medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Title: Computer-aided covid-19 patient screening using chest images (X-Ray and CT scans)

Authors: Xavier P. Burgos-Artizzu.

Transmural Biotech S. L. Barcelona, Spain

e-mail: xpburgos@transmuralbiotech.com

ABSTRACT

Objectives: to evaluate the performance of Artificial Intelligence (AI) methods to detect covid -19 from chest images (X-Ray and CT scans).

Methods: Chest CT scans and X-Ray images collected from different centers and institutions were downloaded and combined together. Images were separated by patient and 66% of the patients were used to develop and train AI image-based classifiers. Then, the AI automated classifiers were evaluated on a separate set of patients (the remaining 33% patients).

Results (Chest X-Ray): Five different data sources were combined for a total of N=9,841 patients (1,733 with covid-19, 810 with bacterial tuberculosis and 7,298 healthy patients). The test sample size was N=3,528 patients. The best AI method reached an Area Under the Curve (AUC) for covid-19 detection of 99%, with a detection rate of 96.4% at 1.0% false positive rate.

Results (Chest CT scans): Two different data sources were combined for a total of N=363 patients (191 having covid-19 and 172 healthy patients). The test sample size was N=121 patients. The best AI method reached an AUC for covid-19 detection of 90.9%, with a detection rate of 90.6% at 24.6% false positive rate.

Conclusions: Computer aided automatic covid-19 detection from chest X-ray images showed promising results to be used as screening tool during the covid-19 outbreak. The developed method may help to manage patients better in case access to PCR testing is not possible or to detect patients with symptoms missed in a first round of PCR testing. The method will be made available online (www.quantuscovid19.org). These results merit further evaluation collecting more images. We hope this study will allow us to start such collaborations.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

INTRODUCTION

Coronaviruses are a family of viruses that can cause illnesses such as the common cold, severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS). In December 2019, a new coronavirus was first identified in Wuhan, China. The virus is now known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), while the disease it causes is called coronavirus disease 2019 (covid-19), a highly infectious respiratory disease[1]. In March 2020, the World Health Organization (WHO) declared the covid-19 outbreak a global pandemic[2]. According to the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)[3], there have been more than 12 million confirmed cases and the pandemic has already caused over 550,000 casualties.

Symptoms of covid-19 vary in severity from having no symptoms at all (being asymptomatic) to having fever, cough, sore throat, general weakness, fatigue, muscular pain and loss of smell. In the most severe cases it can lead to severe pneumonia, acute respiratory distress syndrome (ARDS), possibly precipitated by cytokine storm, multi-organ failure, septic shock, and blood clots, which can lead to death[4–9]. The time from exposure to onset of symptoms is typically around five days, but may range from two to fourteen days[10].

The preferred method of diagnosis of covid-19 is by real-time reverse transcription polymerase chain reaction (rRT-PCR) from a nasopharyngeal swab[11,12]. PCR permits, especially in a few hours, the "acellular cloning" of a DNA fragment through an automated system, which usually takes several days with standard techniques of molecular cloning. PCR is widely used for diagnostic purposes to detect the presence of a specific DNA sequence in a biological fluid. It's use from a sample obtained from a nasal swab is generally considered the gold standard for covid-19 diagnosis and has been used as the main tool to control the pandemic[13].

However, PCR is still prone to errors as covid-19 detector [14,15], its precision varying according to the quality of the biological sample obtained[16,17]. Furthermore, PCR is difficult to be performed routinely on all the population due to its technical requirements[18] and shortages of test strips was experienced in many countries[19]. Therefore, researching other possibilities is clearly worthwhile.

In this context, perhaps the two more prominent alternatives are the use of two "everyday tools" such as chest X-Ray imaging and Computed Tomography (CT). These are widely used by clinicians worldwide to diagnose all sorts of respiratory diseases and viral infections affecting the respiratory system. Given the high tropism of covid-19 for respiratory airways, identification of lung involvement in infected patients can be relevant for treatment and monitoring of the disease. Using these widely used tools has its advantages, especially because of their readiness to be applied in most centers, which is a key factor considering that timely diagnosis of covid-19 is of adamant importance both for the patient's successful treatment and to control the pandemic spread[20–24].

Researchers have already looked at the use of both techniques during covid-19 pandemic[25–30]. Each seems to have in principle its advantages over the other. CT is supposedly more precise, while X-Ray is more practical and can be very useful especially in emergency settings due to the existence of portable devices.

In this study, we wanted to evaluate how well current state-of-the-art Artificial Intelligence (AI) methods could detect covid-19 from patients showing some symptoms using

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

both X-ray and CT chest medical images in a completely automated manner. Of course, pathological findings found through these techniques cannot be specific to the type of virus causing the pathology. Covid-19 is a type of viral pneumonia caused by the SARS-CoV-2 virus, and has many things in common with other types of viral pneumonia and infections such as influenza, H1N1, SARS or MERS. Therefore, strictly speaking, by using only the image without any other clinical information of the patient, these methods cannot confirm whether a patient is infected by covid-19. They can only detect whether the patient has a viral pneumonia. But since during the ongoing covid-19 pandemic most viral pneumonia is caused by SARS-COV-2, detecting these cases quickly in an automated manner can be beneficial. The information provided by the AI methods could be combined with all the other information regarding the patient (clinical history, symptoms, closeness to other covid-19 cases, etc.) to reach a more informed decision[31].

We developed new AI methods and evaluate their performance to judge their usefulness for dealing with the ongoing pandemic. We are by no means the first to attempt detecting covid-19 from chest CT scans and/or X-Ray images using AI methods. Since the beginning of the covid-19 outbreak there has been a rush to test novel methods, and many prior approaches have been proposed[32–47]. We do not claim explicit novelty; but we do offer a careful and fair evaluation on a large set of patients from different centers (especially for X-Ray images). Furthermore, we directly test for the first time if AI methods are capable of separating a viral infection from a bacterial one through the incorporation of tuberculosis patients to the study.

Final X-Ray AI model evaluated here will be incorporated into an online platform (www.quantuscovid19.org). The platform is being developed using Transmural Biotech's long experience and know-how on how to deploy AI methods through Software-as-a-service platforms complying with clinical regulations. Therefore, it will effectively transfer the technology here reported to any professional and/or researcher who desires to make use of it.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

MATERIALS AND METHODS

XRay data

Five different sources were used. Two different sources provided covid-19 infected lungs: *covid-19 Image Data Collection*[48]and *Covid Data Save Lives*[49] (*COVIDDSL).* Then, we also downloaded U.S. National Library of Medicine Tuberculosis Datasets[50] and TB_portals[51], both containing bacterial tuberculosis chest X-ray images.

As stated in the Introduction, the AI methods cannot distinguish covid-19 from other viral infections. However, they should be able to distinguish viral infections from other type of infections, such as bacterial ones. To put these theoretical notions to test, we added bacterial tuberculosis patients to the study. We chose tuberculosis given its importance worldwide (it can actually be considered another pandemic in many countries), availability of images online and also since it can serve as a good baseline for the potential of covid-19 detection, considering that tuberculosis detection from X-Ray images is more established[52–55].

From all these sources we downloaded also all healthy lung images available. Finally, to have more healthy images, we added a random subset of ChestX-ray8[56] dataset (10,000 random healthy lung images). In all cases only Anterior-Posterior or Posterior-Anterior views were used, discarding all lateral images. In the case of covid-19 suspicious patients, we only retained those with a confirmed positive PCR record in the two weeks prior to the image acquisition date.

Table 1 shows the number of patients and X-Ray images from each. A total of 9,841 patients were included in the study, with a covid-19 prevalence of 17.6% (1,733/9,841) and a Tuberculosis prevalence of 8.2% (810/9,841). In average each patient had 2.03 X-Ray images available, for a total of 20,024 images. Demographical information such as age, gender, etc. was not used in this study. The reasons are three-fold: 1) this data was available only from a handful of patients, 2) To avoid overfitting due to dangerous data correlations and 3) to test the power of AI image analysis on its own.

Images from these sources were already provided after a pre-processing done by the dataset authors: clinical image headers were removed for anonymization, and images converted to either as Portable Network Graphics (PNG) or Joint Photographic Experts Group (JPEG) images. The exception was *COVIDDSL*[49], which was provided using original Digital Imaging and Communications in Medicine (DICOM) format images. These images were automatically processed in a similar fashion to the other datasets, removing the image header and converting them to PNG format after adequate conversion to 8-bits. Figure 1 shows some image examples from all sources.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Source	#Covid-19 Patients	#Tuberculosis patients	#Total patients	covid-19 prevalence	# Images	Images/patient
Covid-19 Image Data Collection[48]	265	0	269	98.5%	442	1.64
COVIDDSL[49]	1,468	0	1,468	100%	8088	5.5
U.S. National Library of Medicine Tuberculosis Datasets[50]	0	394	800	0%	800	1
TB_portals[51]	0	416	497	0%	1202	2.41
ChestX-ray8[56]	0	0	6,807	0%	10,000	1.47
TOTAL	1,733	**810**	**9,841**	**17.6%**	**20,024**	**2.03**

Table 1. X-Ray data used for the study.

Power sampling[57] was used to establish the N necessary to validate a model with at least 90% detection rate considering the prevalence of the data (17.6%) and error rates of type I below 5% and type II below 10%. The result was N=548. Although this condition could have been met using only 5% of the data for testing, we performed a 66%/33% training/test split for fairness and to have a large test set. Therefore, 3,528 patients were used for testing and the remaining (6,313) for training. Resulting prevalences on the test set were 23.3% (823/3,528) for covid-19 and 7.6% (269/3,528) for Tuberculosis.

CT Scan data

Data from two different sources was used: the *COVID-CT-Dataset*[58] and the *covid-19 Image Data Collection*[48]. From these, only patients having at least one axial CT image were used. Table 2 shows the number of patients and CT scan images from each source. A total of 363 patients were included in the study, with a covid-19 prevalence of 53% (191/363). In average each patient had 1.92 CT scan images available, for a total of 700 images. Demographical information such as age, gender, etc. was again not used in this case, the reasons being the same as those stated above for X-Ray images.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

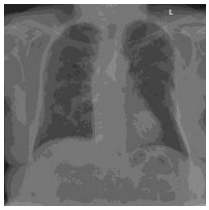
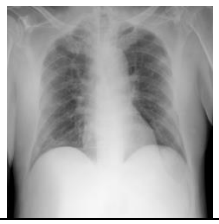
Dataset	Example affected chest XRay image	Example healthy chest XRay image
covid-19 Image Data Collection[48]		
COVIDDSL[49]		-
U.S. National Library of Medicine Tuberculosis Datasets[50]		
TB_portals[51]		
ChestX-ray8[56]	-	

Figure 1. Chest X-Ray Scan image examples.

In both data sources the CT scan images were already processed and directly provided as Portable Network Graphics (PNG) images. Figure 1 shows some image examples from both datasets.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Source	#Covid-19 Patients	#Total patients	covid-19 prevalence	# Images	Images/patient
COVID-CT-Dataset[58]	170	341	50%	663	1.94
covid-19 Image Data Collection[48]	21	22	95%	37	1.68
TOTAL	**191**	**363**	**53%**	**700**	**1.92**

Table 2. CT Scan data used for the study.

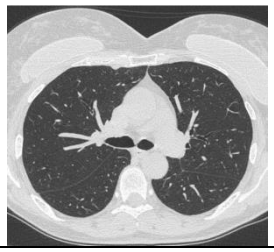
Source	Example covid19 CT scan	Example healthy CT scan
COVID-CT-Dataset[58]		
covid-19 Image Data Collection[48]		

Figure 2. Chest CT Scan image examples.

Power sampling[57] was used to establish the N necessary to validate a model with at least 90% detection rate considering the prevalence of the data (53%) and error rates of type I below 5% and type II below 10%. The result was N=115. To satisfy this condition, a 66%/33% training/testing split was chosen, leaving 121 patients for testing and 242 for training.

AI-based automatic image classification

Several Deep Learning (DL) [59] Convolutional Neural Networks[60] classifiers were benchmarked using Transmural Biotech AI platform. However, for simplicity and given the similarity in the accuracy of many of these models we report only the performance of the best performing model, which used Inception[61] as backbone architecture.

The net was first trained following the original author's guidelines on ImageNet Large Scale Visual Recognition Challenge[62]. Then, the net was fully retrained (allowing changes to the entire network) using our training data, changing the dimension of the last layer to output the probability for each class in each task (3 classes in X-Ray, 2 classes for CT Scans). After training, the net was deployed and applied to the test images, outputting full probability scores

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

for each class. For each patient, if the patient had more than one image, we averaged the output probability scores of the AI method and used this number as final score for the patient.

CNN training details: The model was trained using Transmural Biotech's AI online platform. The Inception classifier was trained following the original author's guidelines using softmax cross-entropy loss and adam optimizer. 10% of the training set was used as validation set. The first training using ImageNet images was performed for 90 epochs with a 5 epoch warm-up. Then, two CNN were trained, one for each task (X-Ray and CT Scans) on the training data for a maximum of 15 epochs, early stopping if loss on validation set was not improved for 5 consecutive epochs. Learning rate was adjusted using a cosine decay starting at 1e-4. Batch size was 64 and weight decay was 0.9. To improve learning, data augmentation was used during training. At each batch, images were randomly flipped, cropped between 0–20%, translated from 0–10 pixels and rotated between [−15; 15] degrees. No data augmentation was used during testing to assure that output was always the same given the same image and to avoid altering the original image.

AI automatic X-Ray lung segmenter

In the case of X-Ray, additionally, since enough training images were available, prior to the classifier, an automatic lung segmenter was also developed. Using a random subset from the classification training patients, (1,100 images), we manually delineated both lungs in the X-Ray image using Transmural Biotech's Graphical User Interface (GUI) online program. Then, we used 1,000 images to train a CNN segmentation network using our online AI platform and the remaining 100 images to test the model. Once trained, the net was capable of segmenting lungs from the 100 test chest X-Ray images with an average Jaccard index of 90.1%. Example automatic lung segmentations are shown in Figure 3.

These automatically generated lung masks were then provided to the Inception classifier both during training and testing. This effectively means that the classifier will use only lung pixels to perform the classification.

X-Ray lung segmenter details: We used a DeepLab model[63] and training was performed in a similar fashion as the Inception classifier, using Transmural Biotech's AI online platform.

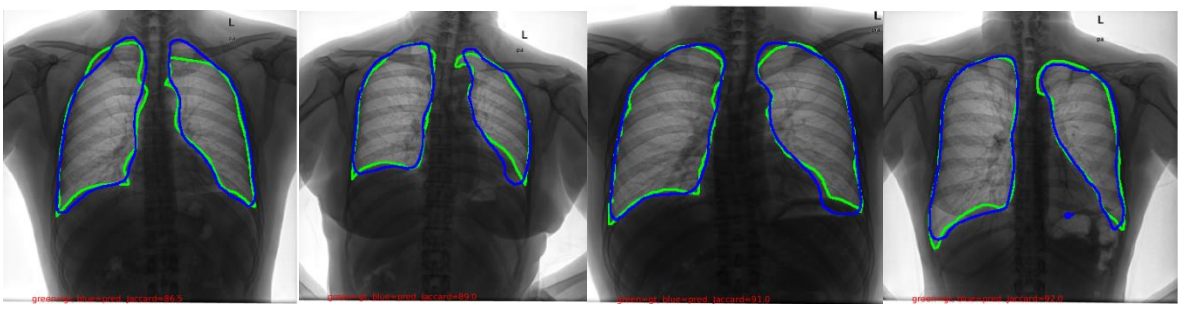


Figure 3. Example automatic X-Ray Lung segmentation masks. Green=manual mask. Blue=predicted mask.

Statistical analysis

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

All methods were trained and applied using Transmural Biotech AI platform. All statistical analysis was then performed using Matlab (Mathworks, USA). All the methods mentioned above (once trained) were applied to all test available images, storing the output estimation (probability score). Whenever several images of the same patient with the same study date were available, the average estimation from all images was used.

The scores were used to draw Receiver Operator Characteristic (ROC) curves and compute full Area Under the Curve (AUC) with their 95% Confidence Intervals. Then, the ROC curves were used to establish the optimal cutoff points as those maximizing accuracy. Detection rate, false positive rate (FPR), positive and negative predictive values (PPV and NPV) and positive and negative likelihood ratios (LR+ and LR-) were calculated with their 95% Confidence Intervals (estimated using bootstrapping on the test set) using these cut-off points.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

RESULTS

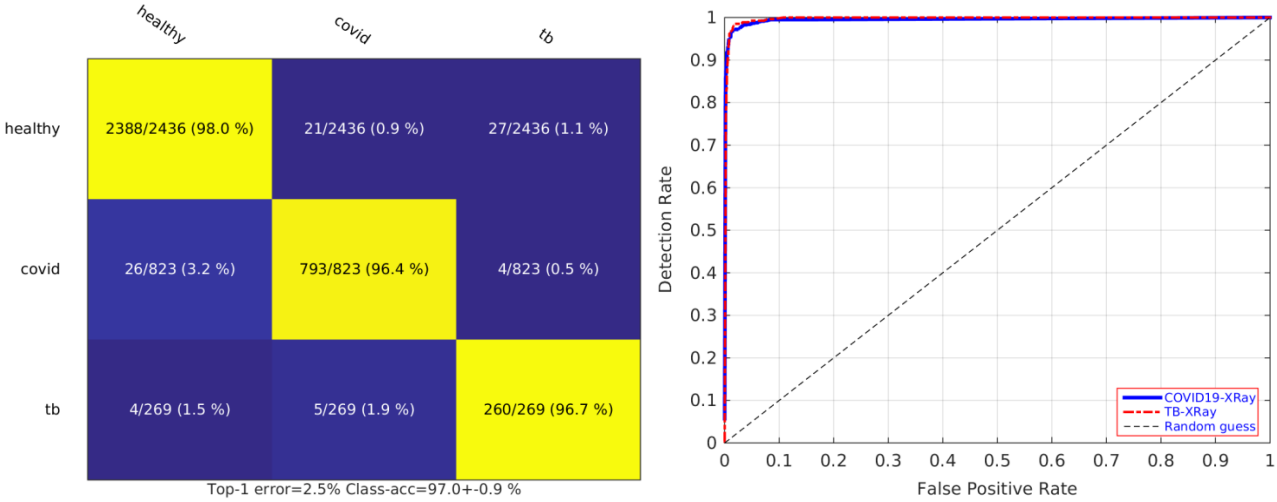
Covid-19 detection from chest X-Ray images

Figure 4 shows the Confusion Matrix and ROC curves for covid-19 and tuberculosis detection from X-Ray images using the testing patients (N=3,528, covid-19 Prevalence 23.3%, Tuberculosis Prevalence 7.6%). Global accuracy on the three classes (healthy, covid, tuberculosis) was 97.5%, with a 97.0+-0.9% average class-accuracy. AUC was 99.5%+-0.1 for covid-19 and 99.7%+-0.1 for tuberculosis.

Table 3 shows the detailed metric scores at the optimal accuracy cut-off points computed from the ROC curve. The CNN was capable of detecting 96.4% of the covid-19 infected patients, while incurring in a 1.0% false positive rate. PPV was 96.8% and NPV 98.9%. This scenario would have implied that of the 823 patients tested that actually had covid-19, the CNN would have successfully detected 793, leaving 30 undetected, and on the other hand it would have falsely flagged 26/3,528 patients as having covid-19.

In comparison, the detection of tuberculosis had very similar results, with a detection rate of 96.7% for a false positive rate of 1.0%. PPV was 89.3% and NPV 99.7%.

Figure 5 shows example correct positive covid-19 detections, with the class-activation maps of the CNN superimposed on top of the image. The CNN appears to be correctly detecting alterations in the lungs, both when only one is affected or both at the same time. Some of the class-activations correspond to manual annotations from the clinician (arrows in first and third images).



(a) *(b)*

Figure 4. Detection results from XRay images (N=3,528 Patients, covid-19 Prevalence 23%, Tuberculosis Prevalence 7.6%). *(a) Full confusion matrix. (b) covid-19* and Tuberculosis ROC curves (blue=covid, red=tuberculosis).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Disease	Accuracy	F1-Score	AUC	Detection Rate	False Positive Rate	PPV	NPV	LR+	LR-
Covid-19	3472/3528 (98.4 % +-0.1 %)	96.6 % +-0.2 %	97.7 % +-0.1 %	793/823 (96.4 % +-0.2 %)	26/2705 (1.0 % +-0.1 %)	793/819 (96.8 % +-0.3 %)	2679/2709 (98.9 % +-0.1 %)	100.2 (+-9.9)	0.0 (+-0.0)
Tuberculosis	3488/3528 (98.9 % +-0.1 %)	92.9 % +-0.5 %	97.9 % +-0.2 %	260/269 (96.7 % +-0.5 %)	31/3259 (1.0 % +-0.1 %)	260/291 (89.3 % +-0.7 %)	3228/3237 (99.7 % +-0.0 %)	101.6 (+-8.4)	0.0 (+-0.0)

Table 3. Chest X-Ray covid-19 & Tuberculosis detection results (N=3528 Patients, covid-19 Prevalence 23%, Tuberculosis Prevalence 7.6%). AUC= Area under the Curve. PPV= Positive Predictive Value. NPV=Negative Predictive Value. LR+= Positive likelihood ratio, LR-=Negative likelihood ratio.

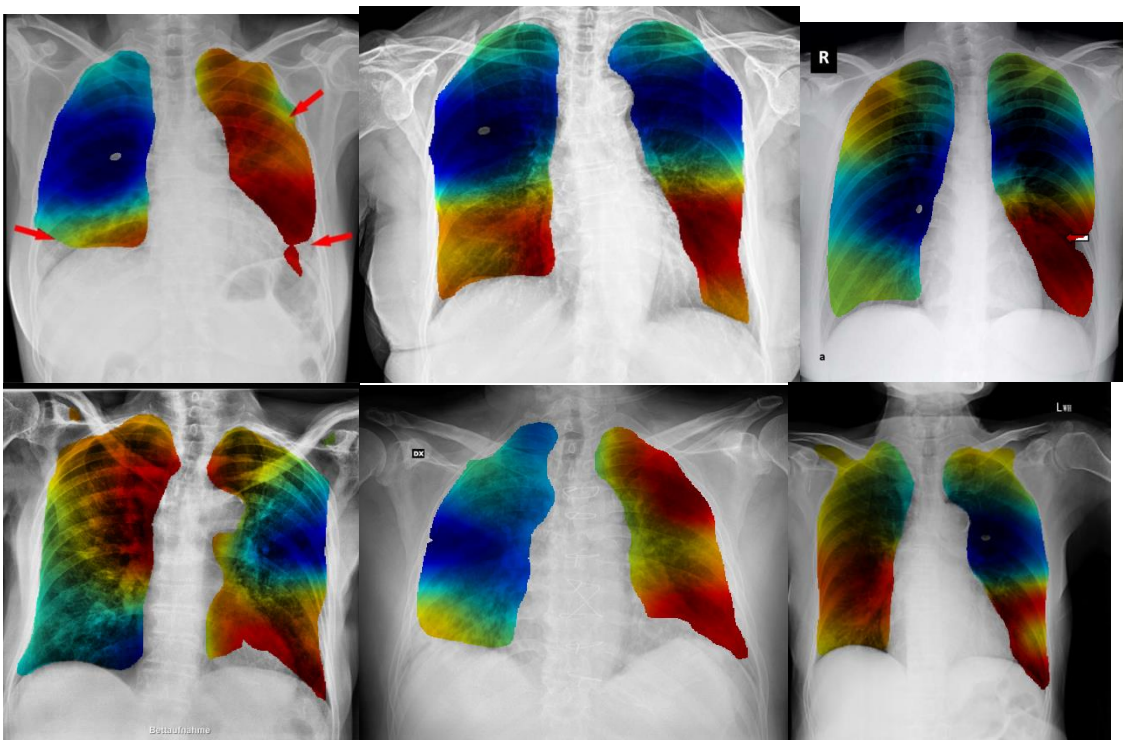


Figure 5. Example correct covid-19 positive predictions from X-Ray images. Images are superimposed with the CNN's Class-activation heatmaps (red= hot, blue=cold), which indicate what parts of the image the CNN model is using to perform the final classification.

Covid-19 detection from chest CT scans

Figure 6 shows the ROC curve for covid-19 detection from CT scans using the testing patients (N=121, covid-19 prevalence 53%). AUC was 90.9%+-0.8%. Table 4 shows the detailed metric scores at the optimal accuracy cut-off point computed from the ROC curve. The CNN was capable of detecting 90.6%+-1.2% of the covid-19 infected patients, while incurring in a

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

24.6%+-1.5% False positive rate. PPV was 80.6%+-2.6% and NPV 87.8%+-1.6%. This scenario would have implied that of the 64 patients tested that actually had covid-19, the CNN would have successfully detected 58, leaving 8 undetected, and on the other hand it would have falsely flagged 14/121 patients as having covid-19.

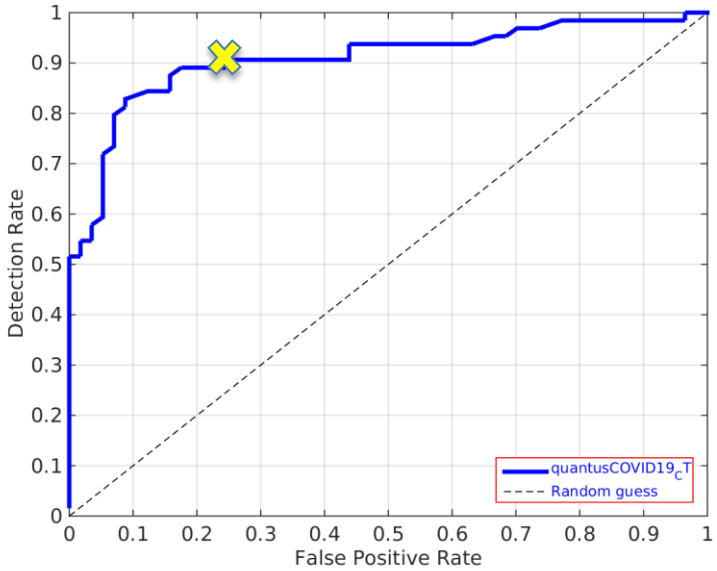


Figure 6. ROC curve for covid-19 detection from CT Scans (N=121 Patients, covid-19 Prevalence 53%). The yellow X marks the optimum cut-off point in terms of Accuracy.

Accuracy	F1-Score	AUC	Detection Rate	False Positive Rate	PPV	NPV	LR+	LR-
101/121 (83.5% +-2.0%)	85.3% +-1.7%	90.9% +-0.8%	58/64 (90.6% +-1.2%)	14/57 (24.6% +-1.5%)	58/72 (80.6% +-2.6%)	43/49 (87.8% +-1.6%)	3.7 (+-0.6)	0.1 (+-0.0)

Table 4. CT-Scan covid-19 detection results (N=121 Patients, covid-19 Prevalence 53%). AUC= Area under the Curve. PPV= Positive Predictive Value. NPV=Negative Predictive Value. LR+= Positive likelihood ratio, LR-=Negative likelihood ratio.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DISCUSSION

Main findings

State-of-the-art AI methods to detect covid-19 in chest X-Ray images and CT scans were developed and evaluated on patients from multiple centers and institutions. Results suggest that covid-19 afflicted patients showing first symptoms can be detected using both medical imaging tools with detection rates of 90% for reasonable false positive rates.

X-Ray images were clearly the more promising of the two. The AI method developed included an automatic lung segmenter and automated detection reached an AUC of 99% on the 3,528 patients tested (823 of which having covid-19). This translated in a detection rate of 96.4% at 1.0% false positive rate and NPV of 98.9%. The AI was also perfectly able to distinguish the viral infection caused by covid-19 from a bacterial infection such as tuberculosis, for which it obtained an AUC of 96%, in line with previously results from tuberculosis large studies[52].

In the case of CT scans, results were also promising although both the number of patients tested and the method's accuracy were lower. The AI method developed was able to reach an AUC of 90.9% on the 121 patients tested, 54 of which had covid-19. This translated into a detection rate of 90.6% at 24.6% false positive rate, with a NPV of 87.8.

Clinical implications

Repeated RT-PCR from nasal swabs, the current gold standard for covid-19 detection, is not perfect (no test ever really is). While it's undoubtedly a very precise test (extremely low false positive rates), prior systematic reviews reported varying detection rates between 63% and 98%[17,64]. The main reasons for these variations are the quality of sampling, the stage of the disease the day of the test, the degree of viral multiplication and the different prevalence of the disease in each center[65,66]. Due to these variations, it has been proposed to combine repeated PCR with clinical history, history of contact with other covid-19 patients and findings from either chest CT scans and/or X-Ray[31].

Results on CT scans are preliminary due to the low number of patients tested. On the other hand, X-Ray results were obtained on a large pool of patients from different centers and therefore are more promising. The new AI method developed extracts information from chest X-Ray automatically, making the process both easier and quicker compared to the evaluation of images by an expert radiologist.

The high detection rate and NPV reported on the 3,528 patients tested makes the X-Ray method suited to help in patient management during the covid-19 pandemic, since it could help reduce the number of false negatives. The method was tested only on symptomatic patients showing first signs of affection in the lungs. The degree of affection or symptoms needed before the method is able to detect covid-19 in a patient remains unclear. However, considering the many reported limitations of PCR discussed above, reducing false negatives in symptomatic patients through a fast and non-invasive procedure could prove useful.

At Transmural Biotech, we have a long experience and know-how of how to deploy AI methods through Software-as-a-service platforms complying with clinical regulations. The X-ray method will be integrated into an online platform (www.quantuscovid19.org), making it available to any professional. The combination of the information provided by this method with all other patient information can undoubtedly help clinicians worldwide during the covid-19

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

pandemic, helping to detect covid-19 infected patients when a PCR test is not available, or helping to detect a fraction of the patients missed by the PCR itself in a first test.

Strengths

This study has several strengths. We have performed the study combining patients from several sources, each of which was in turn constructed using patients from different centers. This means that the data used is inherently multi-center and multiple equipment and operators were responsible for the acquisition of the images, giving more credibility to the results. Furthermore, we have carefully separated patients to train and test methods, and established the number of test patients using power sampling. This resulted in "fair" train/test splits and moderated sized test sets. Moreover, we evaluated methods simulating a real application, where if several images are acquired for each patient all of them are incorporated into the analysis. Another strength worth noting is that we used all data available (as long as clinical outcome was confirmed through a positive PCR), refraining from manual filtering of images and/or patients based on dubious quality criteria that cannot be checked by a third party for correctness. We tested both CT scans and X-Ray images with a common methodology, to establish virtues and defects of each for covid-19 detection. Finally, for the first time, we demonstrated that AI methods seem to be able to distinguish between viral infections and bacterial infections (covid-19 vs tuberculosis).

Limitations

We acknowledge a number of limitations. We were not involved in data acquisition; we trusted the data as it came and weren't able to double-check clinical outcomes of the images. Therefore we cannot be entirely sure of the correctness of the data used. However, these were either public sources widely used by the scientific community, or sources collected in large clinical institutions, so the assumption that data is correct is not entirely senseless. Another limitation is the varying prevalence regarding covid-19 patients found in the datasets used, which undoubtedly difficult a direct comparison of results. Finally, all covid-19 patients evaluated were patients showing symptoms and had lungs affected; it is unclear whether the methods evaluated could be used to detect asymptomatic patients. Larger studies are needed in order to better evaluate the limits of the AI methods proposed.

Conclusions

To conclude, state-of-the-art computational AI methods from both chest CT scans and chest X-Ray images showed promising results. Of the two, X-Ray seemed to be better suited to be used as screening tool during the covid-19 outbreak, providing detection rates >95%. This method could help to manage patients better in case access to PCR testing is not possible or to detect patients missed in a first round of PCR testing. It will be made available through a new platform (www.quantuscovid19.org). These results merit further evaluation and we hope to start larger studies with collaborating centers.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

REFERENCES

1. Hui DS, I Azhar E, Madani TA, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health — The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis*. 2020;91:264-266.
2. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020.
3. Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). No Title. COVID-19 Dashboard.
4. Bikdeli B, Madhavan M V., Jimenez D, et al. COVID-19 and Thrombotic or Thromboembolic Disease: Implications for Prevention, Antithrombotic Therapy, and Follow-Up: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2020;75(23):2950-2973.
5. Cascella M, Rajnik M, Cuomo A, Dulebohn SC, Di Napoli R. *Features, Evaluation and Treatment Coronavirus (COVID-19)*.; 2020.
6. Murthy S, Gomersall CD, Fowler RA. Care for Critically Ill Patients with COVID-19. *JAMA - J Am Med Assoc*. 2020;323(15):1499-1500.
7. Ye Q, Wang B, Mao J. The pathogenesis and treatment of the "Cytokine Storm" in COVID-19. *J Infect*. 2020;80:607-613.
8. Hopkins C. Loss of sense of smell as marker of COVID-19 infection. *Ear, Nose Throat Surg body United Kingdom*.
9. Grant MC, Geoghegan L, Arbyn M, et al. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 148 studies from 9 countries. Hirst JA, ed. *PLoS One*. 2020;15(6):e0234765.
10. Velavan TP, Meyer CG. The COVID-19 epidemic. *Trop Med Int Heal*. 2020;25(3):278-280.
11. Emery SL, Erdman DD, Bowen MD, et al. Real-Time Reverse Transcription-Polymerase Chain Reaction Assay for SARS-associated Coronavirus. *Emerg Infect Dis*. 2004;10(2):311-316.
12. Kadri K. Polymerase Chain Reaction (PCR): Principle and Applications. In: Nagpal ML, Boldura O-M, Baltă C, Enany S, eds. *Synthetic Biology*. IntechOpen; 2020.
13. Day M. Covid-19: identifying and isolating asymptomatic people helped eliminate virus in Italian village. *BMJ*. 2020;368:m1165.
14. Li Y, Yao L, Li J, et al. Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID-19. *J Med Virol*. 2020;92(7):903-908.
15. Xiao AT, Tong YX, Zhang S. False-negative of RT-PCR and prolonged nucleic acid conversion in COVID-19: Rather than recurrence. *J Med Virol*. Published online 2020.
16. Zou L, Ruan F, Huang M, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N Engl J Med*. 2020;382(12):1177-1179.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

17. Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA - J Am Med Assoc*. 2020;323(18):1843-1844.

18. Hope MD, Raptis CA, Shah A, Hammer MM, Henry TS, behalf of six signatories O. A role for CT in COVID-19? What data really tell us so far. *http://www thelancet com/article/S0140673620307285/pdf*. Published online 2020.

19. Vermeiren C, Marchand-Senécal X, Sheldrake E, et al. Comparison of Copan Eswab and FLOQswab for COVID-19 PCR diagnosis: working around a Downloaded from. *J Clin Microbiol*. Published online 2020.

20. Phua J, Weng L, Ling L, et al. Intensive care management of coronavirus disease 2019 (COVID-19): challenges and recommendations. *Lancet Respir Med*. 2020;8(5):506-517.

21. Grasselli G, Pesenti A, Cecconi M. Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy. *JAMA*. 2020;323(16):1545.

22. Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir*. 2020;8:475-481.

23. Xie J, Tong Z, Guan X, Du B, Qiu H, Slutsky AS. Critical care crisis and some recommendations during the COVID-19 epidemic in China. *Intensive Care Med*. 2020;46(5):837-840.

24. Qiu H, Tong Z, Ma P, et al. Intensive care during the coronavirus epidemic. *Intensive Care Med*. 2020;46(4):576-578.

25. Borghesi A, Medica RM-L radiologia, 2020 U. COVID-19 outbreak in Italy: experimental chest X-ray scoring system for quantifying and monitoring disease progression. *Springer*.

26. Wong HYF, Lam HYS, Fong AHT, et al. Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients. *Radiology*. Published online 2019:201160.

27. Bai HX, Hsieh B, Xiong Z, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. *Radiology*. Published online 2020:200823.

28. 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*. Published online 2020:200642.

29. Bernheim A, Mei X, Huang M, et al. Chest CT findings in coronavirus disease 2019 (COVID-19): Relationship to duration of infection. *Radiology*. 2020;295(3):685-691.

30. Li Y, Xia Li LY. Role of Chest CT in Diagnosis and Management. *AJR*. 2020;214(6):1280-1286.

31. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 test result. *BMJ*. 2020;369.

32. Hemdan EE-D, Shouman MA, Karar ME. COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images. *arXiv:200311055*. Published online March 24, 2020.

33. Li L, Qin L, Xu Z, et al. Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT. *Radiology*. Published online March 19, 2020:200905.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

34. Narin A, Kaya C, Pamuk Z. *Automatic Detection of Coronavirus Disease (COVID-19) Using X-Ray Images and Deep Convolutional Neural Networks*.

35. Wang L, Wong A. COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images. *arXiv:200309871*. Published online 2020.

36. Tartaglione E, Barbano CA, Berzovini C, Calandri M, Grangetto M. Unveiling COVID-19 from Chest X-ray with deep learning: a hurdles race with small data. *arXiv:200405405*. Published online 2020.

37. Mei X, Lee H, Diao K, et al. Artificial intelligence–enabled rapid diagnosis of patients with COVID-19. *nature.com*.

38. Abbas A, Abdelsamea MM, Gaber MM. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *arXiv:200313815*. Published online 2020.

39. Apostolopoulos I, In TM-P and ES, 2020 U. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Springer*.

40. Wang S, Kang B, Ma J, et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *medrxiv.org*.

41. Narin A, Kaya C, Pamuk Z. Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. *arXiv:200310849*. Published online 2020.

42. Huang L, Han R, Ai T, et al. Serial Quantitative Chest CT Assessment of COVID-19: Deep-Learning Approach. *Radiol Cardiothorac Imaging*. 2020;2(2):e200075.

43. Gozes O, Ayan Frid-Adar M', Greenspan H, et al. *Title: Rapid AI Development Cycle for the Coronavirus (COVID-19) Pandemic: Initial Results for Automated Detection & Patient Monitoring Using Deep Learning CT Image Analysis Authors*.

44. Zheng C, Deng X, Fu Q, et al. Deep Learning-based Detection for COVID-19 from Chest CT using Weak Label. *medrxiv.org*.

45. Butt C, Gill J, Chun D, Intelligence BB-A, 2020 U. Deep learning system to screen coronavirus disease 2019 pneumonia. *Appl Intell*. 2020;1.

46. Zhang J, Xie Y, Li Y, Shen C, Xia Y. COVID-19 Screening on Chest X-ray Images Using Deep Learning based Anomaly Detection. *arXiv:200312338*. Published online 2020.

47. Shan F, Gao Y, Wang J, et al. *Lung Infection Quantification of COVID-19 in CT Images with Deep Learning Author List*.

48. Cohen JP, Morrison P, Dao L. COVID-19 Image Data Collection. *arXiv:200311597*. Published online March 25, 2020.

49. HM Hospitales. Covid Data Save Lives.

50. Jaeger S, Candemir S, Antani S, Wáng Y-XJ, Lu P-X, Thoma G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant Imaging Med Surg*. 2014;4(6):475-477.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

51. TB Portals | TB Central.

52. Murphy K, Habib SS, Zaidi SMA, et al. Computer aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system. *Sci Rep*. 2020;10(1).

53. Nyein Naing WY, Z. Htike Z. Advances in Automatic Tuberculosis Detection in Chest X-Ray Images. *Signal Image Process An Int J*. 2014;5(6):41-53.

54. van Cleeff MRA, Kivihya-Ndugga LE, Meme H, Odhiambo JA, Klatser PR. The role and performance of chest X-ray for the diagnosis of tuberculosis: A cost-effective analysis in Nairobi, Kenya. *BMC Infect Dis*. 2005;5(1):111.

55. Pasa F, Golkov V, Pfeiffer F, Cremers D, Pfeiffer D. Efficient Deep Network Architectures for Fast Chest X-Ray Tuberculosis Screening and Visualization. *Sci Rep*. 2019;9(1):1-9.

56. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. *Proc - 30th IEEE Conf Comput Vis Pattern Recognition, CVPR 2017*. 2017;2017-Janua:3462-3471.

57. Buderer NMF. Statistical Methodology: I. Incorporating the Prevalence of Disease into the Sample Size Calculation for Sensitivity and Specificity. *Acad Emerg Med*. 1996;3(9):895-900.

58. Yang X, San Diego U, San Diego Jinyu Zhao U, San Diego Yichen Zhang U, San Diego Shanghang Zhang U, Xie P. *COVID-CT-Dataset: A CT Image Dataset about COVID-19*.

59. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444.

60. Bengio Y. *Convolutional Networks for Images, Speech, and Time-Series Unsupervised Learning of Speech Representations View Project Parsing View Project*.; 1997.

61. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol 07-12-June-2015. IEEE Computer Society; 2015:1-9.

62. Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. *Int J Comput Vis*. 2014;115(3):211-252.

63. Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans Pattern Anal Mach Intell*. 2018;40(4):834-848.

64. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, et al. *False-Negative Results of Initial RT-PCR Assays for COVID-19: A Systematic Review*. Cold Spring Harbor Laboratory Press; 2020.

65. Wölfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature*. 2020;581(7809):465-469.

66. Sethuraman N, Jeremiah SS, Ryo A. Interpreting Diagnostic Tests for SARS-CoV-2. *JAMA - J Am Med Assoc*. 2020;323(22):2249-2251.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.16.20155093.this version posted July 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Funding: This work in its entirety was supported by Transmural Biotech SL.

Competing interests: All authors are Transmural Biotech employees

Data and materials availability: All data used comes from external open research sources available online. The rest of data is available in the main text. The final X-Ray model will be provided online at www.quantuscovid19.org