CENTER FOR VIRTUAL IMAGING TRIALS

Virtual vs. Reality: External Validation of COVID-19 Classifiers using XCAT Phantoms for Chest Computed Tomography

Fakrul Islam Tushar, Ehsan Abadi, Saman Sotoudeh-Paima, Rafael B.Fricks, Maciej A. Mazurowski, W. Paul Segars, Ehsan Samei, Joseph Y. Lo

Center for Virtual Imaging Trials, Carl E. Ravin Advanced Imaging Laboratories,

Dept. of Radiology, Duke University School of Medicine;

Dept. of Electrical & Computer Engineering, Pratt School of Engineering, Duke University.

SPIE Medical Imaging, Session # 4 (COVID-19), paper #12033-9, Feb 21, 2022.

Duke Radiology

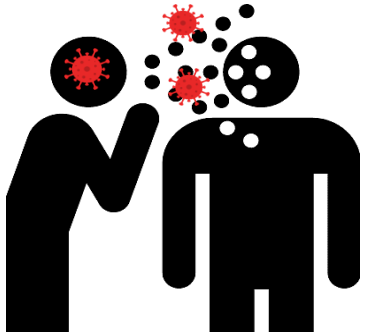
DUKE UNIVERSITY MEDICAL PHYSICS GRADUATE PROGRAM

Duke Clinical Imaging Physics Group

Duke UNIVERSITY | PRATT SCHOOL of ENGINEERING

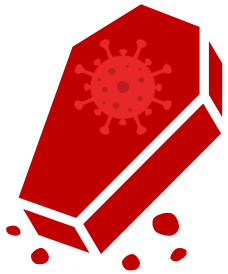
Overview| COVID-19

Cases



414.16 M

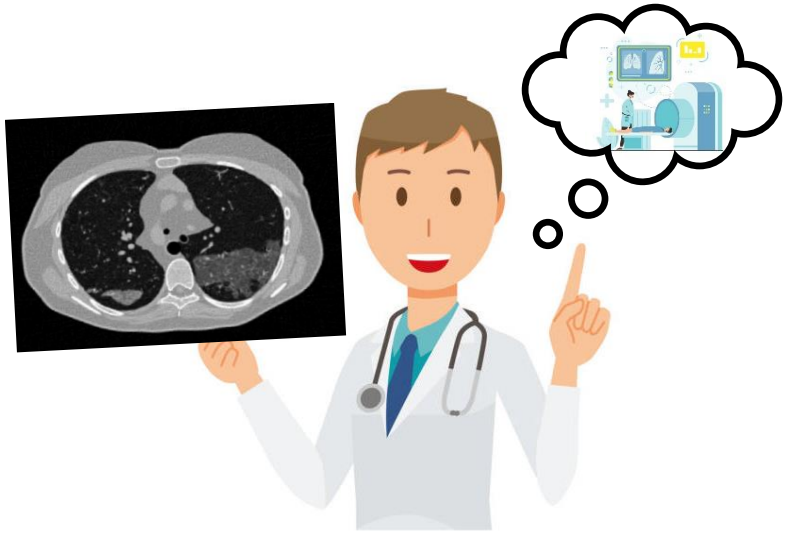
Deaths



5.84 M

Infection statistics world-wide

Chest Computed Tomography (CT)



Effectively determine lung involvement

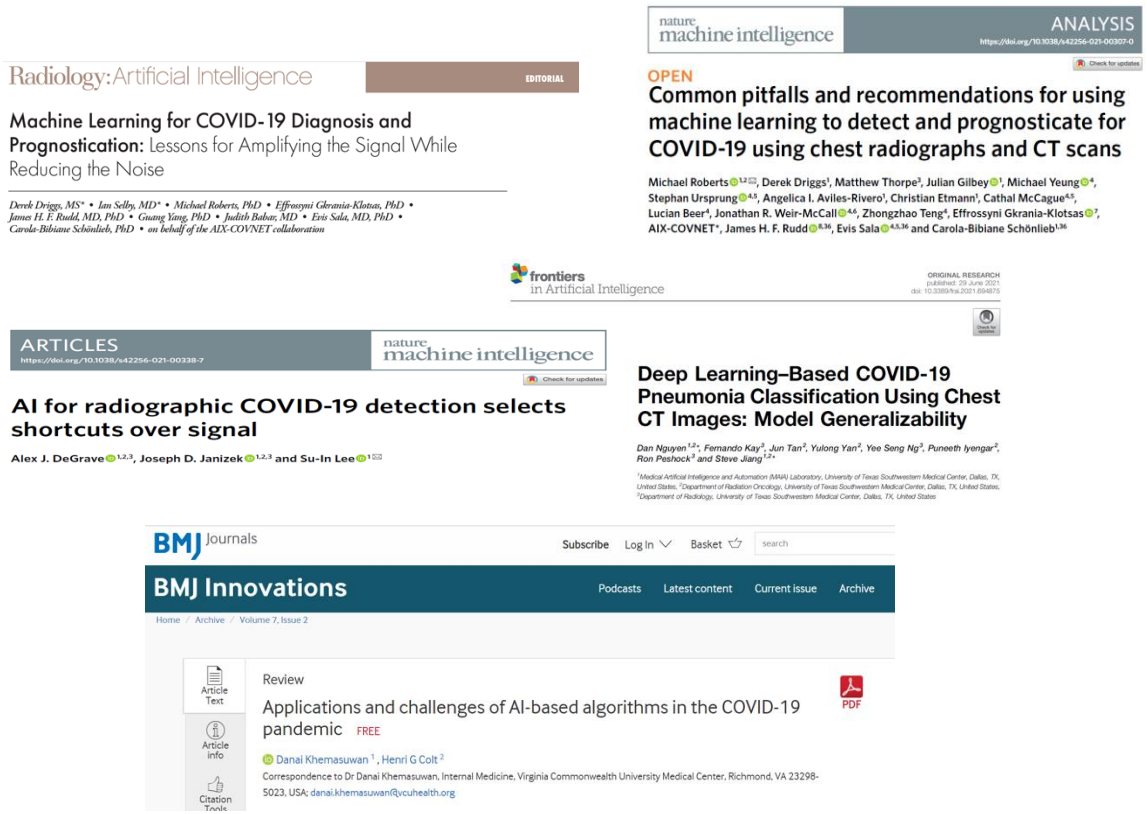
Overview| COVID-19

Google Scholar

91,400 results

Associated to COVID-19, Machine-learning, and Deep-learning.

Overview| **Concerns**



Radiology: Artificial Intelligence — EDITORIAL

Machine Learning for COVID-19 Diagnosis and Prognostication: Lessons for Amplifying the Signal While Reducing the Noise

Derek Driggs, MS • Ian Selby, MD* • Michael Roberts, PhD • Effrosyni Gkrania-Klotsas, PhD • James H. F. Rudd, MD, PhD • Guang Yang, PhD • Judith Babar, MD • Evis Sala, MD, PhD • Carola-Bibiane Schönlieb, PhD • on behalf of the AIX-COVNET collaboration*

nature machine intelligence — ANALYSIS

OPEN

Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans

Michael Roberts, Derek Driggs, Matthew Thorpe, Julian Gilbey, Michael Yeung, Stephan Ursprung, Angelica I. Aviles-Rivero, Christian Etmann, Cathal McCague, Lucian Beer, Jonathan R. Weir-McCall, Zhongzhao Teng, Effrossyni Gkrania-Klotsas, AIX-COVNET, James H. F. Rudd, Evis Sala and Carola-Bibiane Schönlieb

frontiers in Artificial Intelligence — ORIGINAL RESEARCH

Deep Learning–Based COVID-19 Pneumonia Classification Using Chest CT Images: Model Generalizability

Dan Nguyen, Fernando Kay, Jun Tan, Yulong Yan, Yee Seng Ng, Puneeth Iyengar, Ron Peshock and Steve Jiang

ARTICLES — nature machine intelligence

AI for radiographic COVID-19 detection selects shortcuts over signal

Alex J. DeGrave, Joseph D. Janizek and Su-In Lee

BMJ Journals — BMJ Innovations — Podcasts — Latest content — Current issue — Archive

Review

Applications and challenges of AI-based algorithms in the COVID-19 pandemic FREE

Danai Khemasuwan, Henri G Colt

Correspondence to Dr Danai Khemasuwan, Internal Medicine, Virginia Commonwealth University Medical Center, Richmond, VA 23298-5023, USA; danai.khemasuwan@vcuhealth.org

- Near perfect diagnostic performance
- Internal and external variability in performance
- Underlying data biases
- performance can vary depending on data population, image quality, image acquisition protocols, dose differences, and disease appearance.

Objective | Primary Objective

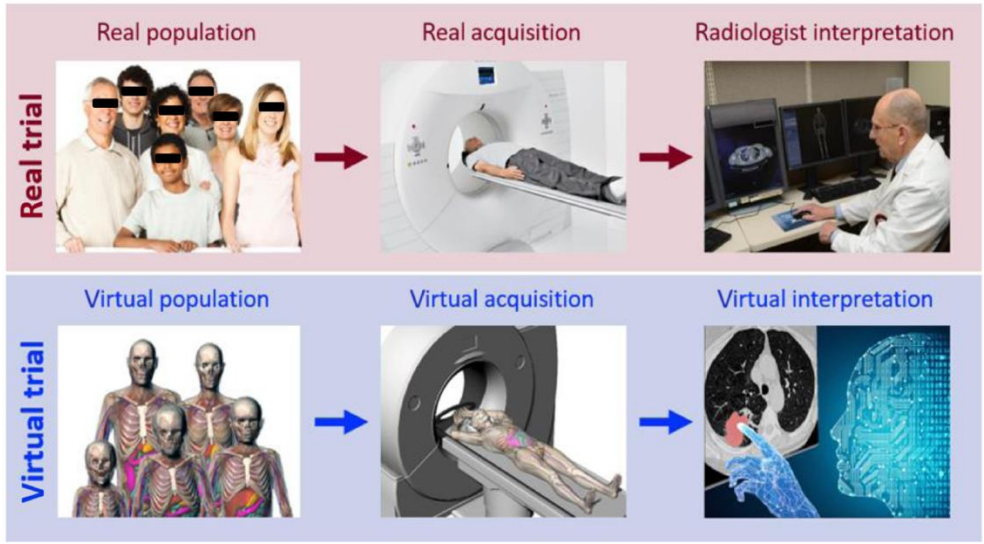
Investigating variability of diagnostic performance based on disease appearance and imaging properties.

Objective | **Challenges**

— Challenging in real clinical setup

— Health concerns- Ionizing radiation

Probable solution -Virtual Imaging Trials (VITs)

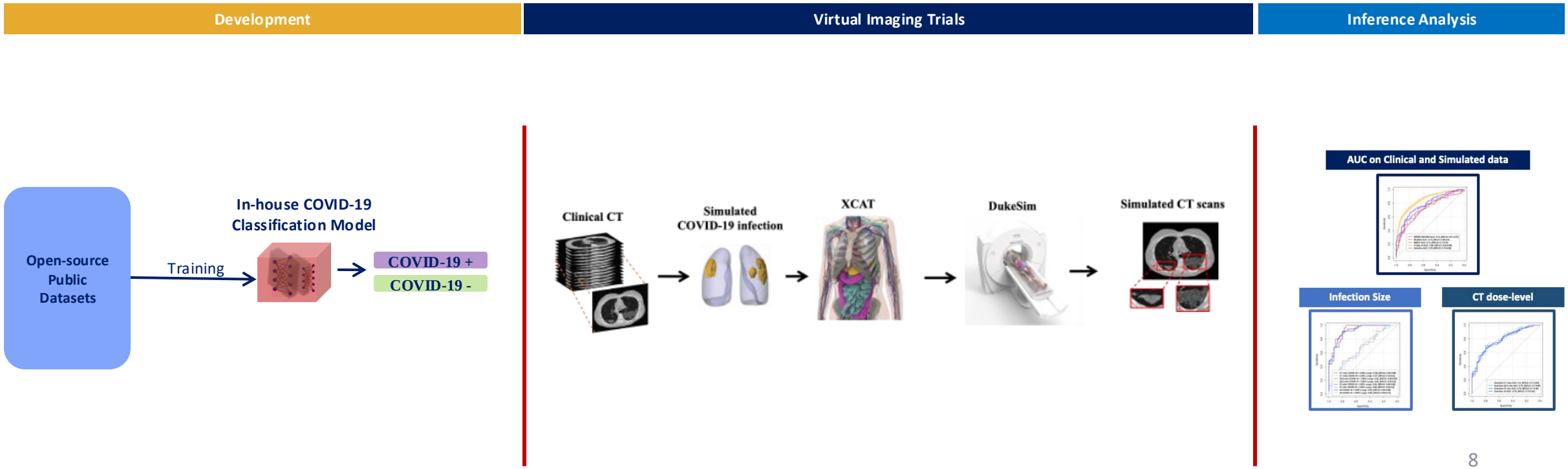


— **Virtual imaging trial** (VIT) is a process of simulating imaging evaluations with varying factors such as **computational human phantoms** (CPs), **imaging scanner systems**, and **virtual readers** .

—**AI** model belongs in the **virtual reader** category.

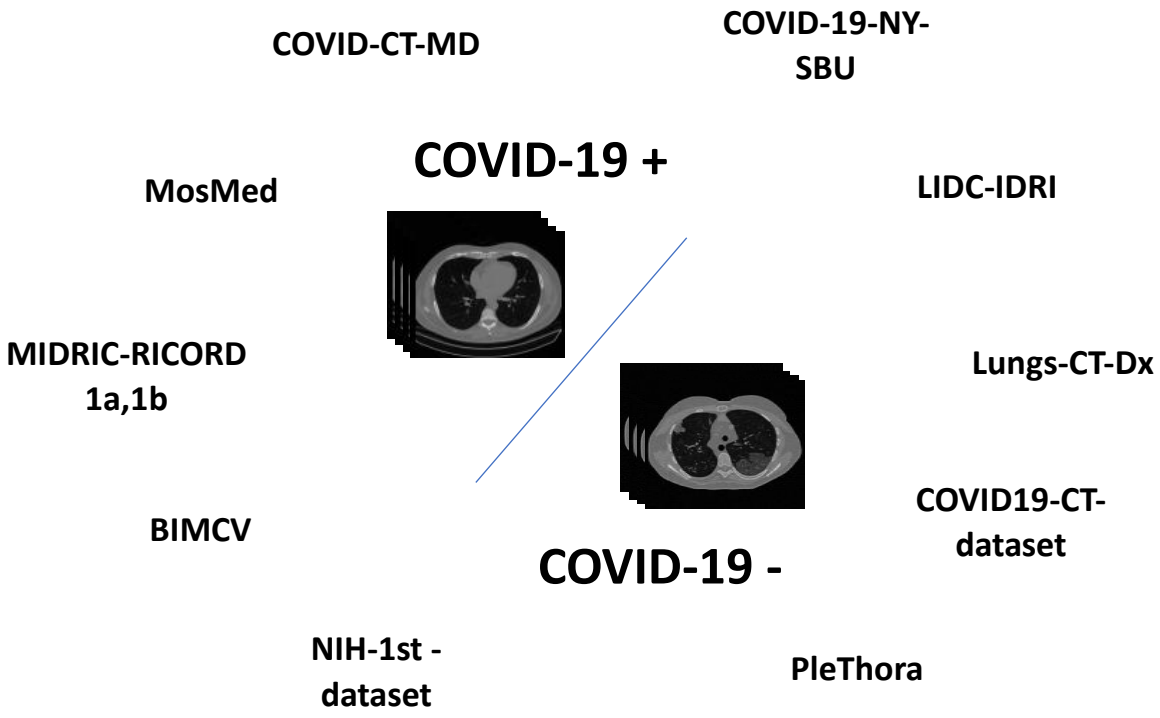
Method| Study Design

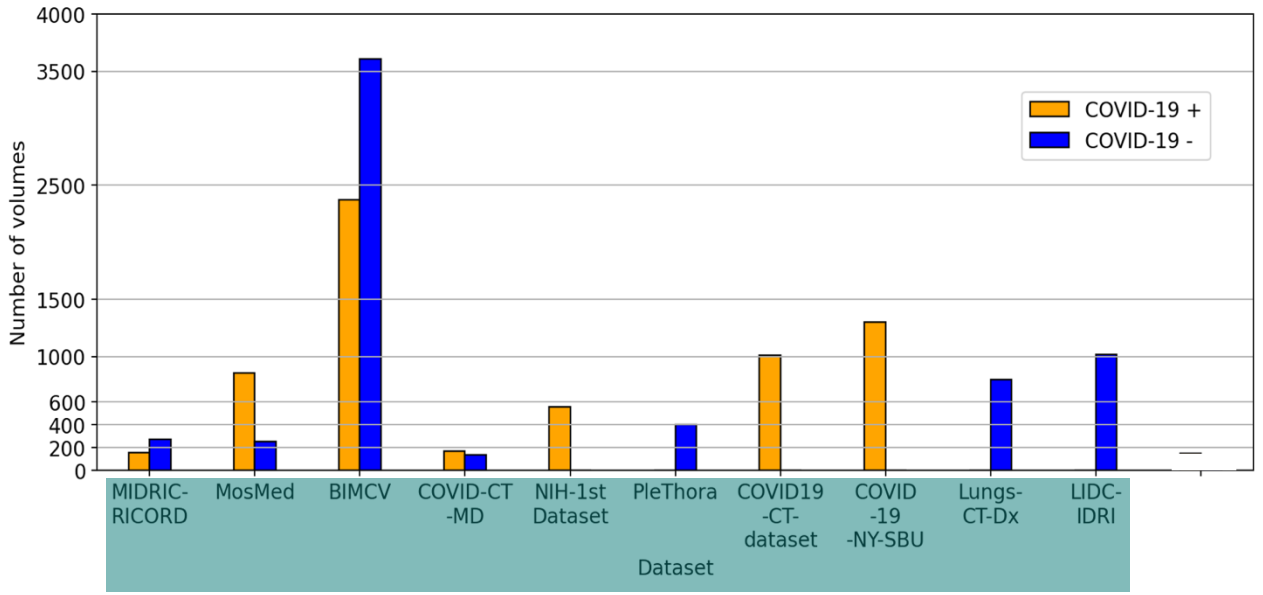
- Open-Source clinical data for model development.
- Generating Simulated CT scans utilizing VITs platform.
- Clinical Vs . Virtual performance analysis.



Method| Open-source Datasets

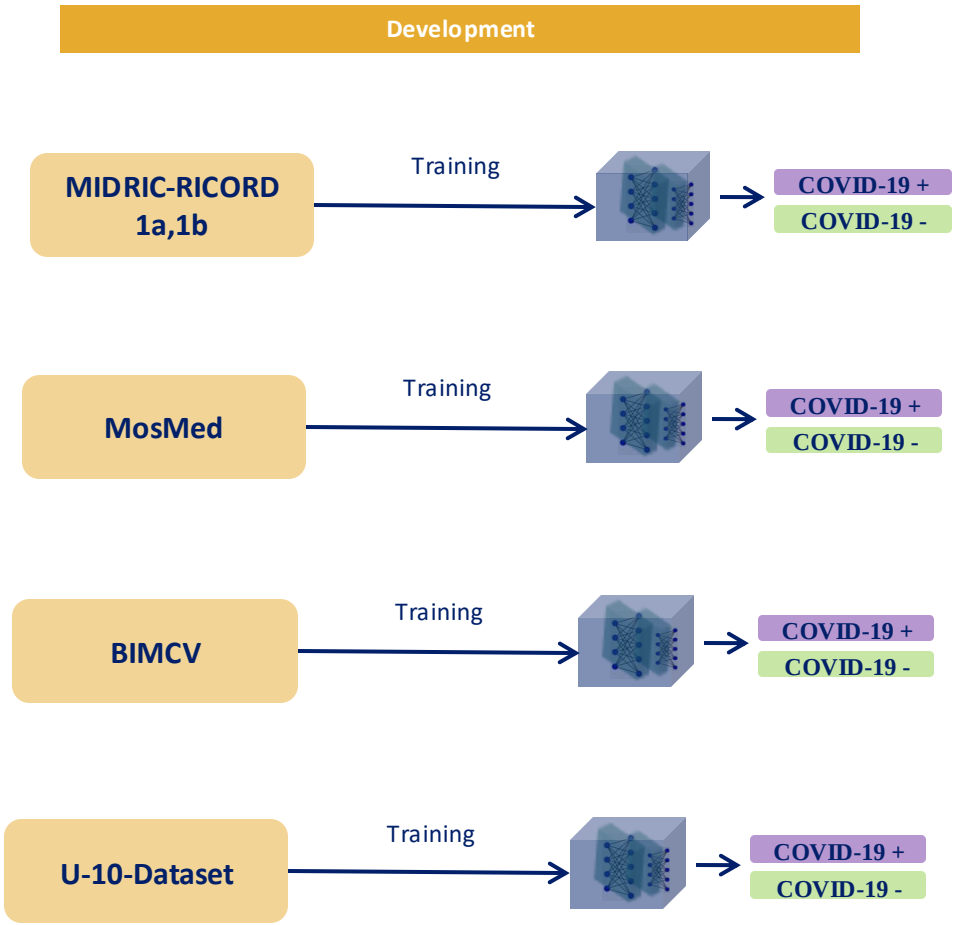
12,000+ open-source clinical CT scans from 6,847 patients.



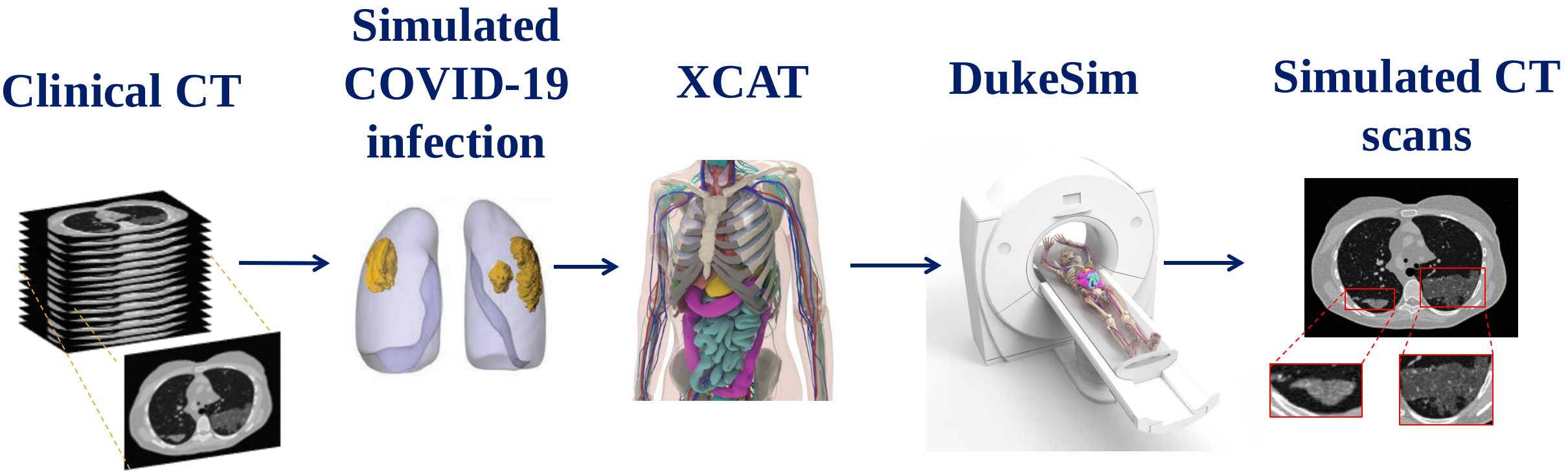


U-10 Dataset

Method| Model Development



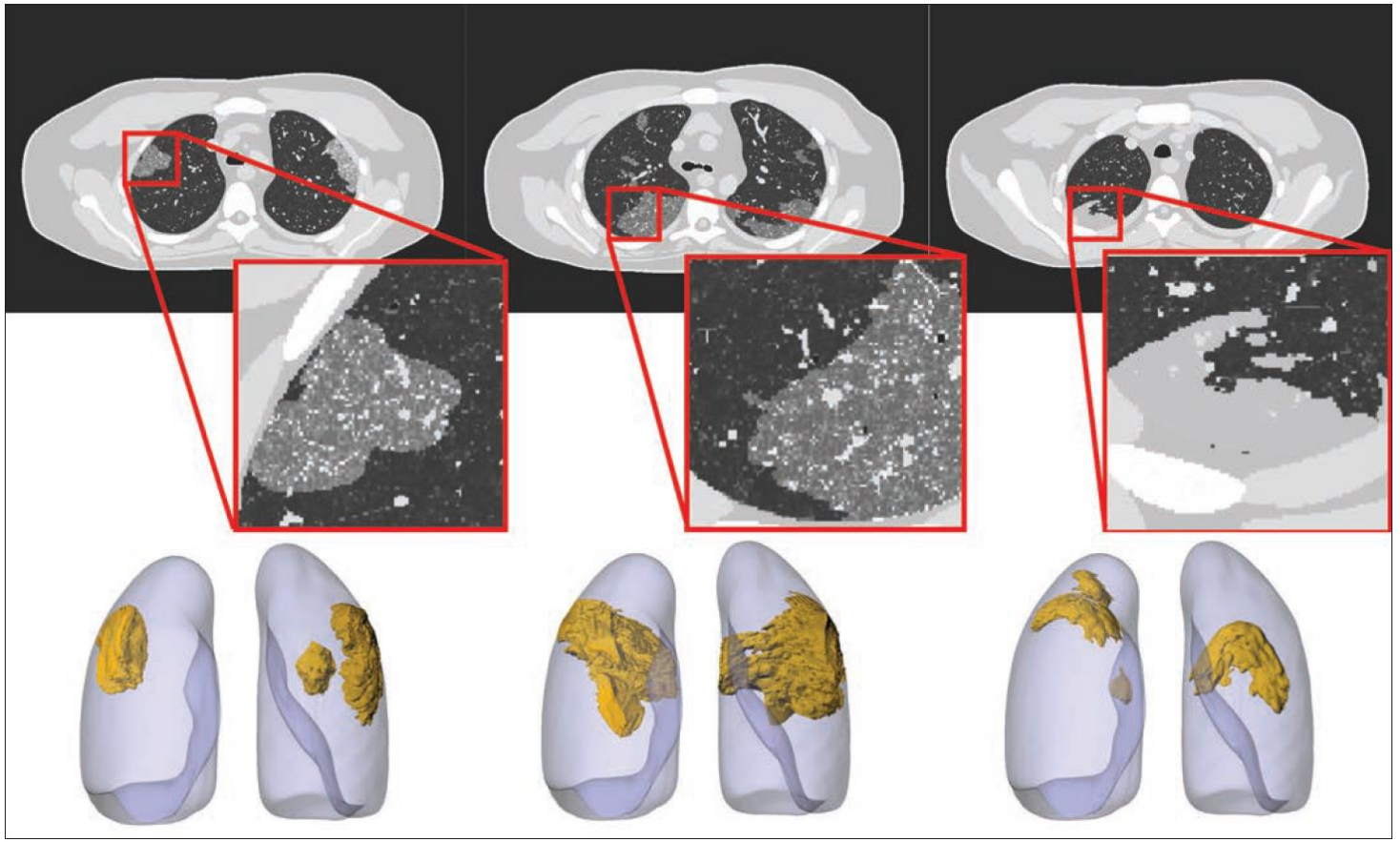
Method| Virtual Imaging Trials (VITs)



E. Abadi, W. Paul Segars, H. Chalian, and E. Samei, "**Virtual Imaging Trials for Coronavirus Disease (COVID-19)**," AJR Am J Roentgenol, vol. 216, no. 2, pp. 362-368, Feb 2021, doi: 10.2214/AJR.20.23429.

Method | Virtual Imaging Trials (VITs)

Ground Truth 4D **XCAT COVID-19 phantoms**



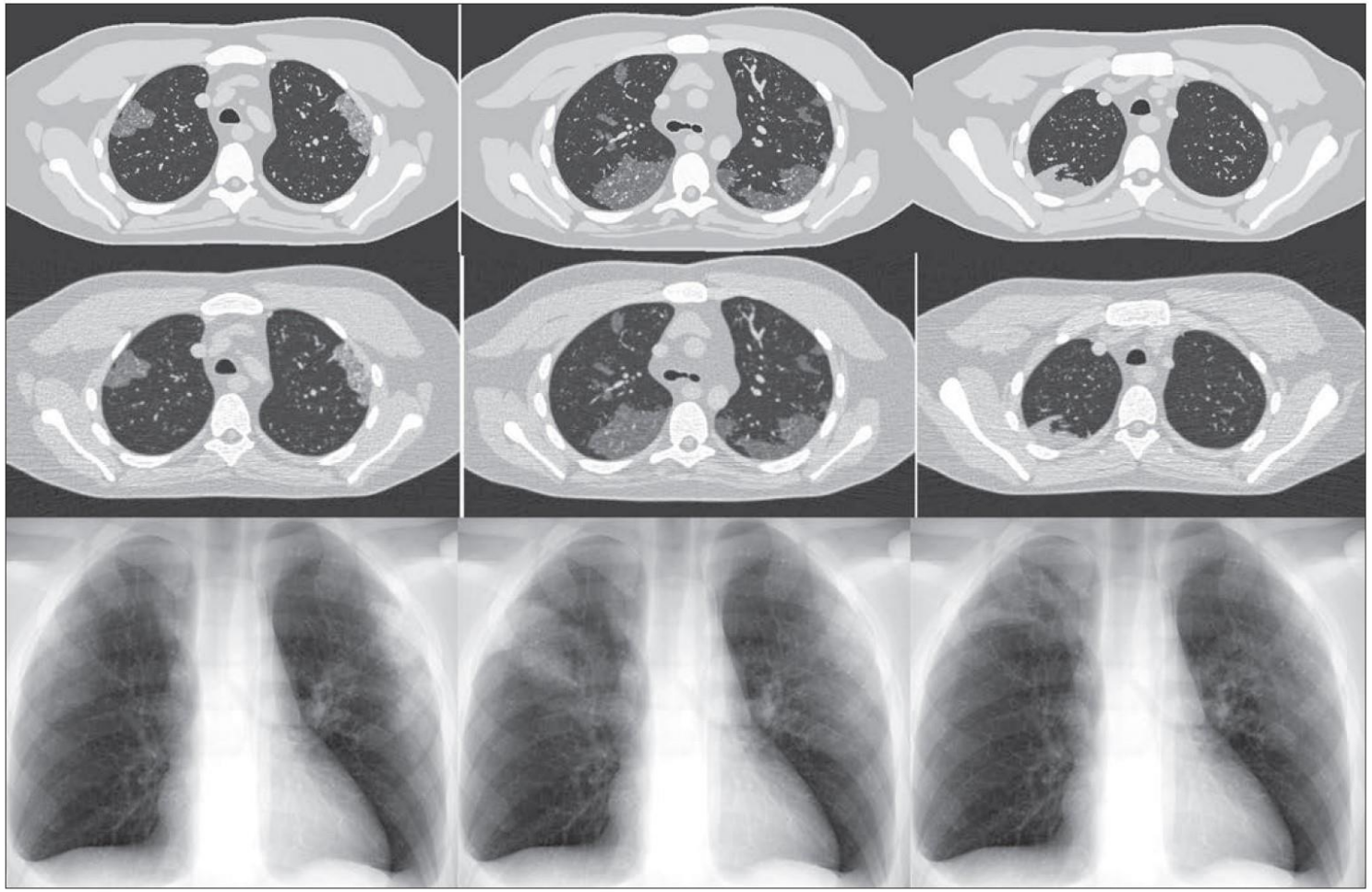
E. Abadi et al. (2021).

Method | Virtual Imaging Trials (VITs)

4D **XCAT phantom** developed at Duke University

Simulated CT

Simulated Xray



E. Abadi et al. (2021).

Method | Virtual Imaging Trials (VITs)

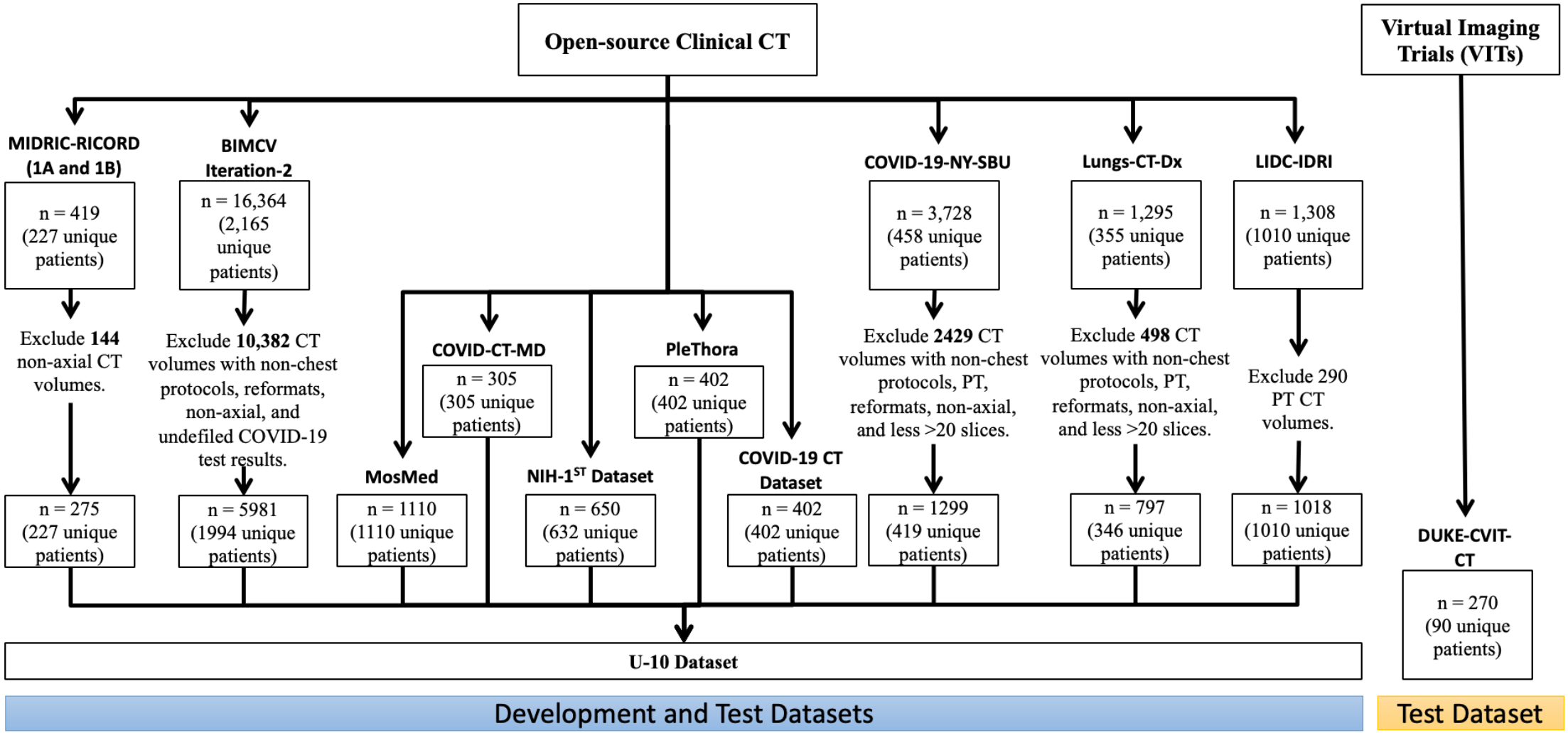
Duke-CVIT-CT / DukeSim Dataset

Dose Level (mAs)	Number of Volume	
	COVID-19	Normal
5.7	50	40
28.5	50	40
57	50	40
Total	**150**	**120**

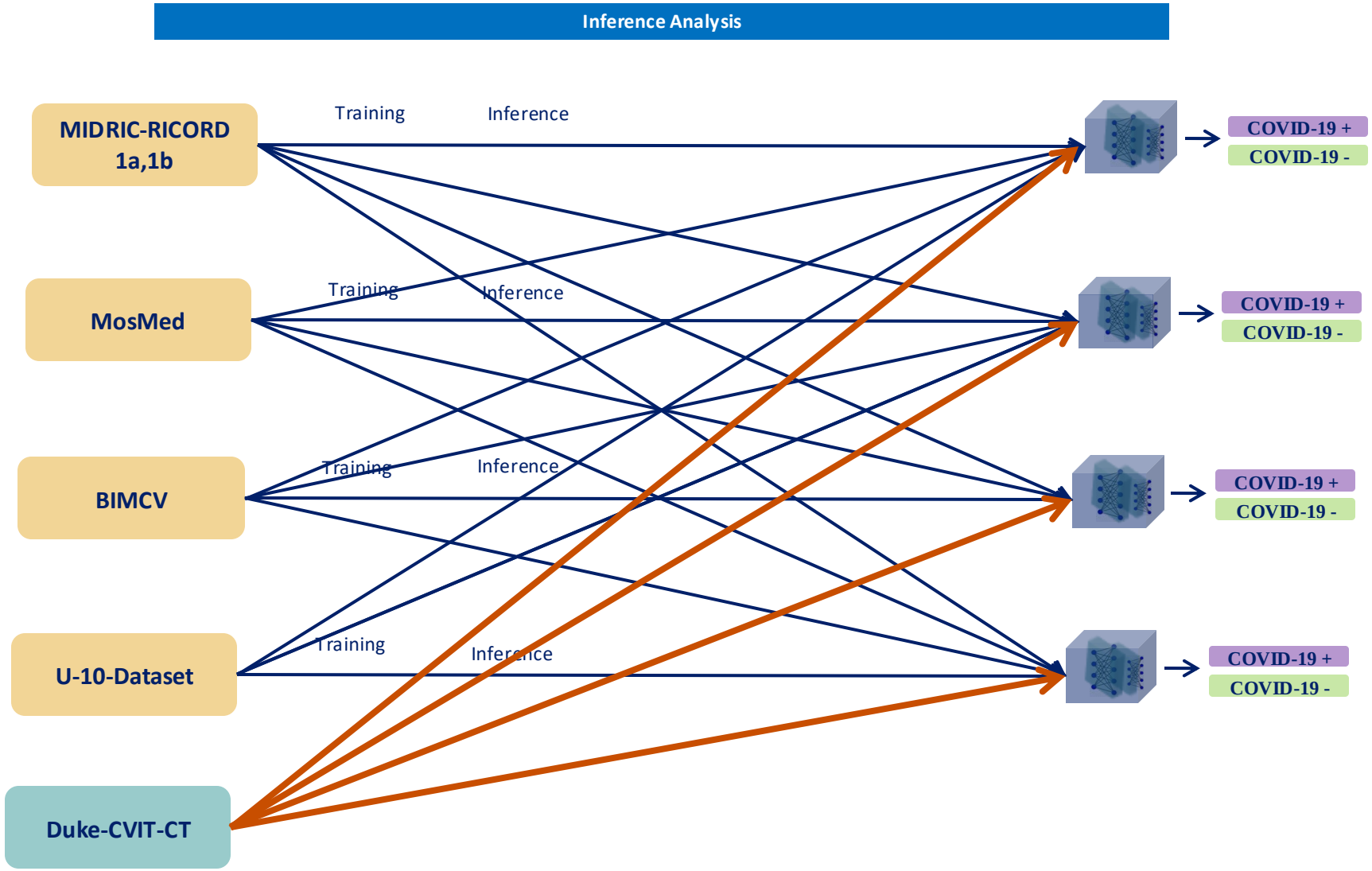
Method | Clinical Vs Virtual Data

Datasets	Class Type		Label Level			Same CT scan with multiple dose levels
	COVID-19 positive	COVID-19 negative	Patient-level	Slice-level	Pixel-level	
MIDRIC-RICORD						

Method | Inclusion and Exclusion



Method| Inference



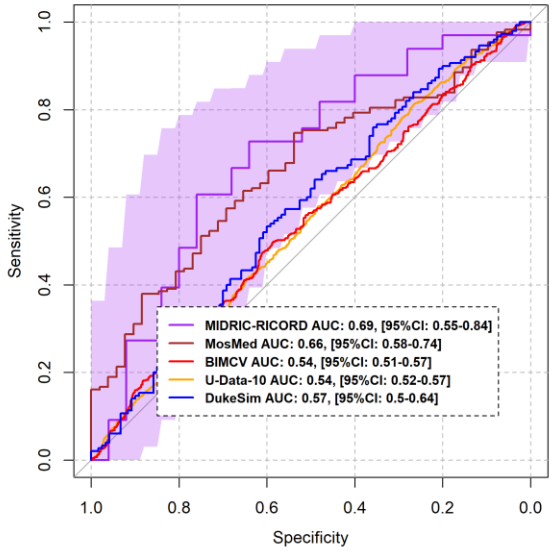
Results | Clinical and Virtual Data

Test (COVID+/-) / Training (COVID+/-)	MIDRIC-RICORD *(33/25)	MosMed (174/52)	BIMCV-V2 (470/823)	U-10-Dataset (1159/1201)	Duke-CVIT-CT (150/120)
MIDRIC-RICORD (90/72)	**0.69** [0.55,0.84]	0.66 [0.58,0.74]	0.54 [0.51,57]	0.54 [0.52-0.57]	0.57 [0.50,0.64]
MosMed (512/152)	0.76 [0.64,0.88]	**0.87** [0.81,0.92]	0.63 [0.60,0.66]	0.68 [0.66,0.70]	0.69 [0.63,0.76]
BIMCV-V2 (1421/2077)	0.64 [0.49,0.78]	0.58 [0.50,0.67]	**0.77** [0.74,0.79]	0.68 [0.66,0.70]	0.68 [0.62,0.75]
U-10-Dataset (3926/3768)	0.74 [0.61,0.87]	0.73 [0.66,0.80]	0.73 [0.70,0.76]	**0.85** [0.84,0.86]	0.79 [0.73,0.84]

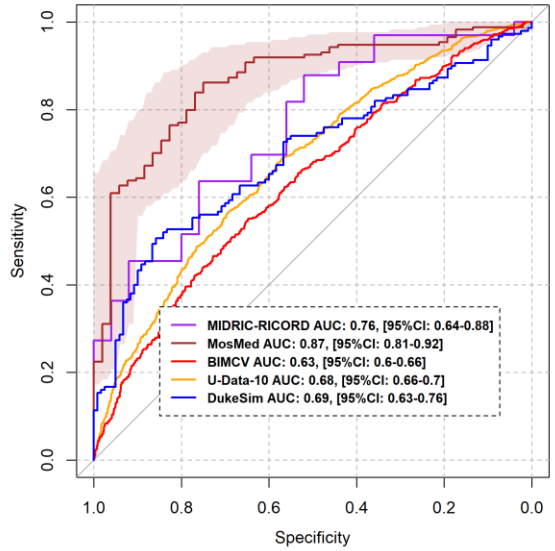
Results | Clinical and Virtual Data

Performance dropped due to domain-shift, simulated data consistent with multiple clinical test set.

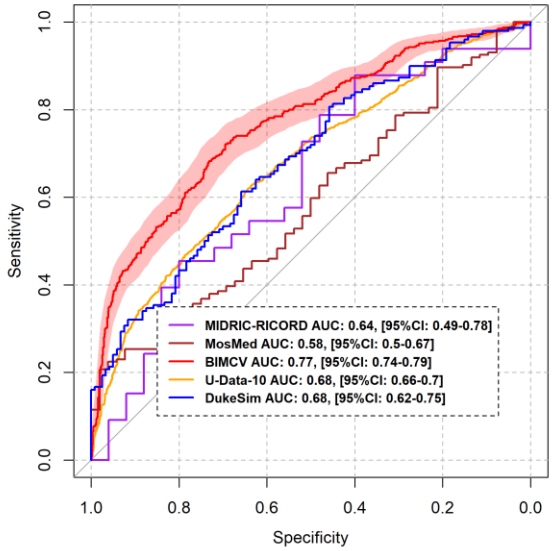
Model MIDRIC-RICORD



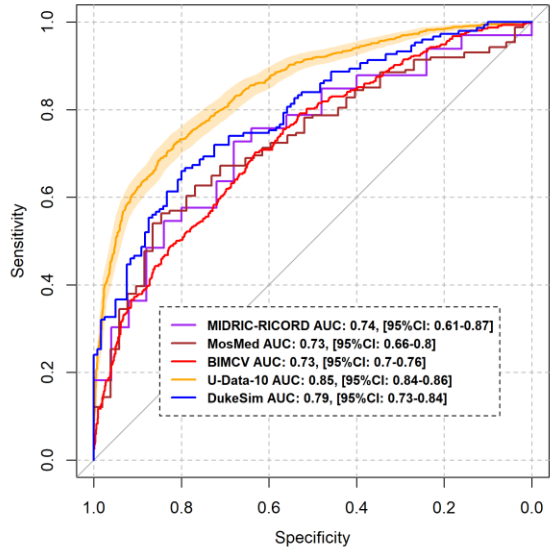
Model MosMed



Model BIMCV

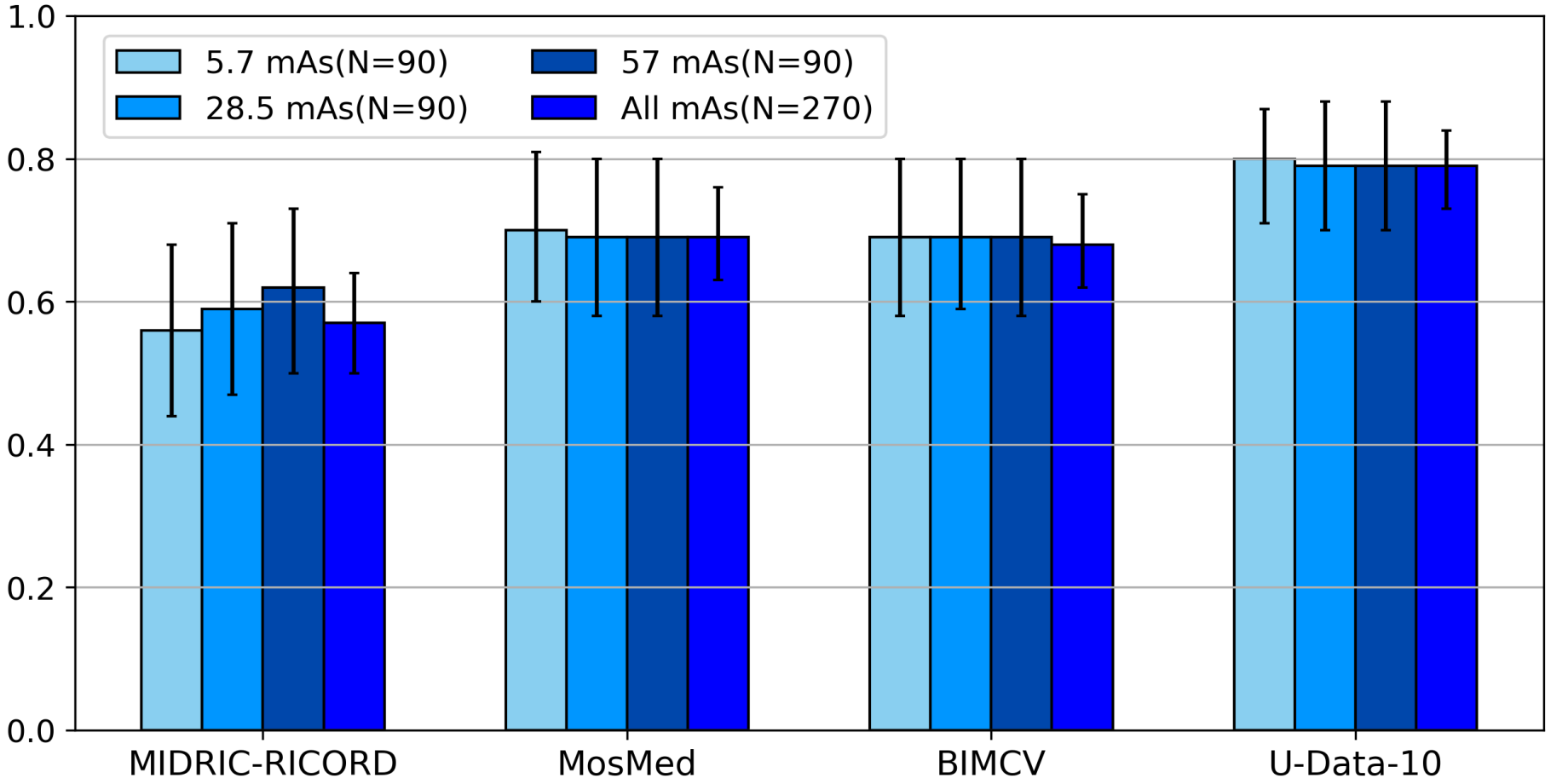


Model U-DATA-10

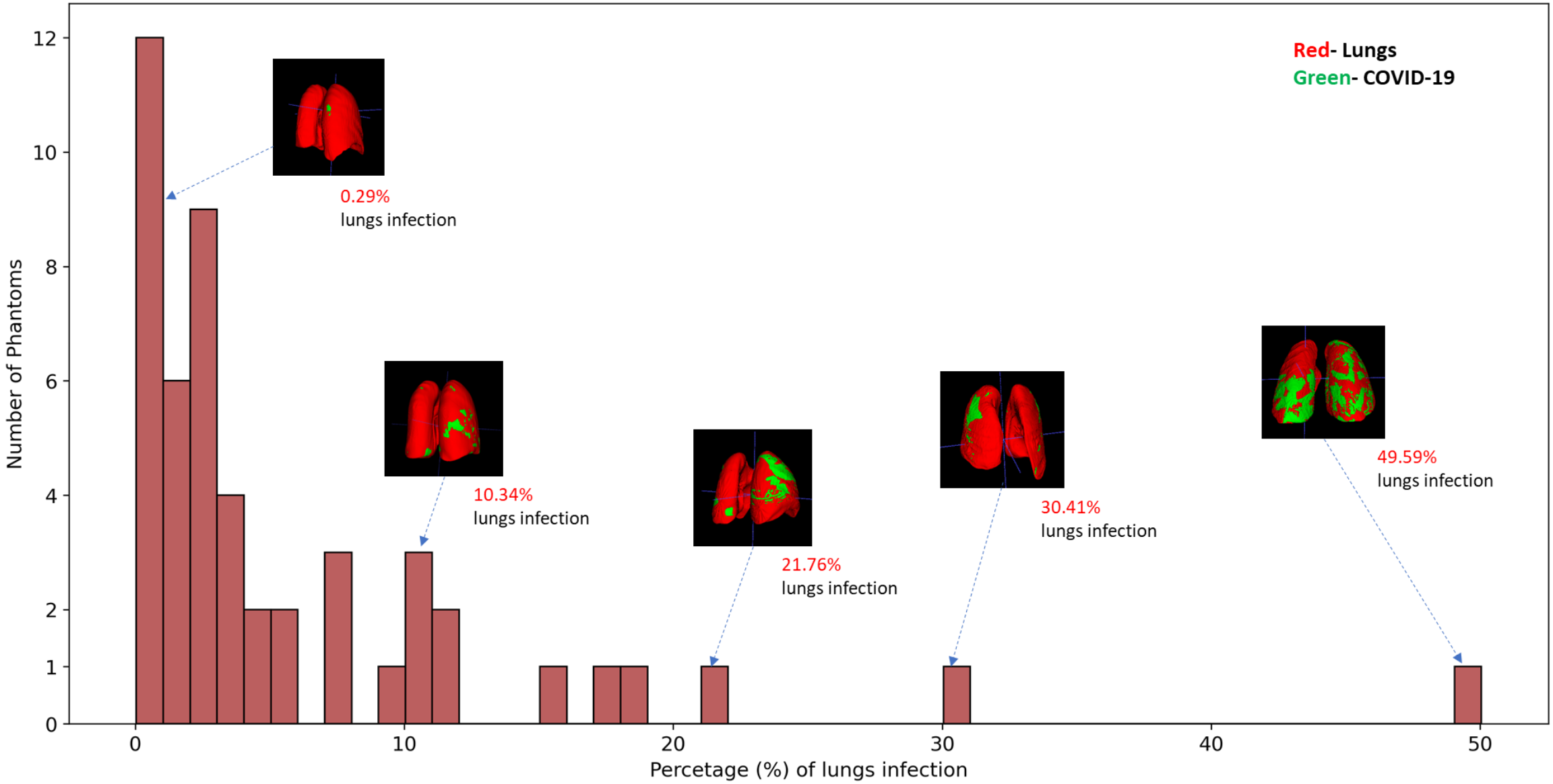


Results | Dose-level Analysis

No dose dependence for all models tested on simulated data



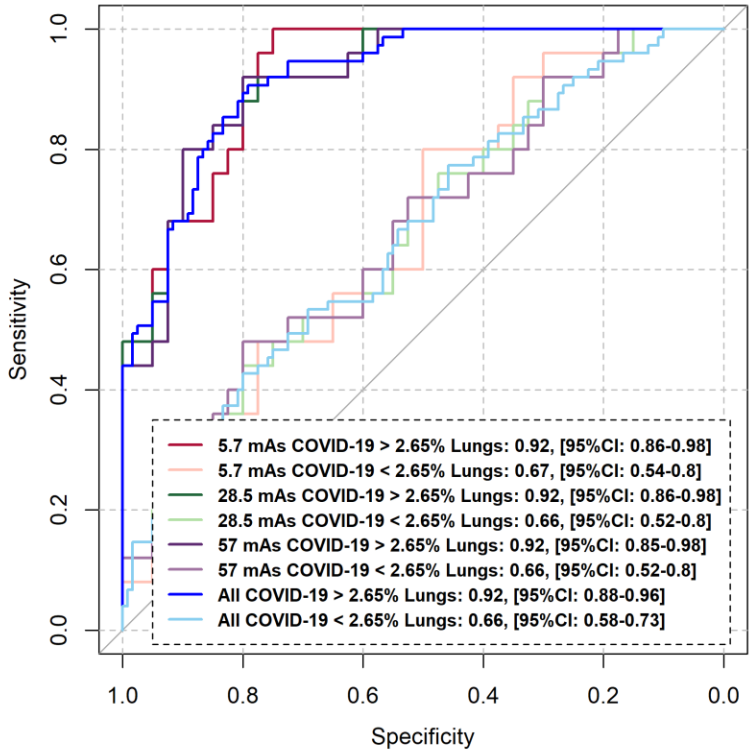
Results | COVID-19 Infection Size



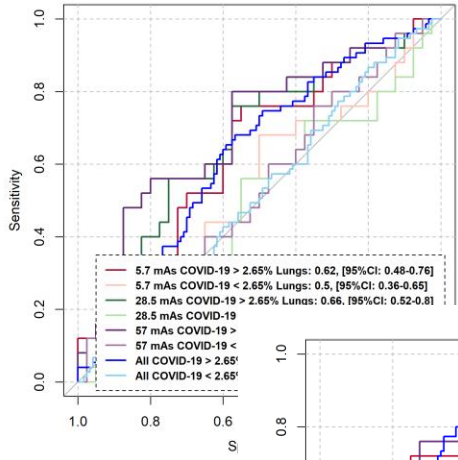
Results | COVID-19 Infection Size

Infection volume strongly affects performances across all models and doses

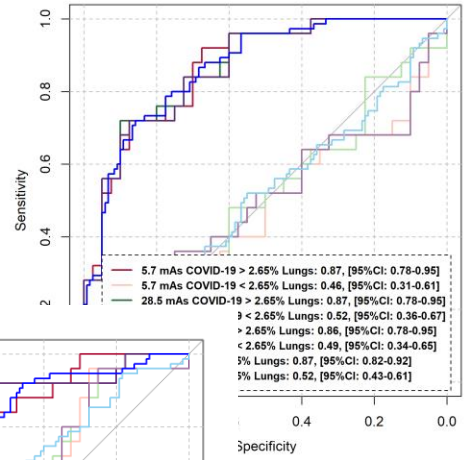
Model U-DATA-10



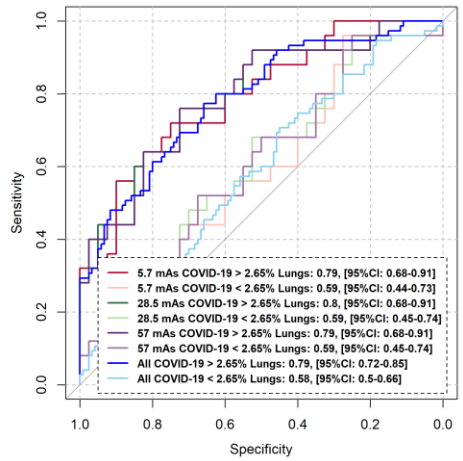
Model MIDRIC-RICORD



Model MosMed



Model BIMCV



Conclusion

- Performance dropped due to domain-shift, simulated data consistent with multiple clinical test set.
- No dose dependence for all models tested on simulated data.
- Infection volume strongly affects performances across all models and doses
- Virtual Imaging Trials make it possible to answer clinically relevant questions.

CVIT: A growing community…

tushar.ece@duke.edu