

## Radiomics Cluster

The University of Chicago/Rush University CTSA – NCATS grant UL1TR002389

Maryellen L. Giger, PhD; Samuel Armato III, PhD; Chun-Wai Chan, MS; John Papaioannou, MS; Heber MacMahon, MB, BCh; Julian Solway, MD

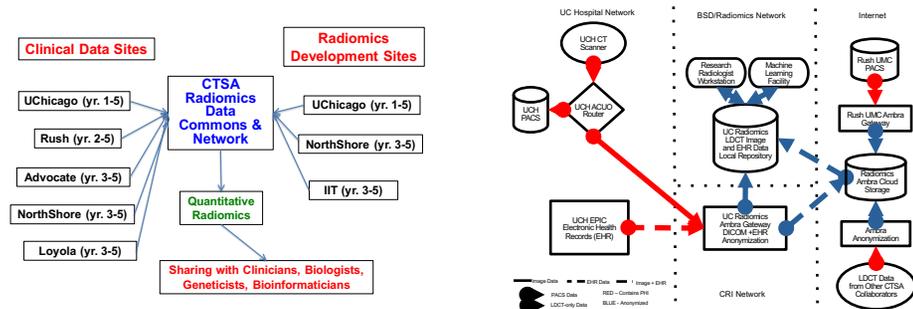
The Institute for Translational Medicine (ITM) aims to make research breakthroughs happen and to put those discoveries into the real world to improve health. The ITM is a partnership between the University of Chicago and Rush in collaboration with Advocate Health Care, the Illinois Institute of Technology, Loyola University Chicago, and NorthShore University HealthSystem.

The goal of the ITM Radiomics Cluster is to enhance and support CTSA Consortium-wide studies through the development and sharing of databases, quantitative radiomic methods & software, and algorithmic outputs for the purpose of discovery and translation through big data science. Radiomics deals with pulling data from images, such as CT and MRI scans, that can be analyzed and used to make diagnoses quantitative, faster, and better. The cluster works to bring imaging and big data studies to the next level by coordinating databases, software, computer calculations, and more across all six ITM institutions.

### Aim 1

Develop a multi-institutional radiomics infrastructure – a platform including processes and implementation for data collection, data annotation, data distribution, quantitative radiomics (QR) methodology and software, and data sharing – for the clinical institutions within the ITM (UChicago, Rush, NorthShore, Loyola, and Advocate). We have collected and are currently annotating hundreds of CT scans from three of the institutions.

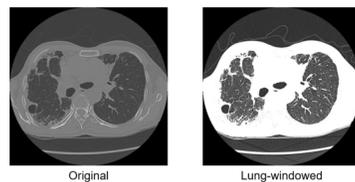
#### CTSA Radiomics Data Commons for Low-Dose CT



### Aim 2

Demonstrate the utility of the Radiomics infrastructure focusing on a variety of diseases assessed on thoracic CT images in order to show "how to effectively and efficiently do multi-disease radiomics research." We aim to understand the added value arising from quantitative radiomics of incidental findings on low-dose CT screening for lung cancer.

- Radiomics of interstitial lung and pleural diseases
- Quantitative assessment of cardio-vascular risk status
- Radiomics of osteoporosis
- Radiomics assessment of liver diseases



### Aim 3

Disseminate the knowledge gained, databases, radiomics, and algorithms via a novel CTSA Radiomics Website so that other CTSA and investigators can initiate or extend their radiomics programs. CTSA Radiomics Data Commons will be integrated and accessible from the ITM website (<http://chicagoitm.org>).

For specific medical imaging tasks, detailed procedure lists will be given so that imaging technologists and clinical coordinators can effectively and efficiently either deposit imaging data for later analysis or run developed algorithms within the clinical workflow.

NIBIB COVID-19 Contract 75N92020D00021

An Open Discovery Data Repository

--- creation, testing, quality assurance, diversity, and data connectivity

Machine Intelligence Computational Capabilities

--- clinically relevant algorithms and software tools

Medical images play an essential clinical role in the detection, triage, diagnosis, management, and therapy monitoring of COVID-19 because they provide unique, definitive, and quantitative information about the lungs and other organs, which are the most impactful targets of the virus.

The Medical Imaging and Data Resource Center (MIDRC) is a multi-institutional initiative driven by the medical imaging community and aimed at accelerating the transfer of knowledge and innovation in the current COVID-19 pandemic. MIDRC, funded by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and hosted at the University of Chicago, is co-led by the American College of Radiology® (ACR®), the Radiological Society of North America (RSNA), and the American Association of Physicists in Medicine (AAPM). The PI of the MIDRC NIBIB contract, M. Giger, is applying her experience from the ITM Radiomics Cluster to the MIDRC project. Also, MIDRC is discussing collaboration with NCATS N3C to effectively and efficiently link imaging and other COVID-19 clinical data to enhance and expedite research findings. The aim of MIDRC is to foster machine learning innovation through data sharing for rapid and flexible collection, analysis, and dissemination of imaging and associated clinical data by providing researchers with unparalleled resources in the fight against COVID-19.

The MIDRC infrastructure and processes is being created through five Technology Development Projects, which are being conducted collaboratively:

1. Creating an open discovery platform for COVID-19 imaging and associated data (led by RSNA).
2. Creating a real-world testing and implementation platform with direct real-time connections to health care delivery organizations (led by ACR).
3. Developing and implementing quality assurance and evaluation procedures for usage across the MIDRC (led by AAPM).
4. Enabling data intake, access and distribution via a world-facing data commons portal (led by all three plus Gen3).
5. Linking the MIDRC to other clinical and research data registries (led by all three plus Gen3).

This multi-institutional initiative, involving 23 institutions and the FDA, represents a partnership spearheaded by the medical imaging community aimed at accelerating the transfer of knowledge and innovation, including clinical problem identification, discovery, development, evaluation, translation, implementation, and dissemination. The common goal of this coalition is to build data repositories to fuel COVID-19 machine intelligence research, coupled with optimal standardization, curation, and compliance with ethical responsibilities to honor patients' privacy. In order to leverage existing infrastructure, MIDRC will be a linked collection that coordinates access to data and harmonizes data management activities across all participating organizations at three critical stages: (1) intake, including curation, de-identification, abstraction, and quality assessment (2) annotation and labelling of images and other data using semi-automated approaches and (3) distributed access and query methods. These methods will yield a large data set that is in accordance with the FAIR principles (findable, accessible, interoperable and reusable). The data will be collected from various sources including academic medical centers, community hospitals, and others.

The public access "front door" of MIDRC is hosted by the state-of-the art Gen3 Data Ecosystem housed at the University of Chicago and will be expanded to link with additional image and non-image data feeds from multiple registries and repositories. Through the MIDRC Data Commons Portal, images and data, as well as guidelines and recommendations, are disseminated to investigators to expedite research that provides solutions to the COVID-19 pandemic to ultimately maximize patient benefit.

MIDRC also developed twelve research projects to expedite translation of machine intelligence from scientific findings and technical resources to public dissemination and clinical benefit. The investigators were selected from current members of the RSNA-RIC, ACR-DSI, or AAPM-DSC, thus, effectively spanning the nation and spanning the medical imaging community.