Federated Learning in Medical Imaging: Enhancing Data Privacy and Advancing Healthcare

John W Garrett, PhD, DABR

Associate Professor of Radiology, Medical Physics, and Biostatistics and Medical Informatics

Director of Imaging Informatics

Department of Radiology
UNIVERSITY OF WISCONSIN
SCHOOL OF MEDICINE AND PUBLIC HEALTH

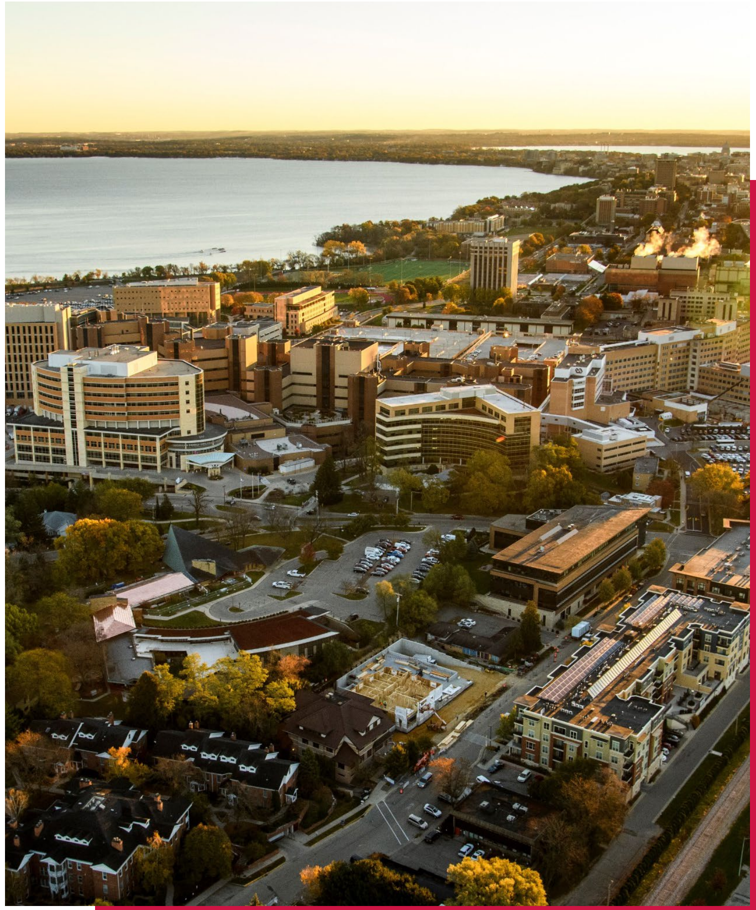
Disclosures

Dr. Garrett is a shareholder in NVIDIA

Dr. Garrett holds equity in and serves on the advisory board of RadUnity, Inc.

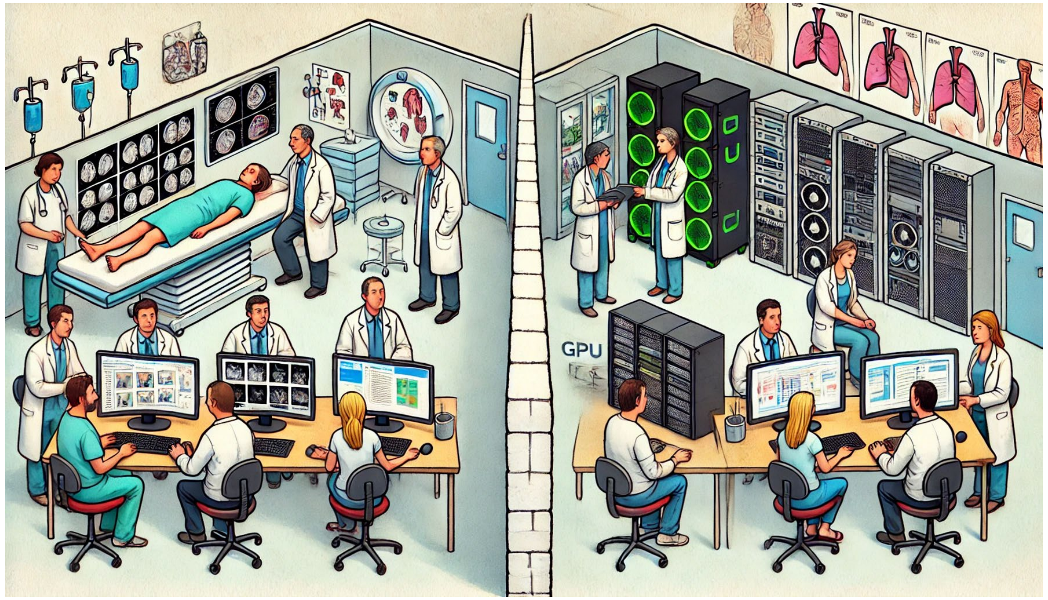
Dr. Garrett is a member of the SIIM Machine Learning Tools and Research Subcommittee

Dr. Garrett is a member of the AAPM Imaging Informatics Committee



Unique Needs in Healthcare Federated Learning can address

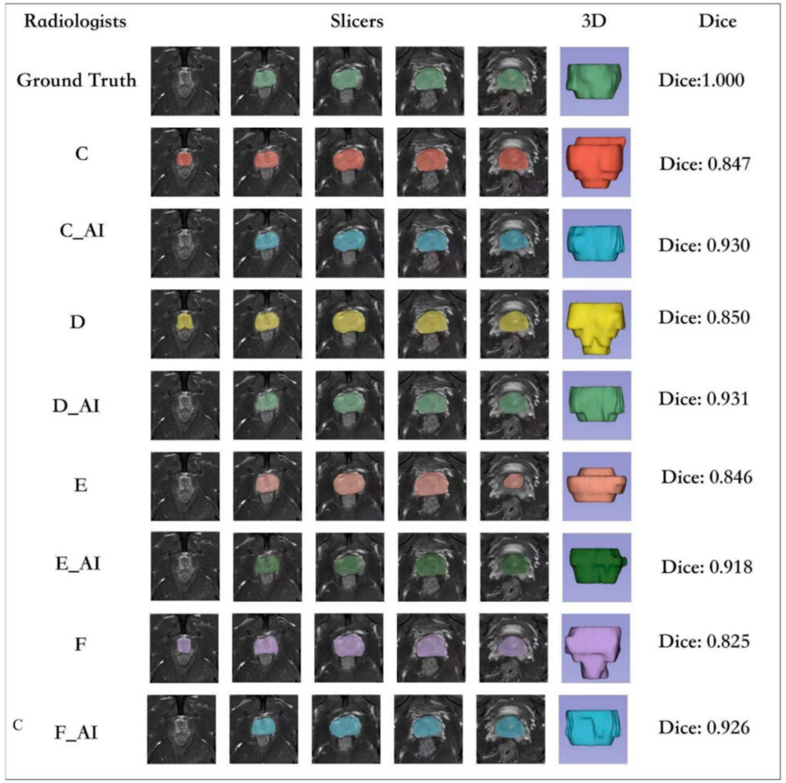
- **Privacy Concerns**
  - Strict regulations (HIPAA/GDPR) limit data sharing.
  - Federated learning enables collaborative research without compromising patient privacy.
- **Data Availability is limited**
  - Data sharing:
    - Extremely difficult and time consuming to get data sharing agreements in place and approved
  - Rare or Novel Diseases:
    - Limited data across individual institutions.
    - Pediatric Imaging: Challenges in gathering large datasets due to fewer cases and stricter regulations.



Generated with DALL-E 2

Challenges in Leveraging Federated Learning for Medical Imaging

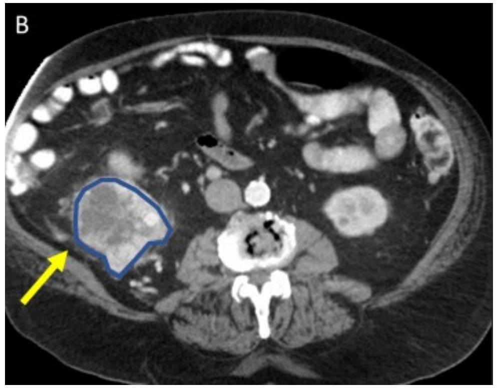
- **Data Harmonization**
  - Variability in imaging protocols and equipment across institutions.
  - Need for standardized data formats and pre-processing techniques to ensure compatibility.
  - Data labels and definitions may vary site to site
- **Compute Availability**
  - Large compute clusters often operate outside of PHI-compliant environments, complicating direct use with medical data.
  - In secure environments may have limited resources
  - May have strict firewalls and security making accessing a federated learning server a challenge

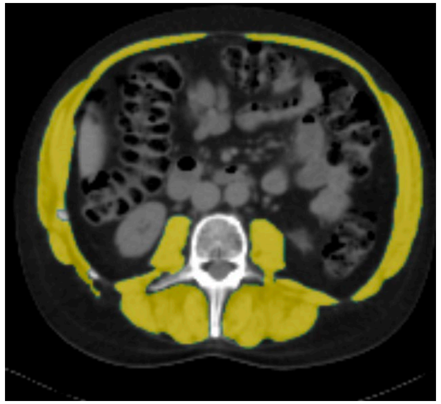


Jin, L.; Ma, Z.; Li, H.; Gao, F.; Gao, P.; Yang, N.; Li, D.; Li, M.; Geng, D. Interobserver Agreement in Automatic Segmentation Annotation of Prostate Magnetic Resonance Imaging. *Bioengineering* **2023**, *10*, 1340. https://doi.org/10.3390/bioengineering10121340

Applications of Federated Learning in Medical Imaging

- There are numerous applications for federated learning in medical imaging
- Today I will highlight a few we've worked on that showcase the utility of the method:
    - Deep learning to predict patient needs for novel disease: Covid 19
    - Tumor segmentation for rare tumor types: AI Assisted annotation for improved data labeling: Renal cell carcinoma
    - Improved access to compute for experimentation and development: muscle segmentation

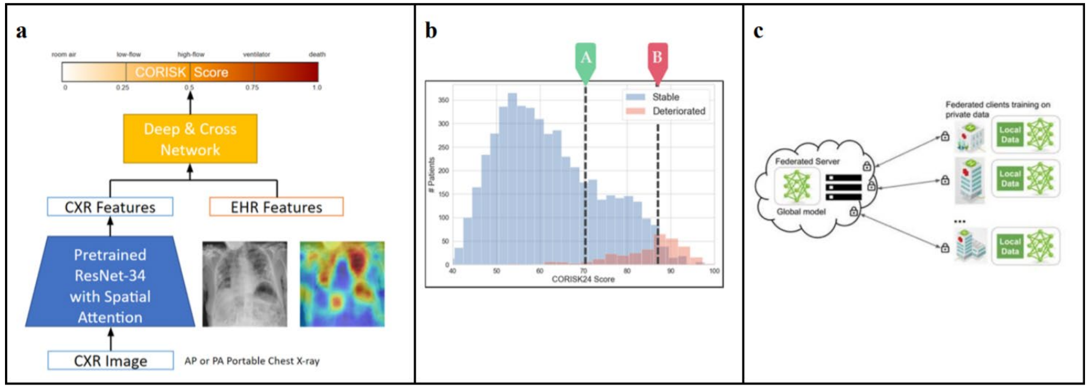


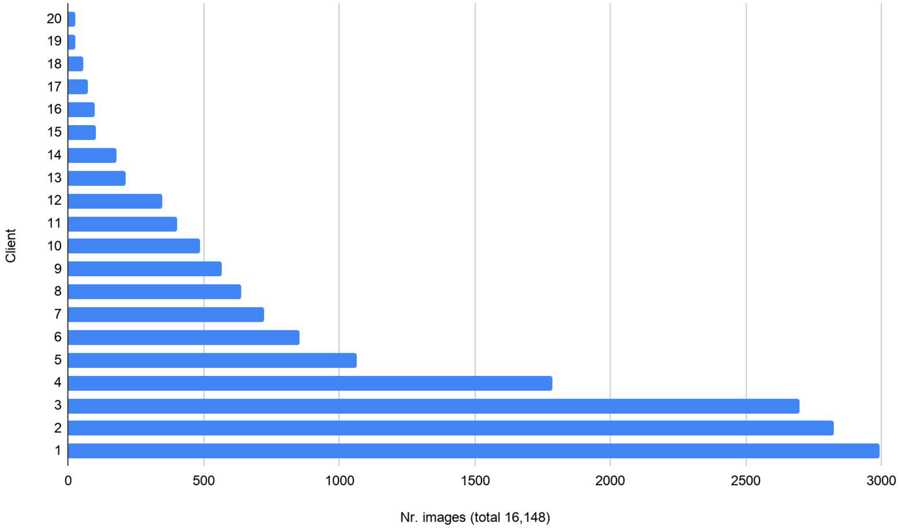


EXAM (EMR CXR AI Model) initiative[1]

- The goal was to develop a model to predict oxygen needs for Covid patients as a quick response to pandemic: EXAM (EMR CXR AI Model) initiative[1]
    - Challenges:
        - Hard to put a large diverse dataset together for this at a single site
        - Getting data sharing agreements in place is not practical/possible for quick response

Solution: Leverage federated learning (Clara at the time) to quickly build a useful model





1. Dayan, I., Roth, H.R., Zhong, A. *et al.* Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat Med* **27**, 1735–1743 (2021). https://doi.org/10.1038/s41591-021-01506-3

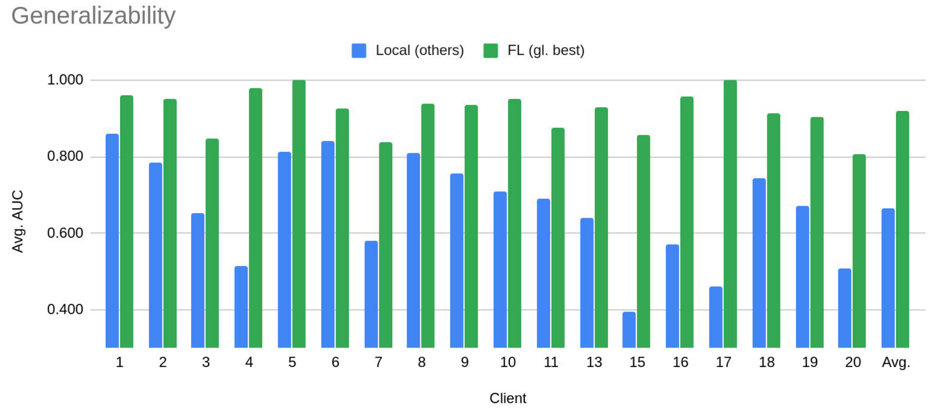
EXAM (EMR CXR AI Model) initiative

Federated learning initiative with 20 global institutions and industrial support resulted on average in

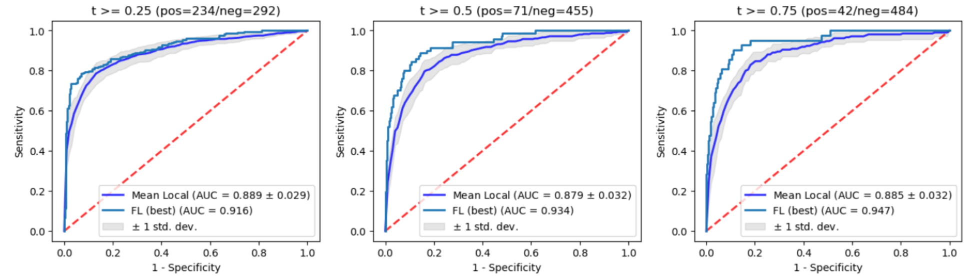
16% performance improvement

38% generalizability improvement

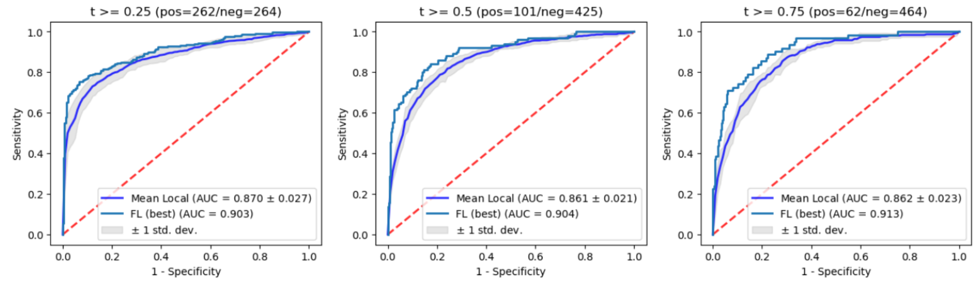
<u>Accomplished without requiring site-to-site data sharing agreements or for any data to move outside our hospital firewall</u>



24h Prediction for COVID positive patients



72h Prediction for COVID positive patients



AI Assisted annotation project with SIIM ML Tools and Research Committee

- One of the major challenges in distributed learning like this however is the variability in data labels, particularly for image segmentation tasks where the tools and rules used to generate the training masks may vary dramatically from site to site.
  - This both reduces overall performance of the final trained model and potentially requires more datapoints for training.
- Recently, AI assisted annotation tools have been developed that may help provide a consistent and efficient mechanism for data to be labeled.
  - Used to great effect in FeTS study but without a control[1]

Hypothesis: Using AI Assisted annotation will help provide a more consistent dataset for federated learning training improving performance or reducing the number of cases needed.

SIIM committee has been leading FL initiatives since 2019[2,3,4]

Team:

Committee Co-Chairs:

Khaled Younis, PhD

Yuankai Huo, PhD Vanderbilt University

Project lead: John Garrett, PhD University of Wisconsin Madison

Members:

Farzana Ali University at Buffalo

Gian Marco Conte, MD, PhD Mayo Clinic - Rochester

Abdussalam Elhanashi University of Pisa

Ahmed Elshaikh Mustafa, MBBS

Shahriar Faghani, MD Mayo Clinic Rochester

Amilcare Gentili San Diego VA Health Care System

Shiba Kuanar, PhD Mayo Clinic - Rochester

Andrew Missert Mayo Clinic - Rochester

Ghulam Rasool, PhD Moffitt Cancer Center

Yashbir Singh, PhD Mayo Clinic - Rochester

Joseph Yacoub, MD Medstar Georgetown University Hospital

1. Pati, S., Baid, U., Edwards, B. *et al.* Federated learning enables big data for rare cancer boundary detection. *Nat Commun* **13**, 7346 (2022). https://doi.org/10.1038/s41467-022-33407-5
2. CMIMI 2020: V. Nath, et al., Empirical Evaluation of Federated Learning for Classification of Chest X-Rays, Society for Imaging Informatics in Medicine Conference on Machine Intelligence in Medical Imaging, 2020 (Thanks Yuankai for sharing the abstract)
3. ECR Oral Presentation 2021: K. Younis, et al., Enabling privacy-preserving training of deep learning models for radiological applications using federated learning, European Congress of Radiology, 2021 (I presented and can dig up the slides).
4. AI Model Winners' showcase at SIIM 22. K. Younis, et al., AI Model Winners' Showcase: Federated Learning for CXR disease classification, Cleveland, OH, The Society for Imaging Informatics in Medicine (SIIM) Annual Meeting, 2022 (**Link to all Models:** **https://siim.org/page/siim21w_breaking_ai_boundaries**)

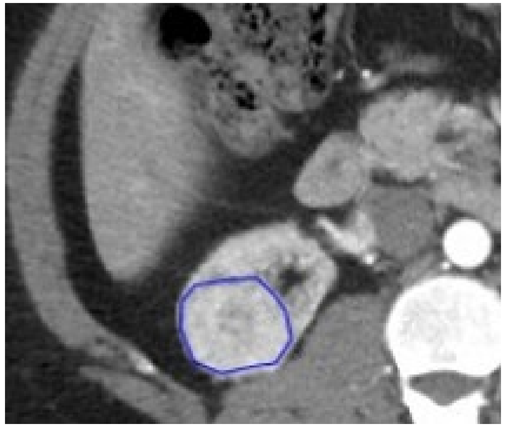
AI Assisted annotation project with SIIM ML Tools and Research Committee: Goal

- The goal of this project is to evaluate whether AI assisted annotation can facilitate federated learning for an automated image segmentation task either by improving overall model performance, reducing the number of samples required for training, or both.

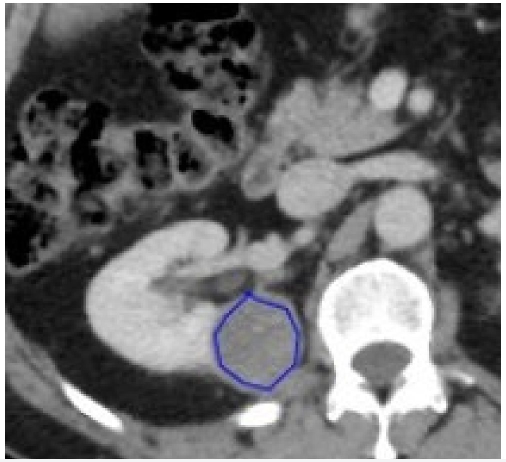
Renal cell carcinoma or RCC selected for this project. Why RCC?

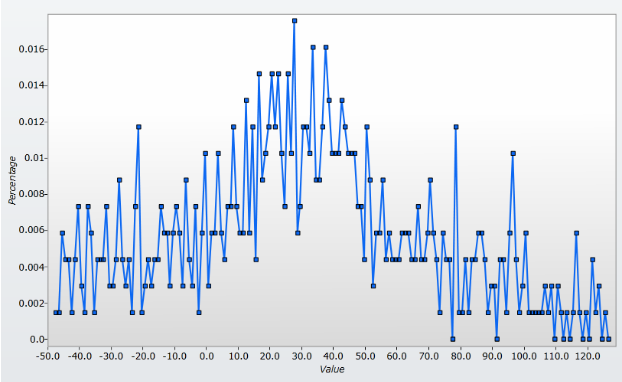
- RCC are commonly seen and yet resected masses are often found to be benign.
- Once segmented, tools such as radiomic texture analysis can help provide immediate feedback, helping distinguish subtype or malignancy vs benign

Clear cell RCC



Papillary RCC





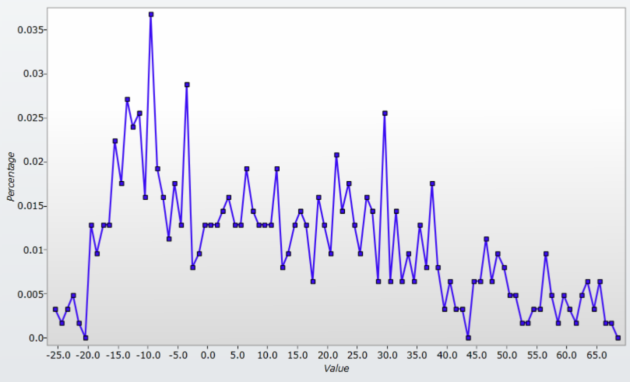
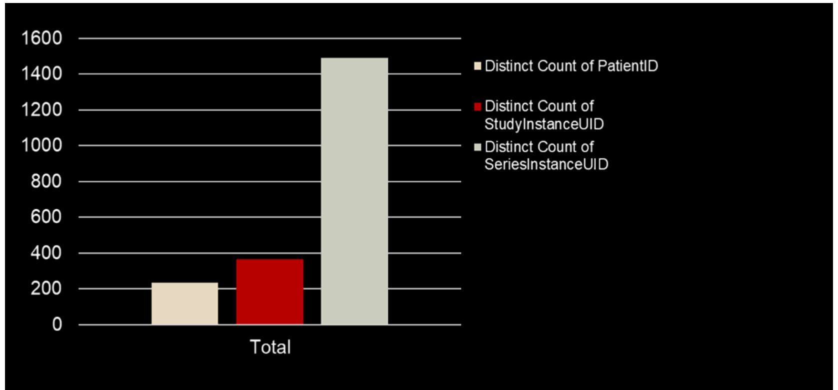


Figure from Dr. Meghan Lubner at UW

AI Assisted annotation project with SIIM ML Tools and Research Committee: IDC and TCIA

- A large, curated imaging dataset for RCC has been made publicly available as part of IDC: TCGA-KIRC - The Cancer Imaging Archive (TCIA) [1]
- This dataset includes MR and CT imaging along with clinical data, outcomes, patient follow-up, etc.
  - This incredibly rich dataset contains all the imaging and clinical data needed to build a model like this and to test this project's hypothesis, but…
  - A segmentation model requires image level annotation/segmentation of masses to provide a useful training dataset.



Row Labels	Distinct Count of PatientID	Distinct Count of StudyInstanceUID	Distinct Count of SeriesInstanceUID
GE MEDICAL SYSTEMS	182	221	861
Philips	8	8	51
SIEMENS	56	136	581
Grand Total	236	365	1493

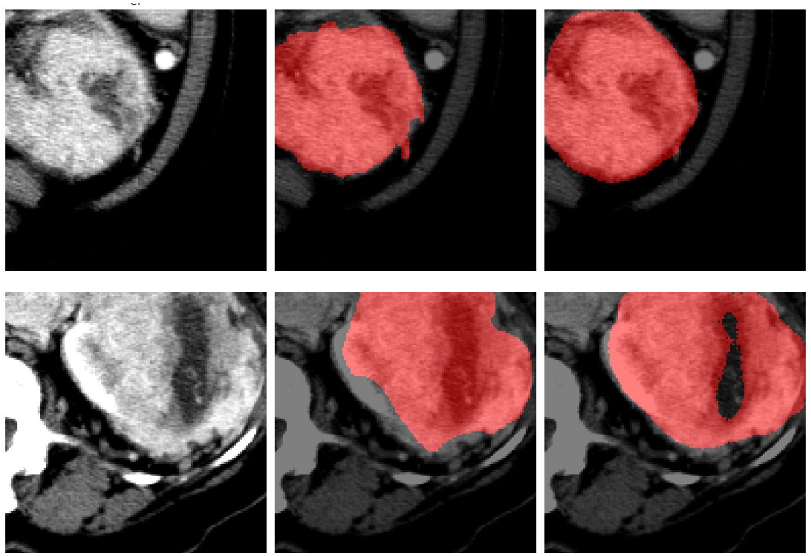
Row Labels	Distinct Count of PatientID	Distinct Count of StudyInstanceUID	Distinct Count of SeriesInstanceUID
(blank)	6	6	24
IV	38	53	121
null	146	240	746
Oral	77	83	105
Oral & IV	136	145	497
Grand Total	236	365	1493

	Distinct Count of SeriesInstanceUID
<=1mm	170
(1mm,3mm]	578
(3mm,5mm]	617
>5mm	128
Grand Total	1493

1. Akin, O., Elnajjar, P., Heller, M., Jarosz, R., Erickson, B. J., Kirk, S., … Filippini, J. (2016). Radiology Data from The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma [TCGA-KIRC] collection. The Cancer Imaging Archive. http://doi.org/10.7937/K9/TCIA.2016.V6PBVTDR (Public Access - Cancer Imaging Archive Wiki: https://wiki.cancerimagingarchive.net/display/Public/TCGA-KIRC)

AI Assisted annotation project with SIIM ML Tools and Research Committee: Training and Initial Results

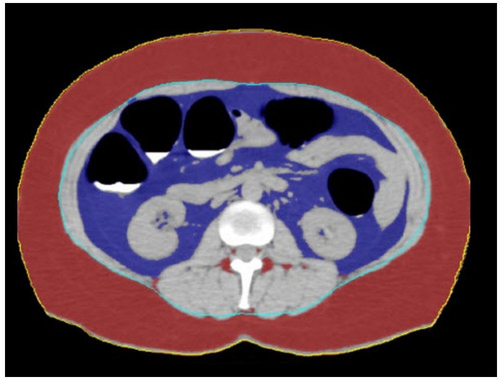
- All annotation is complete and training is currently underway
  - Training different models (Unet, Swin UNETr etc)
- Target output for the project:
  - Sample code and documentation for performing federated learning
  - Further annotated dataset
  - RCC segmentation model
- Full training of AI Assisted annotations and final analysis should be completed soon



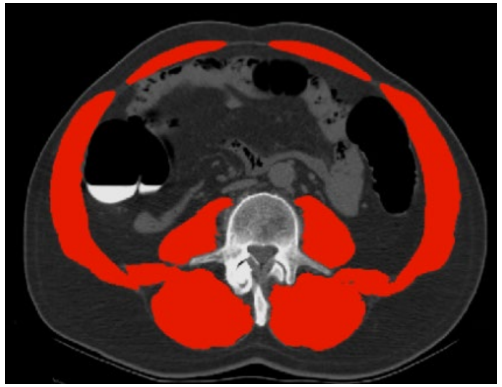
Automated CT Biomarkers for Opportunistic Screening

- Analogous manual, semi-, and fully-automated measures
- Can be applied to nearly any abdominal (or chest) CT
- Adjustments for IV contrast and other acquisition parameters

- Practically can't measure everything…
  - Need to select small number of automated CT parameters (1-2) from each of the abdominal components for predicting future events



V/S Fat Ratio at L3 level



Mean Muscle HU at L3 level

Aortic Agatston Ca^{++} score

Mean Volumetric Liver/Spleen HU

Trabecular HU at L1 level

Relevance of AI in opportunistic screening

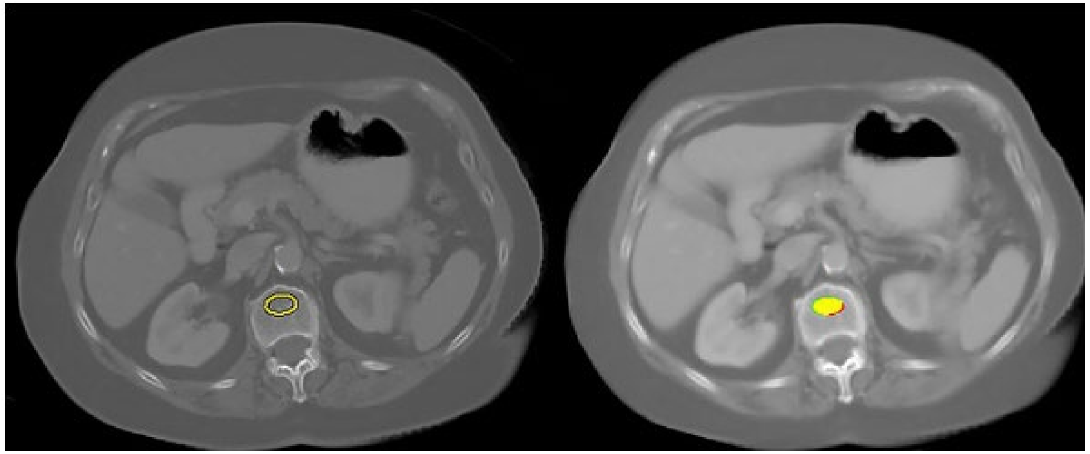
- Almost all commonly used biomarkers are “explainable” in the sense that they are straightforward measurements from identified tissue types or organs
  - These are segmentations that could be performed by a radiologist but are in practice far too tedious
  - The recent emergence of deep learning methods has enabled highly accurate, fast, and automated segmentations

Radiology: Artificial Intelligence

ORIGINAL RESEARCH

Improved CT-based Osteoporosis Assessment with a Fully Automated Deep Learning Tool

Perry J. Pickhardt, MD • Thang Nguyen, MD • Alberto A. Perez, MD • Peter M. Graffy, BA, MPH[1] • Samuel Jang, MD • Ronald M. Summers, MD, PhD • John W. Garrett, PhD



Workflow: Opportunistic screening with biomarkers

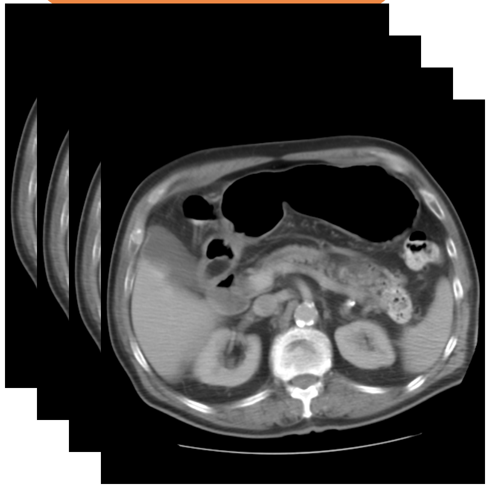
Step 0/1: Identify eligible series and normalize and pre-process input CT images

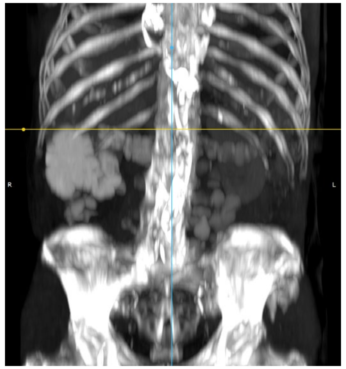
Step 2: Identify relevant anatomical landmarks (vertebral bodies and aortic hiatus/bifurcation)

Step 3: Automatically segment organs/tissues of interest

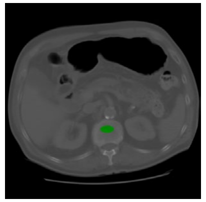
Step 4: Make measurements (density, volumes, cross sectional area, etc.)

Step 5: Develop and plug values into models to predict outcomes and risk-stratify patients

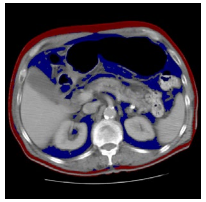




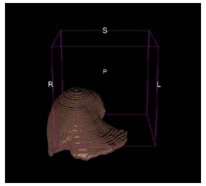
Trabecular bone in L1 vertebrae



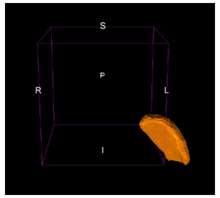
Fat
Visceral and subcutaneous



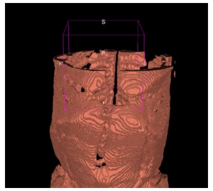
Liver



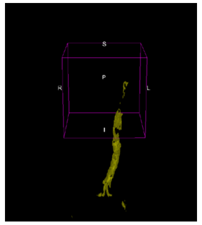
Spleen

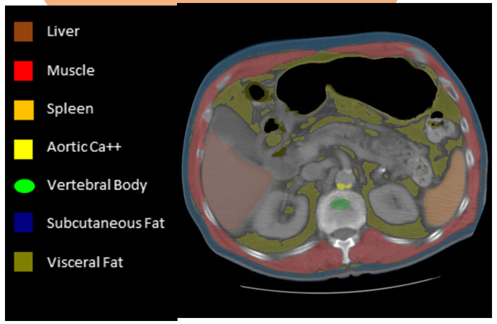


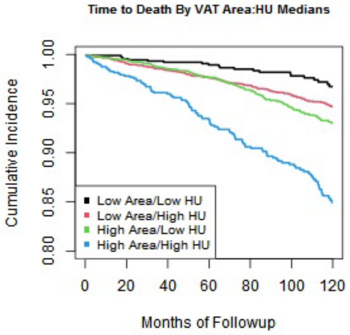
Muscle



Aortic Calcium







Muscle Values

L1MuscleMeanHU	30.4
L1MuscleArea	117
L3MuscleMeanHU	30.2
L3MuscleArea	174.5

Fat Values

L3TATArea	245.6
L3VATArea	132.6
L3SATArea	113
L3VATSATRatio	1.17

Spleen Values

SpleenMedianHU	83.8
SpleenVolume	255

Calcium Scoring

AbdominalAgatston	31809

Bone Density

BMD (HU)	94.2

Liver Values

LiverMedianHU	93.8
LiverVolume	1878

Federated learning for distribution of compute in a heterogeneous compute environment

- Challenge:
  - Datasets and compute resources required to build modern computer vision models are massive:
    - For opportunistic body composition segmentation tools for each organ getting measured and these should be refined and updated as new technology emerges → Many organs and many models!
  - Good, curated datasets are very difficult to fully anonymize and move around freely
  - In ePHI environment, compute may be limited or scattered (rarely a full-fledged cluster)
  - Project purpose: Use FLARE to distribute compute for a large computer vision task to speed up training and facilitate training many models

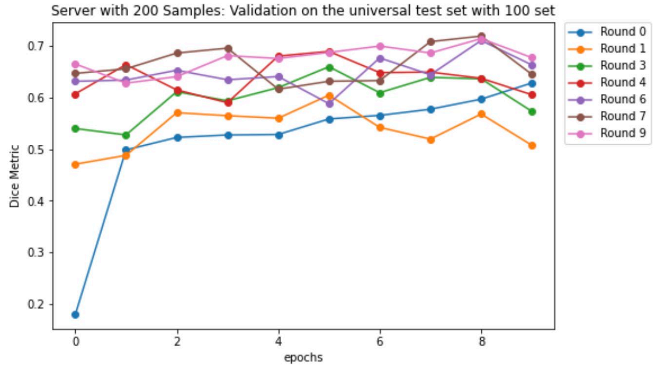
Federated learning for distribution of compute in a heterogeneous compute environment

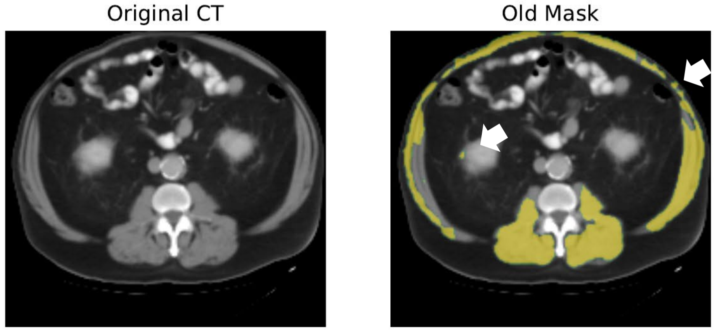
- Methods:
  - Six simulated sites with 20,000 CT images for muscle segmentation from diverse UW CT datasets including routine CT colonography exams, outside studies read at UW, and random UW exams
  - Using MONAI and 5 model classes with an older model for reference:
    - Previously validated/published U-Net Model from our group
    - **UNET:** U-Net trained on UW Curated cohort w/ fed. Learning
    - **Standard SwinUNETR:** U-NETr (State of the art transformer based U-Net) trained on UW Cohort (https://arxiv.org/abs/2103.10504)
    - **Custom SwinUNETR:** This model is essentially a generator model like a GAN (https://arxiv.org/abs/1406.2661) where the encoder layer is a 3D SWIN Transformer (https://arxiv.org/abs/2103.14030) and the decoder layer is another 3D SWIN transformer.
    - **Custom SwinUNETR w/ Connector:** A custom 3D swin generator transformer (https://arxiv.org/pdf/2103.14030) based model with additional residual connections.
  - NVidia's Flare package for federated learning.
  - Model performance observed over training epochs and rounds.

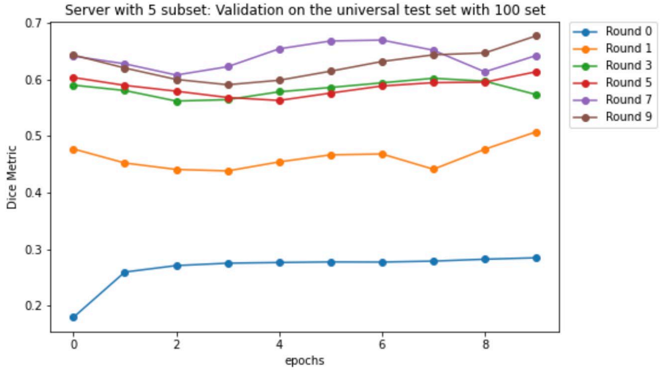
Federated learning for distribution of compute in a heterogeneous compute environment

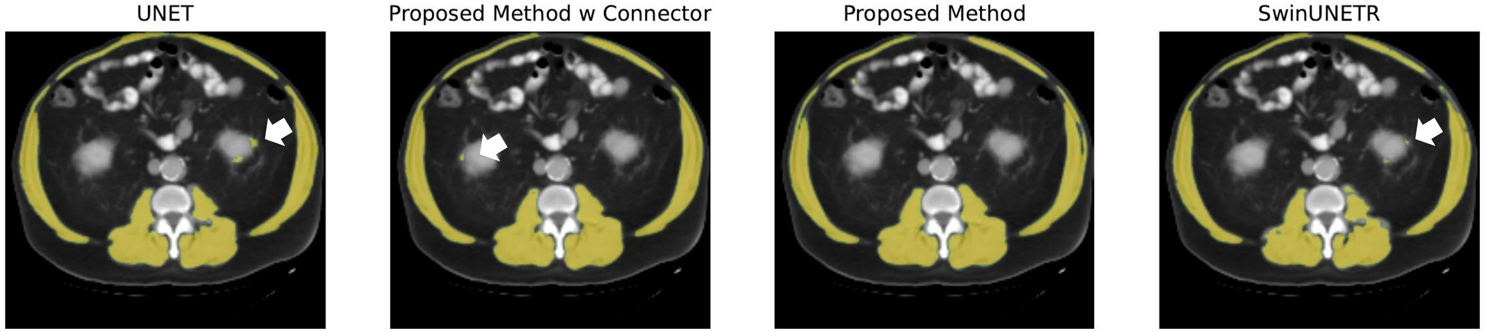
Results:

- Improved site performance, notably for sites with smaller datasets.
- Substantial contributions from sites with unique or rare cases.
- Significant reduction in training time per epoch compared to single GPU training <u>from two days using single GPU training to just a few hours with federated learning.</u>









Conclusions

- Data Privacy and Collaboration:
  - FL provides a solution to the challenge of data sharing across institutions by enabling collaborative model training without the need to centralize sensitive patient data.
  - This is crucial for privacy compliance and data ownership concerns in healthcare.
- Enhanced Model Generalization:
  - By leveraging diverse datasets from multiple institutions, FL enhances the generalizability of models, especially for rare conditions where single-site data is limited, leading to more robust and clinically relevant results.
- Distribution of compute in resource starved clinical environments:
  - Federated learning tools can be used to orchestrate distributed compute in heterogenous environments where large clusters aren't readily available

Thank you!
Department of Radiology
UNIVERSITY OF WISCONSIN
SCHOOL OF MEDICINE AND PUBLIC HEALTH