

**Medical Imaging Data Resource Center (MIDRC) - RSNA
International COVID-19 Open Radiology Database
(RICORD) Release 1a - Chest CT Covid+ (MIDRC-RICORD-
1A)**

Summary

Redirection Notice

This page will redirect to <https://www.cancerimagingarchive.net/collection/midrc-ricord-1a/> in about 5 seconds.

Background

The COVID-19 pandemic is a global healthcare emergency. Prediction models for COVID-19 imaging are rapidly being developed to support medical decision making in imaging. However, inadequate availability of a diverse annotated dataset has limited the performance and generalizability of existing models.

Purpose

The Radiological Society of North America (RSNA) assembled the [RSNA International COVID-19 Open Radiology Database \(RICORD\)](#), a collection of COVID-related imaging datasets and expert annotations to support research and education. The RICORD datasets are made freely available to the research community and will be incorporated in the [Medical Imaging and Data Resource Center \(MIDRC\)](#), a multi-institutional research data repository funded by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health.

Materials and Methods

MIDRC-RICORD dataset 1a was created through a collaboration between the RSNA and the Society of Thoracic Radiology (STR). Pixel-level volumetric segmentation with clinical annotations by thoracic radiology subspecialists was performed for all COVID positive thoracic computed tomography (CT) imaging studies in a labeling schema coordinated with other international consensus panels and COVID data annotation efforts.

Results

MIDRC-RICORD dataset 1a consists of 120 thoracic computed tomography (CT) scans from four international sites annotated with detailed segmentation and diagnostic labels.

Patient Selection: Patients at least 18 years in age receiving positive diagnosis for COVID-19.

Data Abstract

1. 120 Chest CT examinations (axial series only, any protocol).
2. Annotations comprised of
 - a) Detailed segmentation of affected regions;
 - b) Image-level labels (Infectious opacity, Infectious TIB/micronodules, Infectious cavity, Noninfectious nodule/mass, Atelectasis, Other noninfectious opacity)
 - c) Exam-level diagnostic labels (Typical, Indeterminate, Atypical, Negative for pneumonia, Halo sign, Reversed halo sign, Reticular pattern w/o parenchymal opacity, Perilesional vessel enlargement, Bronchial wall thickening, Bronchiectasis, Subpleural curvilinear line, Effusion, Pleural thickening, Pneumothorax, Pericardial effusion, Lymphadenopathy, Pulmonary embolism, Normal lung, Infectious lung disease, Emphysema, Oncologic lung disease, Non-infectious inflammatory lung disease, Non-infectious interstitial, Fibrotic lung disease, Other lung disease)

d) Exam-level procedure labels (With IV contrast, Without IV contrast, QA- inadequate motion/breathing, QA- inadequate insufficient inspiration, QA- inadequate low resolution, QA- inadequate incomplete lungs, QA- inadequate wrong body part/modality, Endotracheal tube, Central venous/arterial line, Nasogastric tube, Sternotomy wires, Pacemaker, Other support apparatus).

3. Supporting clinical variables: MRN*, Age, Study Date*, Exam Description, Sex, Study UID*, Image Count, Modality, Testing Result, Specimen Source (* pseudonymous values).

How to use the JSON annotations

More information about how the JSON annotations are organized can be found on <https://docs.md.ai/data/json/>. Steps 2 & 3 in this [example code](#) demonstrate how to load the JSON into a Dataframe. The JSON file can be downloaded via the data access table below; it is not available via MD.ai. This [Jupyter Notebook](#) may also be helpful.

Code for converting CT scan segmentation labels for lung opacities from MD.ai JSON to DICOM-SEG : <https://github.com/QIICR/dcmqi/blob/add-mdai-converter/util/mdai2dcm.py>

Research Benefits

As this is a public dataset, RICORD is available for non-commercial use (and further enrichment) by the research and education communities which may include development of educational resources for COVID-19, use of RICORD to create AI systems for diagnosis and quantification, benchmarking performance for existing solutions, exploration of distributed/federated learning, further annotation or data augmentation efforts, and evaluation of the examinations for disease entities beyond COVID-19 pneumonia. Deliberate consideration of the detailed annotation schema, demographics, and other included meta-data will be critical when generating cohorts with RICORD, particularly as more public COVID-19 imaging datasets are made available via complementary and parallel efforts. It is important to emphasize that there are limitations to the clinical “ground truth” as the SARS-CoV-2 RT-PCR tests have widely documented limitations and are subject to both false-negative and false-positive results which impact the distribution of the included imaging data, and may have led to an unknown epidemiologic distortion of patients based on the inclusion criteria. These limitations notwithstanding, RICORD has achieved the stated objectives for data complexity, heterogeneity, and high-quality expert annotations as a comprehensive COVID-19 thoracic imaging data resource.

Acknowledgements

We would like to acknowledge the individuals and institutions that have provided data for this collection: This dataset was created through a collaboration between the RSNA and Society of Thoracic Radiology (STR). Data in RICORD will be made available through the Medical Imaging Data Resource Center, funded through a contract with the National Institute for Biomedical Imaging and Bioengineering (NIBIB).

Data Access

Data Access

Data Type	Download all or Query/Filter	License
Images (DICOM, 11 GB)	Download Search (Download requires the NBIA Data Retriever)	CC BY-NC 4.0

Annotations (JSON, 13.62 kB)	Download	CC BY-NC 4.0
Clinical data (csv, 24 kB)	Download	CC BY-NC 4.0

Click the Versions tab for more info about data releases.

Please contact help@cancerimagingarchive.net with any questions regarding usage.

Additional Resources for this Dataset

The NCI Cancer Research Data Commons (CRDC) provides access to additional data and a cloud-based data science infrastructure that connects data sets with analytics tools to allow users to share, integrate, analyze, and visualize cancer research data.

- [Imaging Data Commons \(IDC\)](#) (Imaging Data)

Detailed Description

Detailed Description

Image Statistics	
Modalities	CT
Number of Patients	110
Number of Studies	120
Number of Series	229
Number of Images	31856
Images Size (GB)	11

Citations & Data Usage Policy

Citations & Data Usage Policy

Users must abide by the [TCIA Data Usage Policy and Restrictions](#). Attribution should include references to the following citations:



Data Citation

Tsai, E., Simpson, S., Lungren, M.P., Hershman, M., Roshkovan, L., Colak, E., Erickson, B.J., Shih, G., Stein, A., Kalpathy-Cramer, J., Shen, J., Hafez, M.A.F., John, S., Rajiah, P., Pogatchnik, B.P., Mongan, J.T., Altinmakas, E., Ranschaert, E., Kitamura, F.C., Topff, L., Moy, L., Kanne, J.P., & Wu, C. (2020). **Data from the Medical Imaging Data Resource Center - RSNA International COVID Radiology Database Release 1a - Chest CT Covid+ (MIDRC-RICORD-1A)**. The Cancer Imaging Archive . DOI: <https://doi.org/10.7937/VTW4-X588>

Publication Citation

Tsai, E. B., Simpson, S., Lungren, M., Hershman, M., Roshkovan, L., Colak, E., Erickson, B. J., Shih, G., Stein, A., Kalpathy-Cramer, J., Shen, J., Hafez, M., John, S., Rajiah, P., Pogatchnik, B. P., Mongan, J., Altinmakas, E., Ranschaert, E. R., Kitamura, F. C., ... Wu, C. C. (2021). **The RSNA International COVID-19 Open Annotated Radiology Database (RICORD)**. Radiology, 203957. DOI: <https://doi.org/10.1148/radiol.2021203957>

TCIA Citation

Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057. DOI: [10.1007/s10278-013-9622-7](https://doi.org/10.1007/s10278-013-9622-7)

Other Publications Using This Data

TCIA maintains [a list of publications](#) which leverage TCIA data. If you have a manuscript you'd like to add please [contact the TCIA Helpdesk](#).

Versions

Version 2 (Current): Updated 2020/01/14

Data Type	Download all or Query/Filter
Images (DICOM, 11 GB)	Download Search (Requires NBIA Data Retriever .)
Annotations (JSON)	Download
Clinical data (csv)	Download

Clinical data spreadsheet added.

Version 1: Updated 2020/12/18

Data Type	Download all or Query/Filter
Images (DICOM, 11 GB)	Download Search (Requires NBIA Data Retriever .)
Annotations (JSON)	Download

