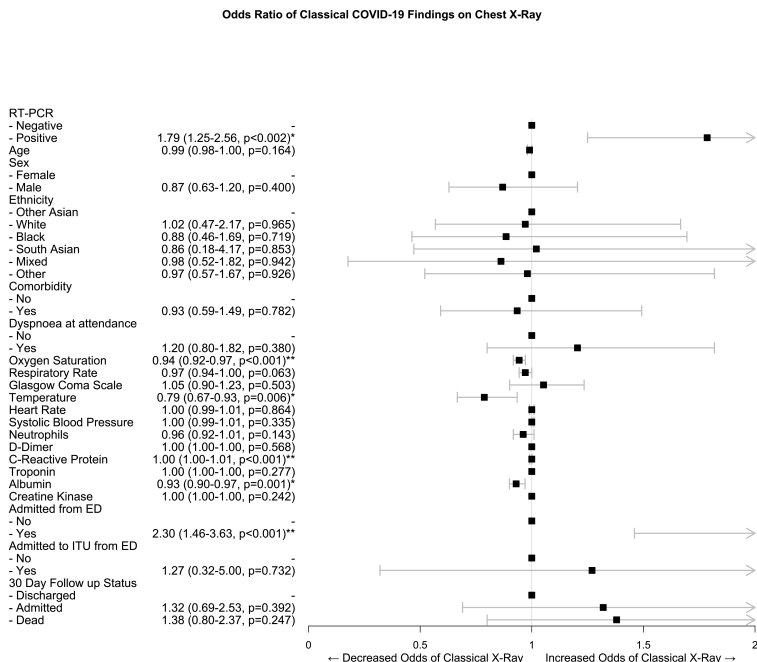
9.1 Forest Plots

43

```

xlab="\u2190 Decreased Odds of Classical X-Ray          Increased Odds
of Classical X-Ray \u2192",
zero=1, cex=0.9, lineheight = unit(6,"mm"), boxsize=0.5,
colgap=unit(6,"mm"),
lwd.ci=2, ci.vertices=TRUE, ci.vertices.height = 0.4,
title="Odds Ratio of Classical COVID-19 Findings on Chest X-Ray",
txt_gp=fpTxtGp(label=gpar(cex=1.25),
               ticks=gpar(cex=1.1),
               xlab=gpar(cex = 1.2),
               title=gpar(cex = 1.2)),
graphwidth = unit(200,"mm")
)

```



9.2 Correlation Matrix

9.2.0.0.1 This section creates a plot of correlation between all the variables in the raw data

```
library(corrplot)
library(Hmisc)
```

9.2.0.0.2 Relevel factors so relevant value is first

```
data$XRPositive <- relevel(data$XRPositive,
  "Negative")

data$Admitted <- relevel(data$Admitted, "Discharged")
data$AdmittedToITU <- relevel(data$AdmittedToITU,
  "No")
```

9.2.0.0.3 New variable for correlation matrix

```
cor <- data
```

9.2.0.0.4 Remove variables with high missings/ data which won't work e.g. date, RT-PCR removed as it only represents initial ED swab, OverallPos used instead as this includes subsequent swabs in 30 days

```
cor <- subset(data, select = -c(CT, DateOfDeath, DateOfDischarge, RTPCR,
  DateOfVisit, DateOfSymptomOnset, FollowUpPos, TimeToDeath, NEWS))'
```

9.2.0.0.5 Format and re-name values

```
cor$CTPositive <- ifelse(cor$CTBSTI == "1",
  "Positive", "Negative")
cor$CTPositive <- as.factor(cor$CTPositive)
cor$CTPositive <- relevel(cor$CTPositive,
```

9.2 Correlation Matrix

45

```
"Negative")
cor$Death <- as.factor(ifelse(cor$ThirtyDayFU ==
  "4", "Dead", "Alive"))
cor$Death <- relevel(cor$Death, "Alive")
cor$OverallPos <- as.factor(cor$OverallPos)
cor <- sapply(cor, as.numeric)
```

9.2.0.0.6 Create new numerical correlation matrix

```
cormatrixall <- cor(cor, method = "spearman",
  use = "pairwise.complete.obs")
```

9.2.0.0.7 This variable also contains p-values so identification of only significant correlations is possible:

```
cormatrixall2 <- rcorr(as.matrix(cor), type = "spearman")
```

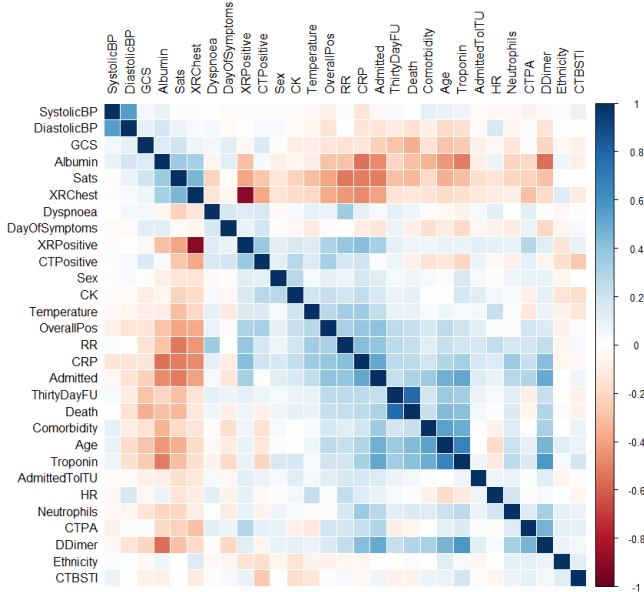
9.2.0.0.8 Function to create and format correlation matrix plot

```
corrplot(cormatrixall2$r, method = "color",
  type = "full", order = "hclust", p.mat = cormatrixall2$p,
  sig.level = 0.05, insig = "blank", tl.col = "black",
  outline = "white", title = "Correlation Matrix of Explanatory and Outcome
  Variables",
  line = -1, cex.main = 2, adj.main = 0.5)
```


46

9 Pooled Regression after Multi...

Correlation Matrix of Explanatory and Outcome Variables



9.3 STARD Flow Diagram

9.3.0.0.1 See instructions from <https://www.r-bloggers.com/flow-charts-in-r/>

9.3.0.0.2 Produces flow charts in Figure 1, (images need to be stretched out, output as svgs)

```
library(grid)
library(Gmisc)

grid.newpage()
# set some parameters to use repeatedly
leftx <- 0.25
```

9.3 STARD Flow Diagram

47

```

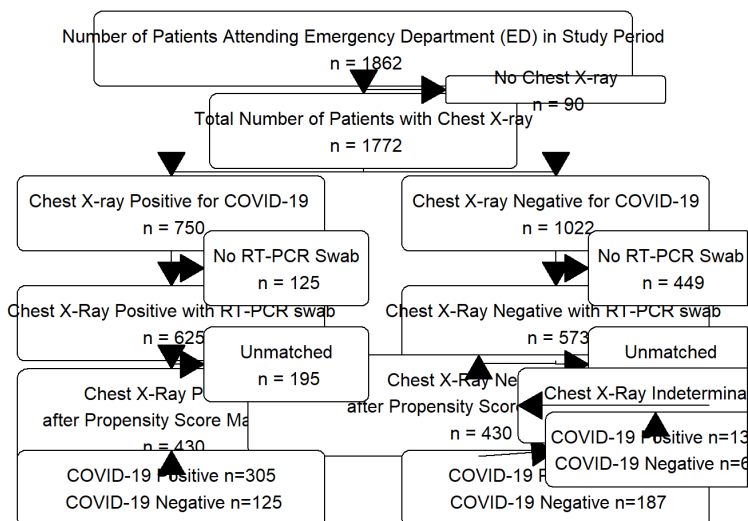
1 midx <- 0.5
2 rightx <- 0.75
3 width <- 0.4
4 gp <- gpar(fill = "white")
5 # create boxes
6 (totalattendance <- boxGrob("Number of Patients Attending Emergency Department
7 (ED) in Study Period\n n = 1862",
8 x = midx, y = 0.9, box_gp = gp, width = 0.7))
9
10 (numberwithxr <- boxGrob("Total Number of Patients with Chest X-ray\n n =
11 1772",
12 x = midx, y = 0.75, box_gp = gp, width = width))
13 # connect boxes like this
14 connectGrob(totalattendance, numberwithxr,
15 "v")
16
17 (numberwithoutxr <- boxGrob("No Chest X-ray\n n = 90",
18 x = rightx, y = 0.825, box_gp = gp, width = unit(2,
19 "inch"), height = 0.05))
20
21 connectGrob(totalattendance, numberwithoutxr,
22 "-")
23
24 (XRPos <- boxGrob("Chest X-ray Positive for COVID-19 \n n = 750",
25 x = leftx, y = 0.6, box_gp = gp, width = width))
26 (XRNeg <- boxGrob("Chest X-ray Negative for COVID-19\n n = 1022",
27 x = rightx, y = 0.6, box_gp = gp, width = width))
28
29 connectGrob(numberwithxr, XRPos, "N")
30 connectGrob(numberwithxr, XRNeg, "N")
31
32 (RTPCRXRPos <- boxGrob("Chest X-Ray Positive with RT-PCR swab\n n = 625",
33 x = leftx, y = 0.4, box_gp = gp, width = width))
34 (RTPCRXRNeg <- boxGrob("Chest X-Ray Negative with RT-PCR swab \n n = 573",
35 x = rightx, y = 0.4, box_gp = gp, width = width))
36
37 connectGrob(XRPos, RTPCRXRPos, "N")
38 connectGrob(XRNeg, RTPCRXRNeg, "N")
39
40 (NoRTPCRXRPos <- boxGrob("No RT-PCR Swab\n n = 125",
41 x = 0.4, y = 0.5, box_gp = gp, width = unit(1.5,
42 "inch")))
43 (NoRTPCRXRNeg <- boxGrob("No RT-PCR Swab\n n = 449",
44 x = 0.9, y = 0.5, box_gp = gp, width = unit(1.5,
45 "inch")))
46
47 connectGrob(XRPos, NoRTPCRXRPos, "-")
48 connectGrob(XRNeg, NoRTPCRXRNeg, "-")
49
50 (MatchedXRPos <- boxGrob("Chest X-Ray Positive \n after Propensity Score
51 Matching\n n = 430",
52 x = leftx, y = 0.225, box_gp = gp, width = width))
53 (MatchedXRNeg <- boxGrob("Chest X-Ray Negative \n after Propensity Score
54 Matching \n n = 430",
55 x = 0.65, y = 0.25, box_gp = gp, width = unit(4.2,
56 "inch")))
57
58 connectGrob(RTPCRXRPos, MatchedXRPos, "N")
59 connectGrob(RTPCRXRNeg, MatchedXRNeg, "N")

```

```
(UnmatchedXRPos <- boxGrob("Unmatched\n n = 195",  
  x = 0.4, y = 0.325, box_gp = gp, width = unit(1.5,  
  "inch")))  
(UnmatchedXRNeg <- boxGrob("Unmatched\n n = 143",  
  x = 0.9, y = 0.325, box_gp = gp, width = unit(1.5,  
  "inch")))  
  
connectGrob(RTPCRXRPos, UnmatchedXRPos, "-")  
connectGrob(RTPCRXRNeg, UnmatchedXRNeg, "L")  
  
(DiagXRPositive <- boxGrob("COVID-19 Positive n=305\n COVID-19 Negative n=125",  
  x = leftx, y = 0.1, box_gp = gp, width = width))  
(DiagXRNegative <- boxGrob("COVID-19 Positive n=243 \n COVID-19 Negative  
  n=187",  
  x = rightx, y = 0.1, box_gp = gp, width = width))  
  
connectGrob(MatchedXRPos, DiagXRPositive,  
  "N")  
connectGrob(MatchedXRNeg, DiagXRNegative,  
  "vertical")  
  
(XRInd <- boxGrob("Chest X-Ray Indeterminate \n n = 197",  
  x = 0.88, y = 0.25, box_gp = gp, width = unit(2.5,  
  "inch")))  
  
connectGrob(MatchedXRNeg, XRInd, "horizontal")  
  
(DiagXRInd <- boxGrob("COVID-19 Positive n=136\n COVID-19 Negative n=63",  
  x = 0.88, y = 0.17, box_gp = gp, width = unit(2,  
  "inch")))  
connectGrob(XRInd, DiagXRInd, "vertical")
```

9.3 STARD Flow Diagram

49



CT Flow Chart####

```

grid.newpage()
(totalattendance <- boxGrob("Number of Patients Attending Emergency Department
(ED) in Study Period\n n = 1862",
  x = midx, y = 0.9, box_gp = gp, width = 0.7))

(numberwithCT <- boxGrob("Total Number with Chest Computed Tomography (CT)\n n
= 319",
  x = midx, y = 0.75, box_gp = gp, width = width))
connectGrob(totalattendance, numberwithCT,
  "vertical")

(numberwithoutCT <- boxGrob("No Chest CT\n n = 1543",
  x = rightx, y = 0.825, box_gp = gp, width = unit(2,
  "inch"), height = 0.05))

connectGrob(totalattendance, numberwithoutCT,
  "-")

(CTPos <- boxGrob("CT Positive for COVID-19 \n n = 232",
  x = leftx, y = 0.6, box_gp = gp, width = width))
(CTNeg <- boxGrob("CT Negative for COVID-19\n n = 87",
  x = rightx, y = 0.6, box_gp = gp, width = width))

connectGrob(numberwithCT, CTPos, "N")
connectGrob(numberwithCT, CTNeg, "N")

(RTPCRCTPos <- boxGrob("CT Positive with RT-PCR swab\n n = 217",
  x = leftx, y = 0.4, box_gp = gp, width = width))

```

50

9 Pooled Regression after Multi...

```
(RTPCRCTNeg <- boxGrob("CT Negative with RT-PCR swab \n n = 85",
  x = rightx, y = 0.4, box_gp = gp, width = width))

connectGrob(CTPos, RTPCRCTPos, "N")
connectGrob(CTNeg, RTPCRCTNeg, "N")

(NoRTPCRCTPos <- boxGrob("No RT-PCR Swab\n n = 15",
  x = 0.4, y = 0.5, box_gp = gp, width = unit(1.5,
  "inch")))
(NoRTPCRCTNeg <- boxGrob("No RT-PCR Swab\n n = 2",
  x = 0.9, y = 0.5, box_gp = gp, width = unit(1.5,
  "inch")))

connectGrob(CTPos, NoRTPCRCTPos, "-")
connectGrob(CTNeg, NoRTPCRCTNeg, "-")

(DiagCTPositive <- boxGrob("COVID-19 Positive n=162\n COVID-19 Negative n=55",
  x = leftx, y = 0.1, box_gp = gp, width = width))
(DiagCTNegative <- boxGrob("COVID-19 Positive n=29\n COVID-19 Negative n=56",
  x = rightx, y = 0.1, box_gp = gp, width = width))

connectGrob(RTPCRCTPos, DiagCTPositive, "N")
connectGrob(RTPCRCTNeg, DiagCTNegative, "N")

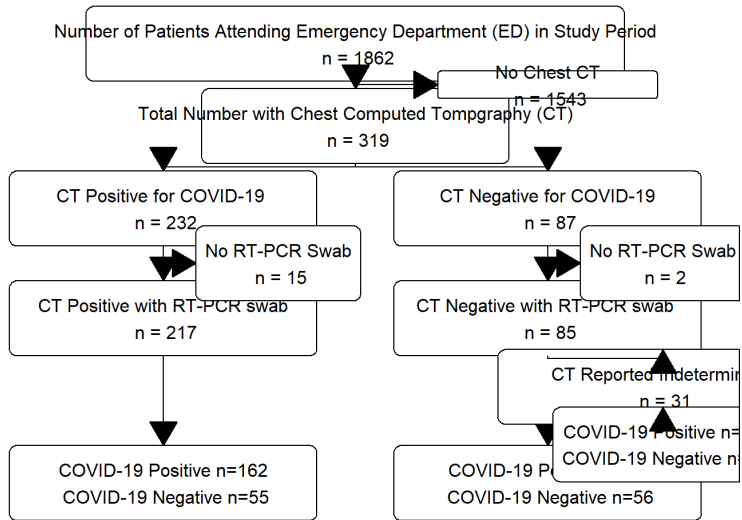
(CTInd <- boxGrob("CT Reported Indeterminate \n n = 31",
  x = 0.9, y = 0.275, box_gp = gp, width = unit(3,
  "inch")))

connectGrob(RTPCRCTNeg, CTInd, "N")

(DiagCTInd <- boxGrob("COVID-19 Positive n=16\n COVID-19 Negative n=15",
  x = 0.9, y = 0.17, box_gp = gp, width = unit(2,
  "inch")))
connectGrob(CTInd, DiagCTInd, "vertical")
```

9.3 STARD Flow Diagram

51



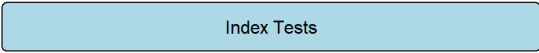
```

24  ### Labels###
25  grid.newpage()
26  (indextest <- boxGrob("Index Tests", x = midx,
27    y = 0.9, box_gp = gpar(fill = "light blue"),
28    width = 0.7))
29  (reftest <- boxGrob("Index Tests and Reference Standards",
30    x = midx, y = 0.4, box_gp = gpar(fill = "light blue"),
31    width = 0.7))
32  (finaldiag <- boxGrob("Final Diagnoses",
33    x = midx, y = 0.1, box_gp = gpar(fill = "light blue"),
34    width = 0.7))
  
```

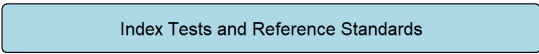
52

9 Pooled Regression after Multi...

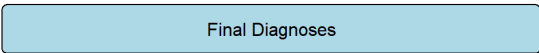
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50



Index Tests



Index Tests and Reference Standards

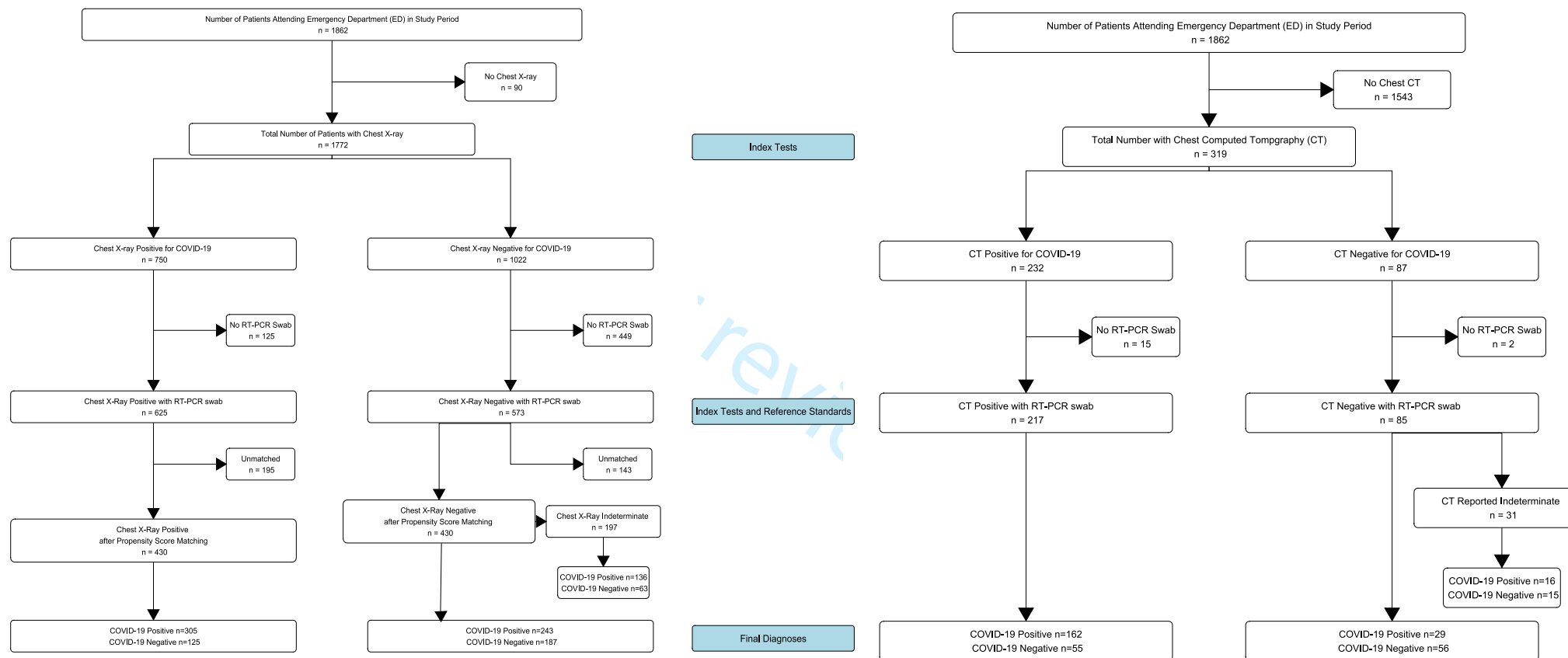


Final Diagnoses

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	5
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
<i>Participants</i>	6	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	5,20
	11	Rationale for choosing the reference standard (if alternatives exist)	N/A
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	5
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	20
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	12
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	6,7
	15	How indeterminate index test or reference standard results were handled	5
	16	How missing data on the index test and reference standard were handled	N/A, excluded
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	N/A
	18	Intended sample size and how it was determined	7
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	22, diagram below
	20	Baseline demographic and clinical characteristics of participants	21
	21a	Distribution of severity of disease in those with the target condition	21
	21b	Distribution of alternative diagnoses in those without the target condition	N/A
	22	Time interval and any clinical interventions between index test and reference standard	N/A
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	22
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	22
	25	Any adverse events from performing the index test or the reference standard	N/A
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12
	27	Implications for practice, including the intended use and clinical role of the index test	14
OTHER INFORMATION			
	28	Registration number and name of registry	N/A
	29	Where the full study protocol can be accessed	N/A
	30	Sources of funding and other support; role of funders	N/A



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Supplementary Figure- STARD Flow Diagram

For peer review only