Medical Imaging and Data Resource Center: Covid-19 and beyond

Rui C. Sá, Ph.D.
Data Scholar
NIBIB & ODSS

Assistant professor
University of California, San Diego
Quantitative functional imaging with MRI
Gas Exchange in the human lung

Feet

Head

gravity

Ventilation ($\dot{V}_A$)

ml/min/ml

Perfusion ($\dot{Q}$)

ml/min/ml

$\dot{V}_A / \dot{Q}$

R.C. Sá et al, JAP123(1), 2017
COVID-19 pandemic response

Role for AI / ML

- Identify infection
- Diagnose disease
- Assess extent
- Monitor therapy
- Detect complications
- Predict outcome
Medical Imaging and Data Resource Center
Rapid Response to Covid-19 Pandemic
Medical Imaging and Data Resource Center
Rapid Response to Covid-19 Pandemic

NIBIB supported resource for medical imaging (CXR, CT)

**Goal:** collect and curate medical images with adjunct clinical data and develop artificial intelligence (AI / ML) methods to aid in the analysis & interpretation of medical images in response to Covid-19 pandemic

**Two Major Scientific Components**

**Creation of Open Discovery Data Repository:**
5 Technology Development Projects

**Machine Intelligence Computational Capabilities:**
12 Collaborative Research Projects
And multiple trans-MIDRC scientific workgroups
Achievements: Infrastructure and standardization of processes

• Infrastructure & standardization:
  • Harmonization of data ingestion, quality control, data flow, common data model, de-identification procedures,...
Achievements: Infrastructure and standardization of processes

MIDRC Dashboard

- **Total ingested into MIDRC**
  - # of Imaging Studies: 62,595

- **Undergoing MIDRC Data Quality and Harmonization**
  - # of Imaging Studies: 52,127

- **Released by MIDRC**
  - # of Imaging Studies: 10,468

- **AI / ML Development**
  - > 24 Algorithms

- **Current MIDRC data (Chest Xray & CT)**

- **Quality assessment**
  - Diversity assessment
  - Sequestration - algorithm validation (FDA)

- **Publicly available**
  - curated, high quality, diverse and representative imaging studies

- **AI / ML Development**
  - e.g.:
    - Segmentation: Lung, opacities
    - Severity & length of hospital stay
Critical gaps in AI/ML deployment
Lack of diverse and representative data

Geographic Distribution of Data to Train AI Algorithms


https://doi.org/10.1148/radiol.2019190613


“...report all results by relevant clinical and demographic group...”

Need for representative dataset
Achievements: Diversity and Representativity of MIDRC Data

Bar chart showing the distribution of different racial and ethnic groups in MIDRC DATA and Census 2020.

- White: 70% (MIDRC DATA), 65% (Census 2020)
- Black or African American: 20% (MIDRC DATA), 15% (Census 2020)
- American Indian or Alaska Native: 5% (MIDRC DATA), 3% (Census 2020)
- Asian: 15% (MIDRC DATA), 10% (Census 2020)
- Native Hawaiian and Other Pacific Islander: 10% (MIDRC DATA), 7% (Census 2020)
- Other, Two or more: 5% (MIDRC DATA), 3% (Census 2020)

Map showing 23 states with data represented.
Medical Imaging and Data Resource Center: Collaborations

NIH - National Center for Advancing Translational Sciences

N3C

NIH - National Heart, Lung, and Blood Institute

BioData CATALYST

NIH - National Institute of General Medical Sciences

IDeA-CTR

Medical Device Research Interest Group

Trans-NIH, FDA, CMS

AIAA

Medical Device Innovation Consortium
Achievements: algorithms by MIDRC investigators

• Algorithms (two examples)
  • Extracting information from radiology reports

  Cascaded deep transfer learning on thoracic CT in COVID-19 patients treated with steroids

  Jordan D. Fuhrman, a Jun Chen, b Zegang Dong, c Fleming Y. M. Lure, c Zhe Luo, de,* and Maryellen L. Giger, e a,*

• Sequestration:
  • Diverse & representative sequestered dataset as a tool for independent testing of AI/ML algorithms for regulatory/translation.
Immediate future

• **Collect and disseminate Covid-19 medical imaging** data for discovery and technology deployment
  • Expand from chest CXR & CT to **other organs/systems** (heart, brain, ...)
  • **Increase the range of modalities** (MRI, ultrasound, ...)

• **Support Post Acute Sequelae of SARS-CoV-2 infection (PASC)** response, including longitudinal monitoring

• Expand to other **acute and chronic diseases**

• Develop, validate and **deploy AI/ML algorithms for medical imaging**
Sustainability: MIDRC as a National resource

- Support the Medical Imaging AI/ML ecosystem
  - High quality, representative, trustworthy data
  - Culture of collaboration
  - Promote standards, sharing, transparency, best practices
  - Lower barrier of access

- Accelerate translation of AI/ML
  - Real-world quantification of algorithm performance (sequestered dataset)
It takes a village…

NIH National Institute of Biomedical Imaging and Bioengineering
- Kris Kandarpa, MD, PhD
- Behrouz Shabestari, PhD
- Guoying Liu, PhD
- Julia Ringel
- Qi Duan, PhD

Office of Data Science Strategy
- Natasha Hurwitz, MS
- Allissa Dillman, PhD
- DATA Scholars cohort, in particular
  - Judy W. Gichoya, MD
  - Mohammad Ghassemi, PhD

NIH National Institutes of Health
Turning Discovery Into Health

MIDRC Medical Imaging and Data Resource Center.
- Maryellen L. Giger, PhD
- Paul E. Kinahan, PhD
- Etta Pisano, MD

UC San Diego Pulmonary Imaging Laboratory
NIH funding:
  - NHLBI, NIA
NASA funding (NSBRI)
Amazon AWS grant
- Michael Tilkin, MS
- Curtis P. Langlotz, MD, PhD
- Adam Flanders, MD
- Robert L. Grossman, PhD