Supplementary Methods

Training for the RadGen AI Model

RadGenX represents a collaboratively updated AI prediction model that is based on the previously reported RadGen CXR AI prediction model [1; 2]. RadGen was originally trained on a total of 146,876 CXRs, including: 112,120 CXR films from the ChestX-ray14 dataset used for model pre-training, and 6,326 CXRs with corresponding SARS-CoV-2 confirmed test results which included CXRs collected from 4 institutions from Hong Kong (714 and 2,471 COVID-19 positive and negative), public datasets (accessed before May 1, 2020) including BIMCV-COVID19 (2,474 positive), Italian Society of Medical and Interventional Radiology (SIRM) (83 positive) and GitHub (452 and 132 COVID-19 positive and negative) and 28,430 randomly selected historic negative CXRs from ChestX-ray14, CheXpert, PadChest, and RSNA Kaggle [3-6].

Private study patients for the RadGenX AI Model

At the beginning of the study, all 9 private study sites were invited to contribute a subset of their data for AI model updating. Cases for model updating and for validation consisted of a subset of "private" study patients collected over a prospectively pre-designated enrollment period from February 1, 2020, to July 31, 2020, at a ratio of 1:2 COVID-19 positive to negative cases. On September 15, 2020, following the close of the private study data collection period, we distributed the AI software to all private study sites for training and for voluntary data contribution for AI model updating. Sites that opted to participate could volunteer "contributed data," which was prospectively pre-specified as the initial ¾ of their cases (in chronological order by date of CXR), for both COVID-19 positive and negative cases and historical negatives,

with the remaining non-overlapping ¼ of their cases (the most recent chronologically) completely held out and untouched for contribution to the international test set. If sites chose not to contribute cases for model updating, then by definition all of their cases were held out for use in the international test set

Modified Federated Learning System

All technical training and development were performed using PyTorch version 1.7 on a Linux workstation (AMD1950X central processing unit with 128 Gigabytes (GB) random access memory (RAM) and 4 Nvidia 2080TI graphical processing units). We first trained a simple U-Net like segmentation model to identify frontal CXRs, and to then segment and crop the chest from these CXRs [7]. This model was trained with CXR images from the NIH Chest-X-ray set consisting of 20,000 frontal, 9,000 non-frontal and 100 normal frontal, randomly selected CXRs with an additional 170 randomly selected frontal CXRs with the chest manually segmented. This segmentation model was incorporated into RadGen which was then converted into a standalone software application using the PyTorch C++ API [1; 2]. To mitigate privacy concerns for data contribution for model updating we implemented multiple steps. First, this software runs on any personal computer running Microsoft Windows 10 or later with 8 GB of RAM or greater as a fully independent desktop application and does not require an internet connection to operate. Second, input images are stripped of all DICOM header information, and a sha1 hash of the patient id is used for folds allocation and then discarded. Further, in order to eliminate potential text or identifier information that may be present on the actual image, the images are further de-identified via chest cropping. Finally, after chest segmentation, the cropped chest image is then inputted into the RadGen SE-ResNeXt model; the first 3 layers of the network are frozen, and the 3rd layer extracted as “anonymized and encoded intermediate layer image representation

data”. Finally, if a user then chooses to contribute CXR image data for model updating, they must select the file containing the anonymized and encrypted image representation, which is placed in a separate file directory, and send this to the investigators for model updating.

Supplementary Tables

Table S1. Demographic data on patients used for RadGenX model updating.

	Number of CXRs (Positive CXRs)	**Number of Patients (Positive Patients)**	**Mean Age (95% CI), years**	**Gender, % male**
Study Site Contributed Datasets (city, country)				
Rome, Italy	1,361 (668)	1,062 (370)	63·19 (62·08,64·31)	54·0%
Brescia, Italy	1,618 (868)	1,125 (375)	70·72 (69·79,71·64)	58·0%
Pavia, Italy	671 (646)	389 (364)	66·41 (64·83,67·98)	62·5%
Public Datasets				
BIMCV-COVID19[6]	18,072 (11,953)	6,980 (3,855)	62·74 (62·31,63·17)	49·4%
COVID-AR[8]	252 (252)	104 (104)	54·03 (50·59,57·47)	49·0%
All datasets	21,974 (14,387)	9,660 (5,068)	63·77 (63·41,64·13)	51·5%

Table S2. Experience levels of the three international radiologists in the reader study.

	Location	**Years of Experience**	**Thoracic Imaging Fellowship trained**
Reader 1	Hong Kong SAR, China	13	No
Reader 2	Italy	23	Yes
Reader 3	United States	15	No

Table S3. All cases collected and cases eligible for inclusion in the international test set cohort.

Region	**Overall**		**COVID-19 positive cases**		**COVID-19 negative cases**	
	Collected	**Eligible for validation analysis**	**Collected**	**Eligible for validation analysis**	**Collected**	**Eligible for validation analysis**
Total	7,513	5,894	3,684	2,747	3,829	3,147
United States	3,605	2,671	2,040	1,713	1,565	958
Europe	1,981	1,456	894	614	1,087	842

Hong Kong	1,927	1,767	750	420	1,177	1,347

Table S4. Demographic distribution of patients in the international test set cohort.

Region	Number of Patients	Mean Age (95% CI), years	Gender, % male
Europe	1,456	65.89 (64.99,66.79)	60.3%
United States	2,671	58.51 (57.83, 59.20)	53.1%
Hong Kong	1,767	61.85 (60.88,62.83)	55.4%

References

1 Chiu WHK, Vardhanabhuti V, Poplavskiy D et al (2020) Detection of COVID-19 Using Deep Learning Algorithms on Chest Radiographs. J Thorac Imaging 35:369-376

2 Chiu WHK, Poplavskiy D, Zhang S, Ho Yu PL, Kuo MD (2021) Dynamic Prediction of SARS-CoV-2 RT-PCR status on Chest Radiographs using Deep Learning Enabled Radiogenomics. medRxiv. 10.1101/2021.01.10.21249370:2021.2001.2010.21249370

3 Bustos A, Pertusa A, Salinas J-M, de la Iglesia-Vayá M (2019) PadChest: A large chest x-ray image dataset with multi-label annotated reports. arXiv e-prints

4 Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2017) ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. arXiv e-prints

5 Irvin J, Rajpurkar P, Ko M et al (2019) CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison.

6 de la Iglesia Vayá M, Saborit JM, Montell JA et al (2020) BIMCV COVID-19+: a large annotated dataset of RX and CT images from COVID-19 patients.

7 Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentationInternational Conference on Medical image computing and computer-assisted intervention. Springer, pp 234-241

8 Desai S, Baghal A, Wongsurawat T et al (2020) Chest imaging representing a COVID-19 positive rural U.S. population. Sci Data 7:414