Using artificial intelligence to risk stratify COVID-19 patients based on chest X-ray findings

Diego A. Hipolito Canario [a],[*], Eric Fromke [a], Matthew A. Patetta [a], Mohamed T. Eltilib [a], Juan P. Reyes-Gonzalez [b], Georgina Cornelio Rodriguez [b], Valeria A. Fusco Cornejo [c], Seymour Duncker [c], Jessica K. Stewart [d]

[a] *UNC School of Medicine, University of North Carolina at Chapel Hill, Bondurant Hall, CB 9535, Chapel Hill, NC, 27599-3280, United States*
[b] *Department of Radiology, Angeles del Pedregal Hospital, Camino de Sta, Teresa 1055-S, Héroes de Padierna, La Magdalena Contreras, 10700, Ciudad de México, Mexico*
[c] *Mindscale, 800 W El Camino Real Suite 180 Mountain View, CA, 94040, United States*
[d] *Division of Interventional Radiology, Department of Radiology, David Geffen School of Medicine, University of California at Los Angeles. 757 Westwood Plaza, Suite 2125, Los Angeles, CA, 90095, United States*

ARTICLE INFO

ABSTRACT

Background: Deep learning-based radiological image analysis could facilitate use of chest x-rays as a triaging tool for COVID-19 diagnosis in resource-limited settings. This study sought to determine whether a modified commercially available deep learning algorithm (M-qXR) could risk stratify patients with suspected COVID-19 infections.
Methods: A dual track clinical validation study was designed to assess the clinical accuracy of M-qXR. The algorithm evaluated all Chest-X-rays (CXRs) performed during the study period for abnormal findings and assigned a COVID-19 risk score. Four independent radiologists served as radiological ground truth. The M-qXR algorithm output was compared against radiological ground truth and summary statistics for prediction accuracy were calculated. In addition, patients who underwent both PCR testing and CXR for suspected COVID-19 infection were included in a co-occurrence matrix to assess the sensitivity and specificity of the M-qXR algorithm.
Results: 625 CXRs were included in the clinical validation study. 98% of total interpretations made by M-qXR agreed with ground truth (p = 0.25). M-qXR correctly identified the presence or absence of pulmonary opacities in 94% of CXR interpretations. M-qXR's sensitivity, specificity, PPV, and NPV for detecting pulmonary opacities were 94%, 95%, 99%, and 88% respectively. M-qXR correctly identified the presence or absence of pulmonary consolidation in 88% of CXR interpretations (p = 0.48). M-qXR's sensitivity, specificity, PPV, and NPV for detecting pulmonary consolidation were 91%, 84%, 89%, and 86% respectively. Furthermore, 113 PCR-confirmed COVID-19 cases were used to create a co-occurrence matrix between M-qXR's COVID-19 risk score and COVID-19 PCR test results. The PPV and NPV of a medium to high COVID-19 risk score assigned by M-qXR yielding a positive COVID-19 PCR test result was estimated to be 89.7% and 80.4% respectively.
Conclusion: M-qXR was found to have comparable accuracy to radiological ground truth in detecting radiographic abnormalities on CXR suggestive of COVID-19.

1. Introduction

The novel coronavirus disease 2019 (COVID-19) was first reported in Wuhan, China and has since spread throughout the world, causing the World Health Organization (WHO) to declare the virus a global pandemic on March 11, 2020 [1]. Despite extensive research, clinical,

and governmental efforts, the prevalence of this virus continued to escalate globally as of this writing [2]. Due to the prevalence and highly infectious nature of this disease, there is a need to accurately and quickly identify patients, particularly in resource-limited regions [1].

Although initial efforts in mainland China first utilized chest computed tomography (CT)-scans to identify patients with likely

COVID-19 infection, recent studies indicate that chest x-ray (CXR) may be preferable for this purpose due to its widespread availability and lower cross-infectivity risks compared to CT scanners [3,4]. Common radiographic CXR findings in patients infected with COVID-19 include bilateral pulmonary consolidations predominantly in the lower lobes and in the periphery of the lungs [5–7].

When available, real time reverse transcription polymerase chain reaction (RT-PCR) testing for SARS-Cov-2 nucleic acid from the novel coronavirus is currently the preferred diagnostic testing, with a significantly higher sensitivity (91%) for detecting COVID-19 infection when compared to CXR (69%) [5,8]. However, in resource limited settings where RT-PCR testing is not readily available, the Fleischner Society recommends evaluation with imaging to triage suspected COVID-19 patients [8]. CXR has been shown to be a useful tool in triaging patients with suspected COVID-19 infection who could benefit most from early intervention and hospitalization [9,10].

Since many of the regions that would potentially utilize CXR to classify possible COVID-19 patients are already resource-limited, there is an opportunity for artificial intelligence (AI) to unburden radiologist's workload by quickly identifying potentially infected patients [11]. While a few preliminary studies have investigated the use of AI, specifically utilizing a deep learning model, to interpret structural abnormalities on CT scans or CXR in order to improve COVID-19 detection, the use of AI in a resource-limited region for medical triage has not yet been investigated [12–14]. The objective of this study is to determine the efficacy of a newly designed algorithm (M-qXR) in stratifying patients with suspected COVID-19 infections in a resource-limited setting.

2. Algorithm development

qXR, a clinically validated, proprietary deep learning (DL) algorithm (Qure.ai, Mumbai, India) trained to identify abnormal imaging findings suggestive of tuberculosis (TB), is currently deployed in over 28 countries as an assistive tool for radiologists [15–17]. This algorithm was developed using a dataset of 2.5 million CXRs (CXRs included radiographic findings suggestive of infectious etiologies for pneumonia, tuberculosis, and other pulmonary pathologies) and demonstrated an area under the curve (AUC) as a primary accuracy measure of 0·92 (CI 0·91-0·94) for detection of abnormal scans [15]. This DL algorithm was modified (M-qXR) to detect radiographic abnormalities that are indicative, such as opacities and consolidation with bilateral and peripheral distribution, and contra-indicative, such as hilar enlargement, discrete pulmonary nodule, calcification, cavity, and pleural effusion, of COVID-19 to assign a COVID-19 risk score. M-qXR uses the same model architecture as qXR, which is a convolutional neural network (CNN).

The algorithm uses features derived from qXR generated segmentation maps of radiological findings clinically relevant to COVID-19 to generate COVID-19 risk scores. The specific training steps and post-processing algorithm for M-qXR are propriety; however, the architectures that form the basic blocks in the systems that detect individual abnormalities in M-qXR are similar to those of q-XR and are versions of resnets with squeeze-excitation modules [18,19].

In summary, CNNs are algorithms used to analyze data as a whole and further categorize this data into sections based on similar characteristics. The use of CNNs has become popular in medical image classification for disease diagnosis due to high accuracy rates [20–23]. These algorithms follow a general architecture, consisting of an input image, convolution layers, pooling layers, and fully connected layers that create an output (Fig. 1). An input image is analyzed using several filters, referred to as channels, via convolution operations to form convolution layers. A down-sampling operation occurs, yielding further data extraction and a pooling layer. Then, each input is mapped fully to form a connected layer [21,24]. CNNs combine spatial and channel-wise information to analyze an image by placing varying weights on filters based on the desired output. More recently, Squeeze-Excitation Networks (SENets), improve the efficiency of CNNs by reducing the computational power needed to analyze input images. These algorithms can evaluate channel-interdependencies when analyzing input images and decipher which channels should be assigned greater weight for the overall interpretation of an image [19].

2.1. Evaluation of M-qXR

An independent test-set of 11,479 CXRs was created prior to deployment to evaluate the modified algorithm (M-qXR) as a tool to risk stratify patients based on CXR imaging findings. The test-set consisted of 515 positive CXRs and 10,964 negative CXRs. Positive CXRs were defined as any CXRs acquired from a patient with a RT-PCR positive test, the gold-standard of care for diagnosis of novel coronavirus-19 based on the current World Health Organization (WHO) guidelines. Positive CXRs were acquired after PCR-diagnosis was obtained. The algorithm was trained on a set of 300 CXRs from COVID-19 positive patients from India and Europe. Negative CXRs were defined as those acquired before the first case of COVID-19 was discovered on December 31, 2019 [25]. The negative CXR set used was sampled from a distinct set of 250,000 CXRs and was comprised of 1609 cases of bilateral pulmonary opacities and 547 cases of pulmonary consolidations, where the final diagnosis was not COVID-19. This set also included 355 CXRs with abnormalities not related to pulmonary opacities. The respective sizes of the negative and positive CXR sets were created to match the available prevalence of the
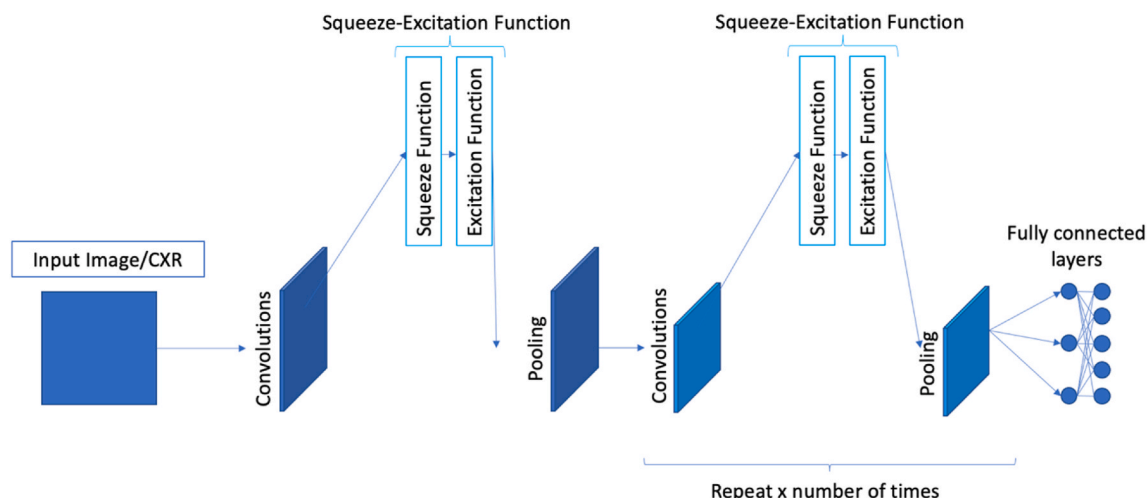

Fig. 1. General CNN architecture with Squeeze-Excitation Network structure.

disease as reported in the literature at the time of algorithm development during the first trimester of 2020, allowing for evaluation of the algorithm's ability to detect non-COVID-19 pulmonary opacities and other findings [26]. M-qXR was able to rule out COVID-19 and suggest alternative possible etiologies based on CXR imaging findings. M-qXR was observed to have an area under the curve (AUC) of 0.9 (95% CI: 0.88–0.92) on this test set. At the operating threshold, M-qXR was observed to have a sensitivity of 0.912 (95% CI: 0.88–0.93) and specificity of 0.775 (95% CI: 0.77–0.78) to detect imaging findings indicative of COVID-19 (Table 1).

3. Methods

3.1. Study Design

A dual track clinical validation study was designed to compare the accuracy of M-qXR in detecting CXR imaging findings for COVID-19 to the formal interpretations made by radiologists. In collaboration with Hospital Angeles del Pedregal, the algorithm was deployed at a private hospital in Mexico City, Mexico, during the initial stages of the COVID-19 pandemic in the country (April to May 2020) as a tool to stratify at-risk patients to receive further testing to help contain the early spread of the virus. Of note, the study period was eight weeks. Patient consent was waived as the data used for algorithm development was retrospective and de-identified and data processes were carried out in a controlled environment compliant with all Indian IT laws and the Health Insurance Portability and Accountability Act (HIPAA).The dual track consisted of:

3.2. Track 1: M-qXR

M-qXR evaluated all CXRs performed at Hospital Angeles del Pedregal during the study period for abnormal findings consistent with COVID-19. The algorithm was configured to process all CXRs performed, regardless of reason for hospital visit, and evaluate for the presence of radiographic abnormalities commonly found with COVID-19 infection. The presence of pulmonary opacities and consolidations with a bilateral and peripheral distribution were considered indicative of likely COVID-19 infection. The presence of hilar enlargement, a discrete pulmonary nodule, calcifications, cavitation, and pleural effusions were contra-indicative of COVID-19. M-qXR risk stratified CXRs into no-risk, low-risk, medium-risk, and high-risk for COVID-19 based on the presence or absence of these radiographic abnormalities (Fig. 2). The thresholds for low-risk, medium-risk, and high-risk COVID-19 scores were established by Qure.ai and were provided by M-qXR as a validated threshold. However, the marginal percent probability that the radiographic findings in each CXR were suggestive of COVID-19 were also provided by M-qXR.

3.3. Track 2: ground truth

Four independent radiologists served as ground truth (ideal expected result used to calculate DL algorithm's accuracy) and evaluated all CXRs during the study period for abnormal findings consistent with COVID-
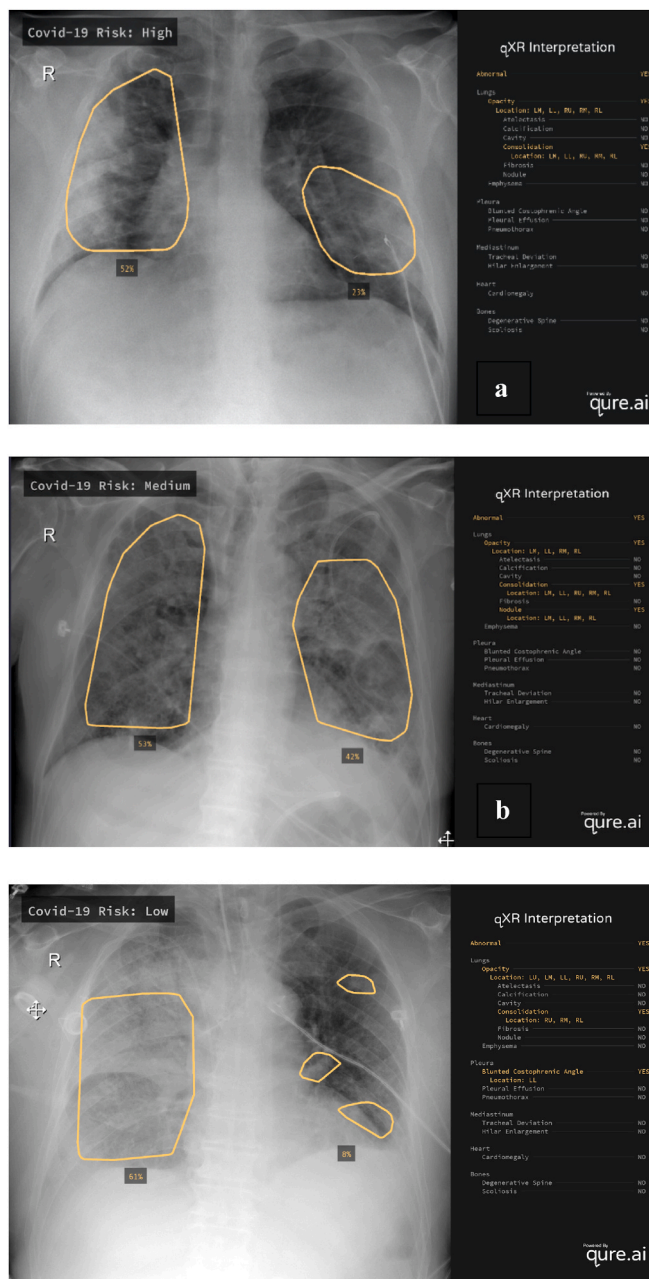

Fig. 2. Examples of chest radiographs evaluated by the M-qXR algorithm. M-qXR assigned a high (a), medium (b), and low (c) COVID-19 risk score based on the presence of radiographic abnormalities on the chest radiographs presented.

19. Ground truth was established by majority consensus interpretations by the four radiologists, who had an average of 22 years (2, 5, 30, 50 years respectively) of experience. The radiologists were

Table 1
M-qXR's metrics for detecting signs of COVID-19 on CXR.

	[a]Sensitivity (%)	[a]Specificity (%)	[a]AUC	PPV (%)		NPV (%)	
M-qXR	91.2 [b](88.0–93.0)	77.5 [b](77.0–78.0)	0.9 [b](0.88–0.92)	[c]69.9	[d]89.7	[c]41.6	[d]80.4

[a] Metrics calculated at the operating threshold during algorithm development.
[b] 95% Confidence Intervals (CI).
[c] PPV and NPV of a medium-high COVID-19 risk score yielding a positive COVID-19 PCR result using the estimated prevalence of COVID-19 among the study patients.
[d] PPV and NPV for a medium-high COVID-19 risk score yielding a positive COVID-19 PCR result using the algorithm's observed sensitivity and specificity at the operating threshold and disease prevalence cited in literature.

provided with the identification number and study date for each CXR. Radiologists were blinded to the patient's reason for hospital visit and did not have access to the patient's medical chart during the study period. The radiologists conducted a preliminary review of all CXRs to classify them as either normal or abnormal. A CXR was marked as normal when majority of radiologists classified the CXR as normal. The same criteria were used to classify CXRs as abnormal. The presence of any of the imaging findings concerning for COVID-19 classified the CXR as abnormal.

3.4. COVID-19 risk score and PCR test result Co-occurrence matrix

In parallel to the dual-track validation study, patients who sought treatment at Hospital Angeles del Pedregal also underwent PCR testing for definitive COVID-19 diagnosis. Testing was offered after clinical evaluation by a physician. The decision to perform PCR-testing was made at the physicians' discretion. A subset of patients who underwent CXR also received PCR-testing. Patients who underwent both PCR testing and CXR for suspected COVID-19 infection were included in a co-occurrence matrix to assess the sensitivity and specificity of the M-qXR algorithm.

3.5. Statistical analysis

The M-qXR algorithm output was compared against radiological ground truth and summary statistics for prediction accuracy were calculated. For each radiographic abnormality, paired categorizations (present/not present) were recorded as determined by M-qXR and by consensus of radiologists. McNemar's test was used to estimate the chance of seeing a given difference in these paired categorizations if there was no underlying difference between groups. Furthermore, the risk score assigned by M-qXR was computed using a post-processing algorithm that combined the model outputs for the above-mentioned imaging findings that are either suggestive or contra-indicative of COVID-19. To simulate triage, only the first scan a patient received during hospitalization was included when calculating M-qXR COVID-19 risk score's prediction accuracy. Statistical measures for accuracy such as sensitivity, specificity, PPV, and NPV were calculated.

4. Results

4.1. M-qXR clinical validation

A total of 722 CXRs were processed by M-qXR during the study period. Of the 722 CXRs, 647 were interpreted by both the M-qXR algorithm and by the radiologists. Seventy-three CXRs were initially excluded from the study because they were not properly uploaded into the picture archiving and communication system (PACS); therefore, a final interpretation for these CXRs was not made. Twenty-two additional CXRs were excluded from the study because majority consensus was not achieved by the radiologists and ground truth could not be established. As a result, 625 total CXRs were included in the clinical validation study. M-qXR classified 524 CXRs as abnormal and proceeded to assign a COVID-19 risk score (Fig. 3). In terms of clinical accuracy, 98% of total interpretations made by M-qXR agreed with ground truth. McNemar's test yielded a p = 0.25 indicating no statistically significant difference between M-qXR's ability to detect a radiographic abnormality concerning for COVID-19 on CXR when compared to ground truth.

M-qXR correctly identified the presence or absence of pulmonary opacities in 94% of CXR interpretations. However, McNemar's test yielded a $p < 0.05$ (p = 0.0001) indicating a statistically significant difference between M-qXR's ability to detect the presence of pulmonary opacity on CXR when compared to ground truth. M-qXR's sensitivity and specificity for detecting pulmonary opacities was 94% and 95% respectively. The calculated PPV and NPV for M-qXR's ability to correctly report pulmonary opacity was 99% and 81% respectively. M-qXR correctly identified the presence or absence of pulmonary consolidation in 88% of CXR interpretations. No significant differences were seen in M-qXR's ability to detect the presence of pulmonary consolidation on CXR when compared to ground truth, $p > 0.05$ (p = 0.48). M-qXR's sensitivity and specificity for detecting pulmonary consolidation was 91% and 84% respectively. The calculated PPV and NPV for M-qXR's ability to correctly report pulmonary consolidation was 89% and 86% respectively.

4.2. COVID-19 risk score and PCR test result Co-occurrence matrix

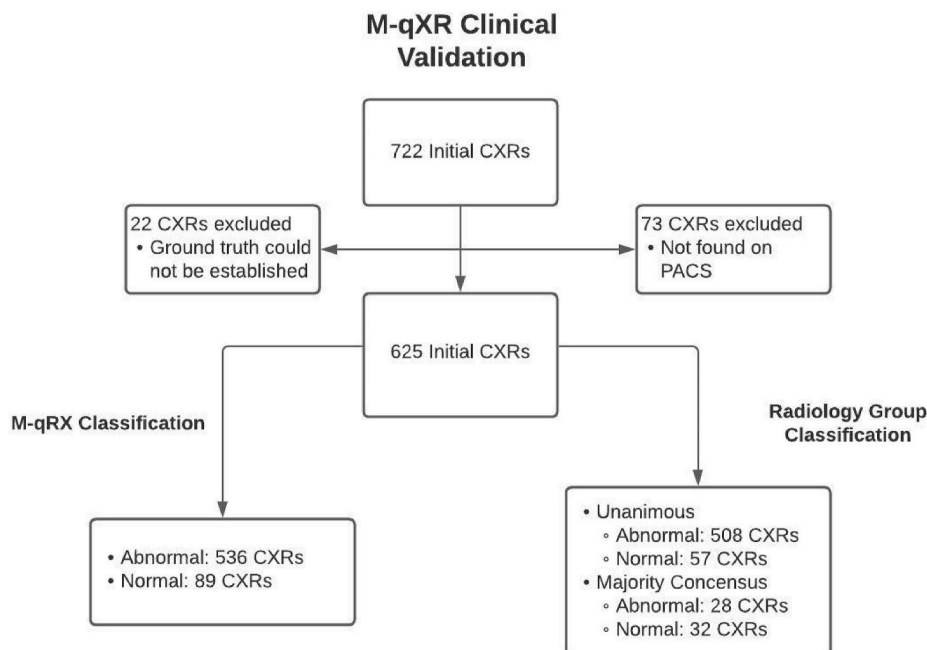A total of 1083 patients underwent PCR testing for definitive COVID-

Fig. 3. Study Design: M-qXR Clinical Validation. PACS: Picture archiving and communication system; CXRs: Chest X-rays.

19 diagnosis during the 8-week study period. Of the 1083 patients, the majority (n=962) underwent PCR testing alone without CXR. A small subset of patients (n=121) underwent both PCR testing and CXR for COVID-19 diagnosis. These CXRs were used to clinically validate M-qXR as stated previously. The CXRs corresponding to these patients were automatically processed by M-qXR and assigned a COVID-19 risk score, which was correlated with a PCR result. M-qXR COVID-19 low risk scores (n=8) were excluded, as they could not be correlated to either positive or negative PCR test results. In total, 113 COVID-19 validated cases were used to create a co-occurrence matrix for the COVID-19 risk score produced by M-qXR and PCR testing (Fig. 4).

58.4% of all PCR results (n=66) were positive for COVID-19. Patients with medium-high COVID-19 risk scores accounted for 86.6% of all positive PCR test results. In contrast, 41.6% of PCR tests (n=47) yielded a negative result for COVID-19. Patients with a no-risk COVID-19 score accounted for 50% of all negative PCR test results. The calculated PPV and NPV of a medium-high COVID-19 risk score yielding a positive COVID-19 PCR result was estimated to be 69.9% and 41.6% respectively using the estimated prevalence of COVID-19 among the study patients. Conversely, the calculated PPV and NPV for a COVID-19 medium-high risk score using the algorithm's observed sensitivity and specificity (91.2% and 77.5% respectively) at the operating threshold was 89.7% and 80.4% respectively (Table 1).

5. Discussion

5.1. Clinical significance

Clinical validation of M-qXR as an assistive tool to risk stratify patients based on CXR imaging findings concerning for COVID-19 was achieved by comparing the algorithm output against radiological ground truth. 98% of interpretations made by M-qXR coincided with ground truth and no significant differences were seen in M-qXR's ability to flag overall radiographic abnormalities when compared to the radiologists (p < 0.05). M-qXR exhibited a moderate to high sensitivity and specificity for radiographic findings suggestive of COVID-19. Among patients with PCR positive COVID-19 test results, 64.4% of the corresponding CXRs were classified as medium-high risk by M-qXR. The PPV and NPV of a COVID-19 risk score yielding a positive PCR test result were 69.9% and 41.6% respectively in our study population. We believe the low specificity was due to selection bias as it may be that only patients with symptoms highly concerning for a respiratory infection received both a CXR and PCR. Using M-qXR's observed sensitivity and specificity (91.2% and 77.5% respectively) at the operating threshold, the PPV and NPV were found to be 89.7% and 80.4%.

We believe that M-qXR could serve as a radiology decision tool to guide management of patients who are deemed at medium to high risk for COVID-19 and may have poor outcomes. In the setting of limited testing, Mexico employed an epidemiological surveillance method known as the Sentinel Model to assess disease burden and allocate resources. By risk stratifying patients using M-qXR's COVID-19 risk scores, limited resources, such as PCR-testing, can be allocated to higher risk groups to help diagnose and treat patients with high likelihood of disease transmission. It is estimated that two-thirds of the world's population lacks access to medical imaging [27]. Patients in these radiological scarce zones tend to have higher mortality and morbidity due to poor access to health services. We believe that M-qXR can help increase access to care in these areas and optimize a radiologist's workflow in a critically resource-limited environment by providing recommendations to help guide medical management.

DL algorithms have recently been used to screen for tuberculosis and other diseases, especially in resource-limited settings, to aid in the interpretation of CXR findings. More recently, several DL algorithms have been developed to screen and risk stratify for COVID-19. However, many of these algorithms lack robust clinical validation. One critical appraisal of dozens of AI algorithms and predictive models to screen and stratify COVID-19 patients concluded that all of the models were at a high risk of bias, via a PROBAST assessment [28]. The risk of bias was
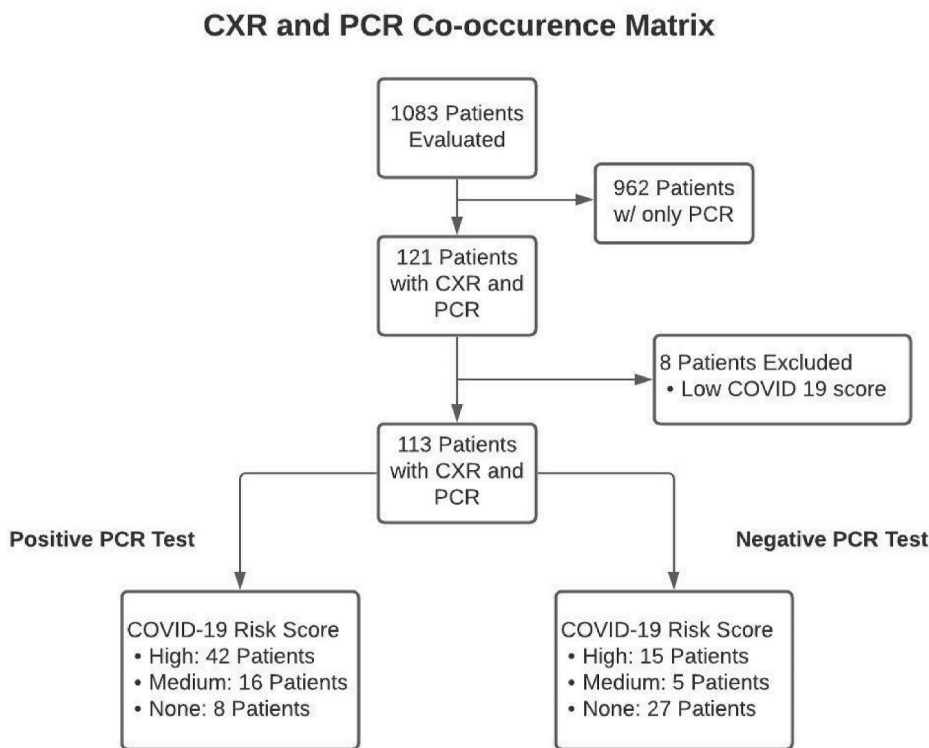

Fig. 4. COVID-19 Risk Score and PCR Test Result Co-occurrence Matrix: Study Design
PCR: Real time reverse transcription polymerase chain reaction (RT-PCR) testing for SARS-Cov-2 nucleic acid from the novel 2019 coronavirus strain; CXR: Chest X-rays.

due to underreporting the use of control patients or the target population, poor description of how regions of interest were assessed, lack of reader consensus to establish ground truth, and a lack of scientific rigor [29]. In addition, many of the prognostic outcomes were not well reported because of lack of long-term follow up. Many of these studies also need to be performed using a large-scale population to improve external validity. The findings of these studies cannot be generalized due to lack of adequate sample size, despite high reported sensitivities and specificities. A study by Chowdhury et al., however, acknowledged the importance of a large dataset and tried to teach an algorithm to read COVID images from a large set of COVID-19 positive CXRs. They were able to achieve high classification accuracy, precision, sensitivity, and specificity in diagnosing COVID-19 [29].

5.2. Limitations of study

The present study had several limitations. The number of CXRs decreased after the 6th week of the study period, despite increasing daily COVID-19 cases in Mexico City, Mexico. To track the spread of COVID-19, the Mexican Health Department redirected all COVID-19 positive patients to government owned hospitals, which significantly decreased the number of CXRs analyzed by the algorithm during the last two weeks of the study. Furthermore, the data that was used to test the performance of the algorithm originated from one institution, and thus may not be an accurate representation of COVID-19's appearance on chest radiographs in other settings. Another limitation is that chest radiographs may be normal early in the disease course of COVID-19. One study found that 31% of patients diagnosed with COVID-19 did not demonstrate any chest radiograph abnormalities [5]. Thus, chest radiographs may not be the most effective way to risk stratify COVID-19 patients who are early in their disease process. Furthermore, to simulate triage, only the first CXR patients received during their hospitalization were used to calculate statistical measures for accuracy.

Of note, patients with low COVID-19 risk scores were not included in the COVID-19 risk score and PCR test result co-occurrence matrix. A low COVID-19 score was assigned to CXRs with even the smallest probability of disease based on radiographic findings. We believe that low-risk scores may not serve as reliable predictive indicators for COVID-19 diagnosis, and the correlation between low-risk scores and PCR test results may be of minimal value. Several limitations of this framework with respect to calculating the PPV and NPV were also identified. To calculate PPV and NPV, the prevalence, sensitivity, and specificity of a test must be known. Unfortunately, the true prevalence for COVID-19 in this study population was unknown. To calculate the PPV and NPV for M-qXR, we had to use M-qXR's observed sensitivity and specificity at the operating threshold using the COVID-19 disease prevalence found in the validation dataset, which mirrored the disease prevalence reported in literature at the time of algorithm deployment. Therefore, further studies are needed to assess the validity and reliability of M-qXR as a triaging tool in patient populations in which the prevalence of COVID-19 is higher than in the reported literature.

5.3. Conclusion

This study evaluated M-qXR's ability to serve as a risk stratification tool to help evaluate patients with possible COVID-19 diagnosis. M-qXR was found to have comparable accuracy in detecting radiographic abnormalities on CXR suggestive of COVID-19 when compared to radiological ground truth. The M-qXR algorithm has the potential to provide benefit in guiding medical management of patients suspected of having COVID-19 who present with a high likelihood of disease, and where timely viral testing is not feasible due to limited resources. We also believe that M-qXR's ability to localize and quantify the affected regions on CXR, will enable us to monitor for progression of infection and evaluate response to treatment in future studies.

This study adds to the growing body of literature utilizing existing CXR and CT datasets in the public domain pertaining specifically to COVID-19 to train an AI model for COVID-19 classification and diagnosis. By adapting q-XR, a validated DL algorithm for CXR interpretation, to identify radiographic findings that are typically seen on CXR and are suggestive of COVID-19, we believe that M-qXR can help predict a patient's likelihood of being diagnosed with COVID-19. In addition, we believe that this study can also help to further validate future studies using a risk stratification approach such as this for a triaging setting.

Funding

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China. JAMA 2020. https://doi.org/10.1001/jama.2020.2648.

[2] World Health Organization. Coronavirus disease (COVID-2019) situation reports-july 2020. 2020.

[3] American College of Radiology. ACR recommendations for the use of chest radiography and computed tomography (CT) for suspected COVID-19 infection.

[4] Jacobi A, Chung M, Bernheim A, Eber C. Portable chest X-ray in coronavirus disease-19 (COVID-19): a pictorial review. Clin Imag 2020. https://doi.org/10.1016/j.clinimag.2020.04.001.

[5] Wong HYF, Lam HYS, Fong AHT, et al. Frequency and distribution of chest radiographic findings in COVID-19 positive patients. Radiology 2019. https://doi.org/10.1148/radiol.2020201160.

[6] Ng M-Y, Lee EY, Yang J, et al. Imaging profile of the COVID-19 infection: radiologic findings and literature review. Radiol. Cardiothorac Imag. 2020. https://doi.org/10.1148/ryct.2020200034.

[7] Yoon SH, Lee KH, Kim JY, et al. Chest radiographic and ct findings of the 2019 novel coronavirus disease (Covid-19): analysis of nine patients treated in korea. Korean J Radiol 2020. https://doi.org/10.3348/kjr.2020.0132.

[8] 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. https://doi.org/10.1148/radiol.2020201365.

[9] Kim HW, Capaccione KM, Li G, et al. The role of initial chest X-ray in triaging patients with suspected COVID-19 during the pandemic. Emerg Radiol 2020. https://doi.org/10.1007/s10140-020-01808-y.

[10] Chung HS, Lee DE, Kim JK, et al. Revised triage and surveillance protocols for temporary emergency department closures in tertiary hospitals as a response to COVID-19 crisis in Daegu Metropolitan city. J Kor Med Sci 2020. https://doi.org/10.3346/JKMS.2020.35.E189.

[11] Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer 2018. https://doi.org/10.1038/s41568-018-0016-5.

[12] Khuzani AZ, Heidari M, Shariati A. COVID-Classifier: an efficient machine learning model to assist in the diagnosis of COVID-19 infection in chest x-ray images. medRxiv 2020. https://doi.org/10.1101/2020.05.09.20096560.

[13] Mei X, Lee HC, yue Diao K, et al. Artificial intelligence–enabled rapid diagnosis of patients with COVID-19. Nat Med 2020. https://doi.org/10.1038/s41591-020-0931-3.

[14] Vaid S, Kalantar R, Bhandari M. Deep learning COVID-19 detection bias: accuracy through artificial intelligence. Int Orthop 2020. https://doi.org/10.1007/s00264-020-04609-7.

[15] Putha P, Tadepalli M, TRB Reddy. Can artificial intelligence reliably report chest x-rays? Radiologist Validation of an Algorithm trained on 2.3 Million X-Rays. 2018. Valid an.

[16] Mertz L. AI-driven COVID-19 tools to interpret, quantify lung images. IEEE Pulse 2020. https://doi.org/10.1109/MPULS.2020.3008354.

[17] Qin ZZ, Sander MS, Rai B, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: a multi-site evaluation of the diagnostic accuracy of three deep learning systems. Sci Rep 2019. https://doi.org/10.1038/s41598-019-51503-3.

[18] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition; 2016. https://doi.org/10.1109/CVPR.2016.90.

[19] Hu J, Shen L, Sun G. Squeeze-and-Excitation networks. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition.; 2018. https://doi.org/10.1109/CVPR.2018.00745.

[20] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, J Am Med Assoc 2016. https://doi.org/10.1001/jama.2016.17216.

[21] Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imag. 2018. https://doi.org/10.1007/s13244-018-0639-9.

[22] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017. https://doi.org/10.1038/nature21056.

[23] Bejnordi BE, Veta M, Van Diest PJ, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA, J Am Med Assoc 2017. https://doi.org/10.1001/jama.2017.14585.

[24] Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. Radiology 2018. https://doi.org/10.1148/radiol.2017170706.

[25] Fan J, Liu X, Pan W, Douglas MW, Bao S. Epidemiology of coronavirus disease in gansu province, China. Emerg Infect Dis 2020. https://doi.org/10.3201/EID2606.200251. 2020.

[26] Verity R, Okell LC, Dorigatti I, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis 2020. https://doi.org/10.1016/S1473-3099(20)30243-7.

[27] Mollura DJ, Azene EM, Starikovsky A, et al. White paper report of the rad-aid conference on international radiology for developing countries: identifying challenges, opportunities, and strategies for imaging services in the developing world. J Am Coll Radiol 2010. https://doi.org/10.1016/j.jacr.2010.01.018.

[28] Wynants L, Van Calster B, Bonten MMJ, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. BMJ 2020. https://doi.org/10.1136/bmj.m1328.

[29] Muhammad E. H. C, Tawsifur R, Amith K, et al. COVID-19 Radiology Database. Can AI help in screening Viral and COVID-19 pneumonia?.