Data Science: State-of-the Art Technologies, Challenges and Opportunities in Radiology and Oncology Imaging

Ronald M. Summers, M.D., Ph.D.

Senior Investigator

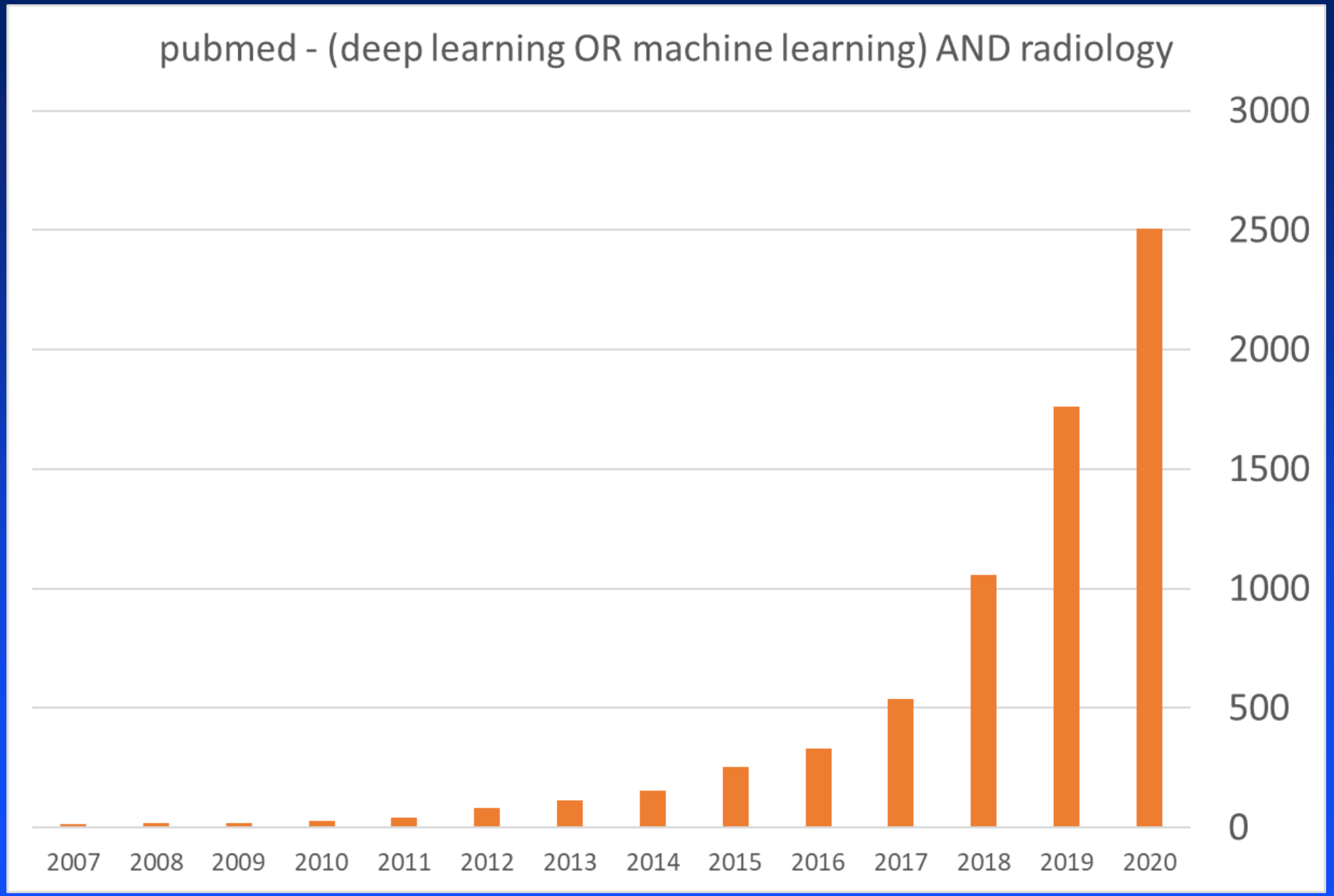
Imaging Biomarkers and CAD Laboratory
Radiology and Imaging Sciences
NIH Clinical Center, Bethesda, MD
rms@nih.gov
github.com/rsummers11
www.cc.nih.gov/drd/summers.html

Disclosures

- Patent royalties from iCAD, ScanMed, PingAn, Philips, Translation Holdings
- Research support from Ping An & NVIDIA

AI

- Detection
- Diagnosis
- Segmentation
- Reconstruction
- Risk prediction and prognosis


pubmed - (deep learning OR machine learning) AND radiology
3000
2500
2000
1500
1000
500
0
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020

PMC

A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version)

In December 2019, a new type viral pneumonia cases occurred in Wuhan, Hubei Province; and then named "2019 novel coronavirus (2019-nCoV)" by the World Health Organization (WHO) on 12 January 2020.

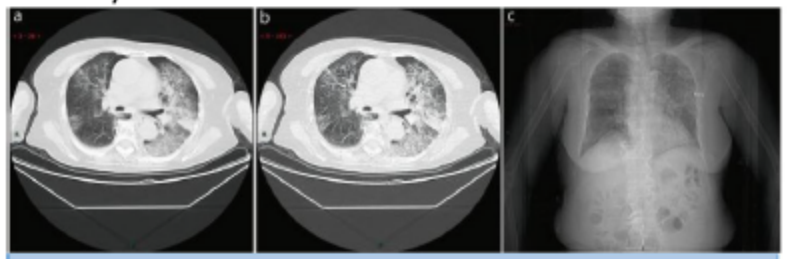


Fig. 4

Atypical CT / X-ray imaging manifestation (case 1). An 83 years old female with fever for 4 days…

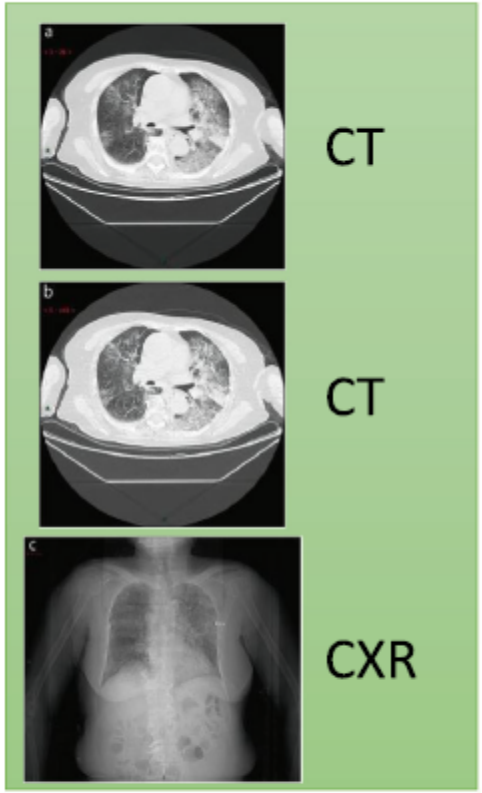
Single, or multiple, or extensive subpleural grid-like or honeycomb-like thickening of interlobular septum, thickening of the bronchial wall, and tortuous and thick strand-like opacity. Several patchy consolidations, occasionally with a small amount pleural effusion or enlargement of mediastinal lymph nodes, can be seen (Fig. 4: 6 cases, 7.2% in a total of 83 cases). This is mostly seen in the elderly.

Text extraction

Figure segmentation (Tsutsui et al., 2017)

Modality classification

Database

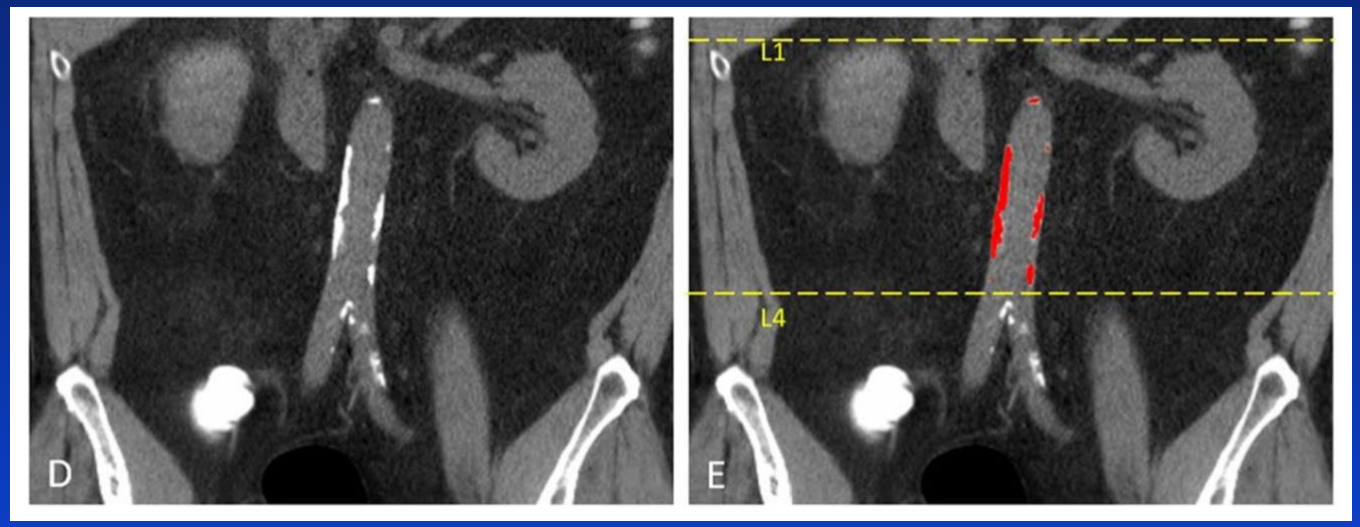


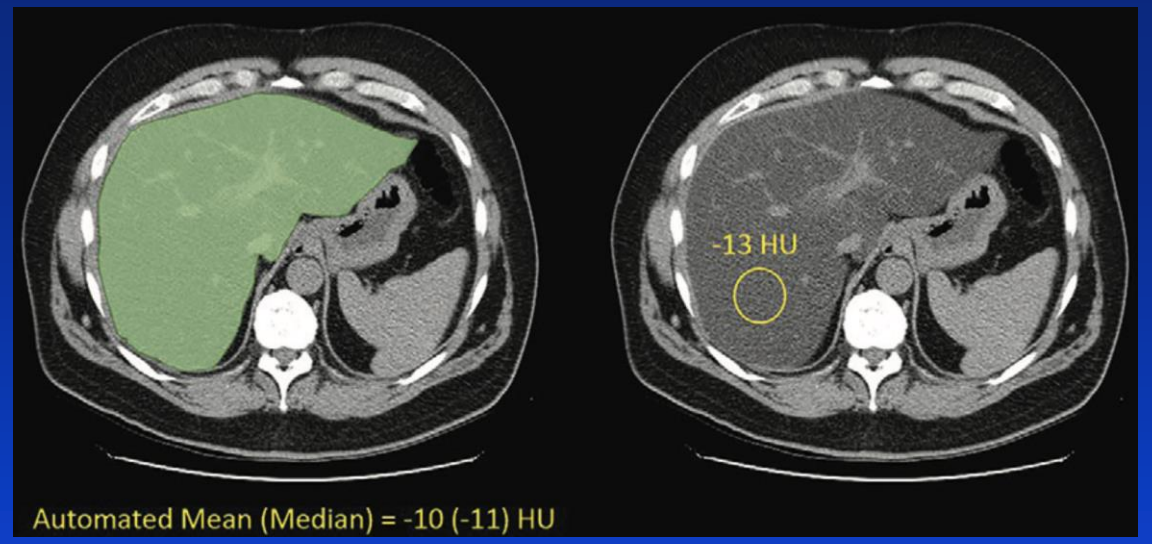
Caption

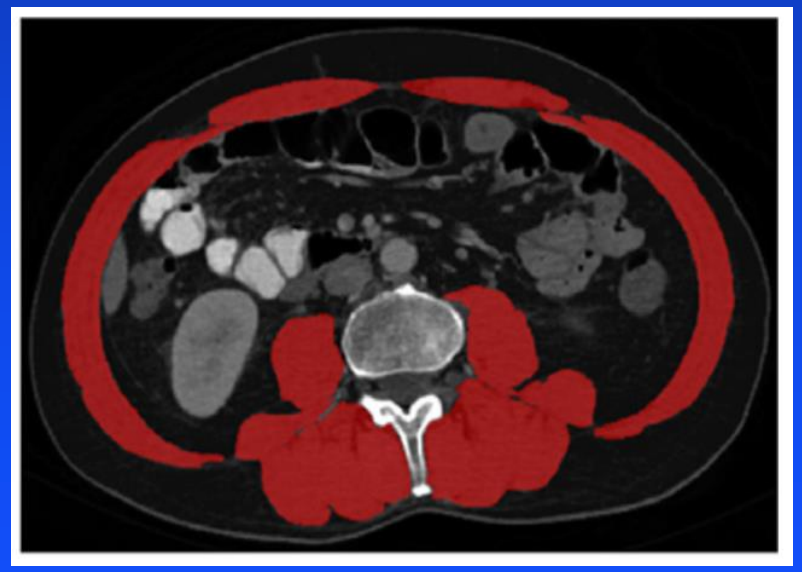
Referred text

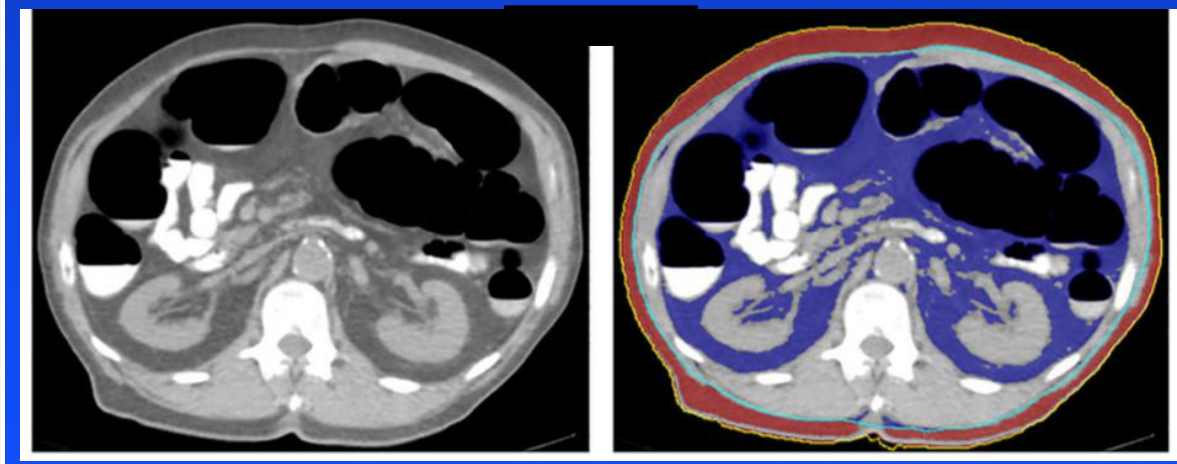
Y Peng et al. IEEE Transactions on Big Data 2021; YH Jin, Mil Med Res 2020

Large-scale Body Composition Analysis



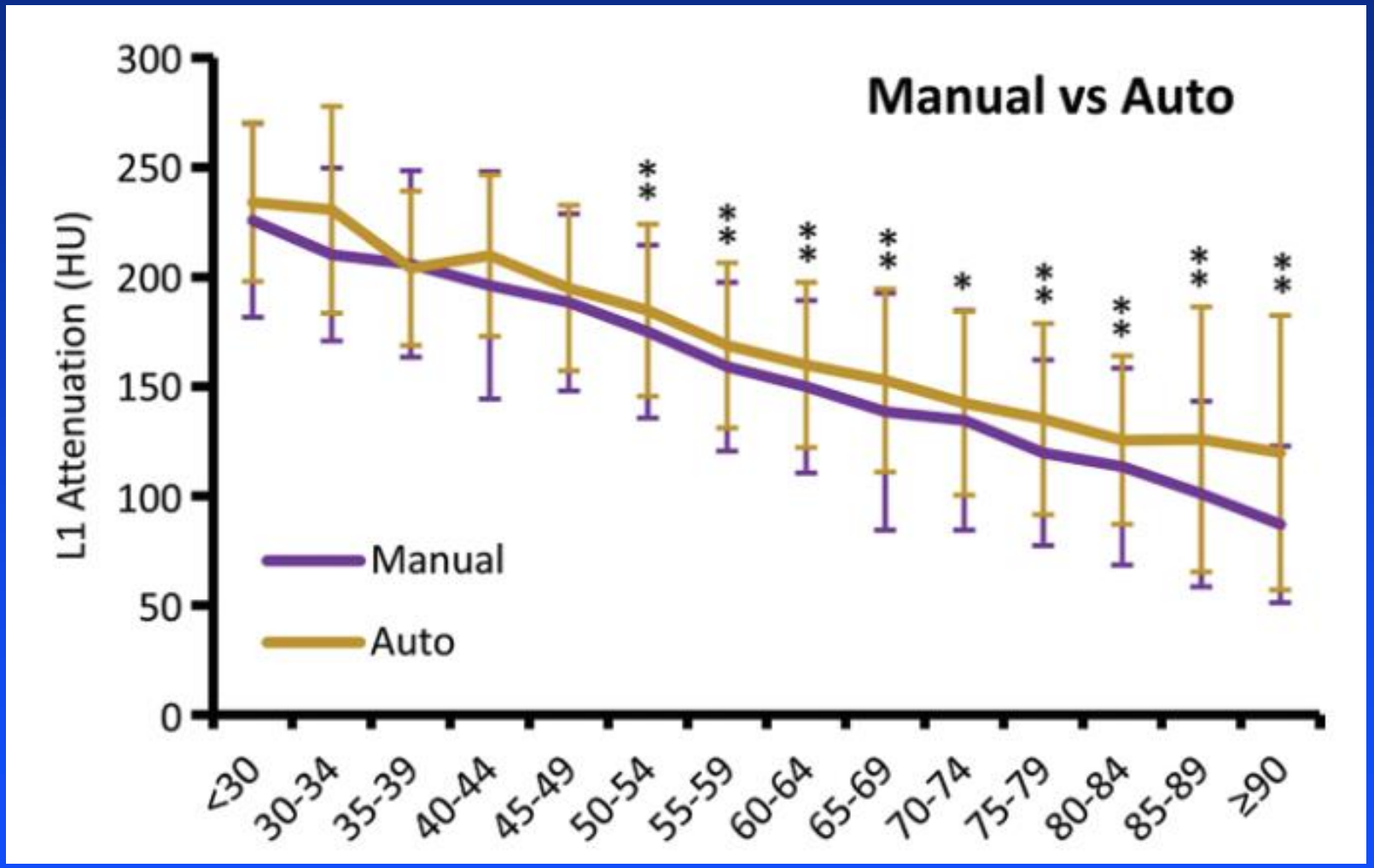


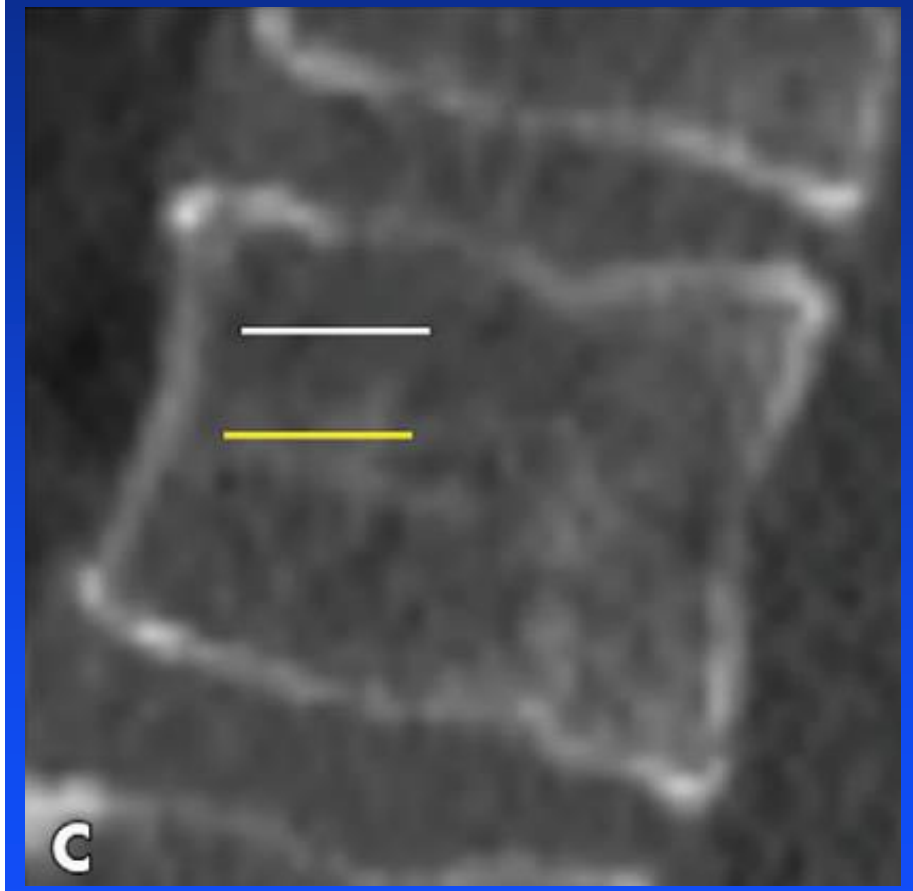




Lee et al. BJR 2018; Burns et al. Acad Radiol 2019; Graffy et al. Abd Radiol 2019; Jang et al. Radiology 2019; Graffy et al. Radiology 2019

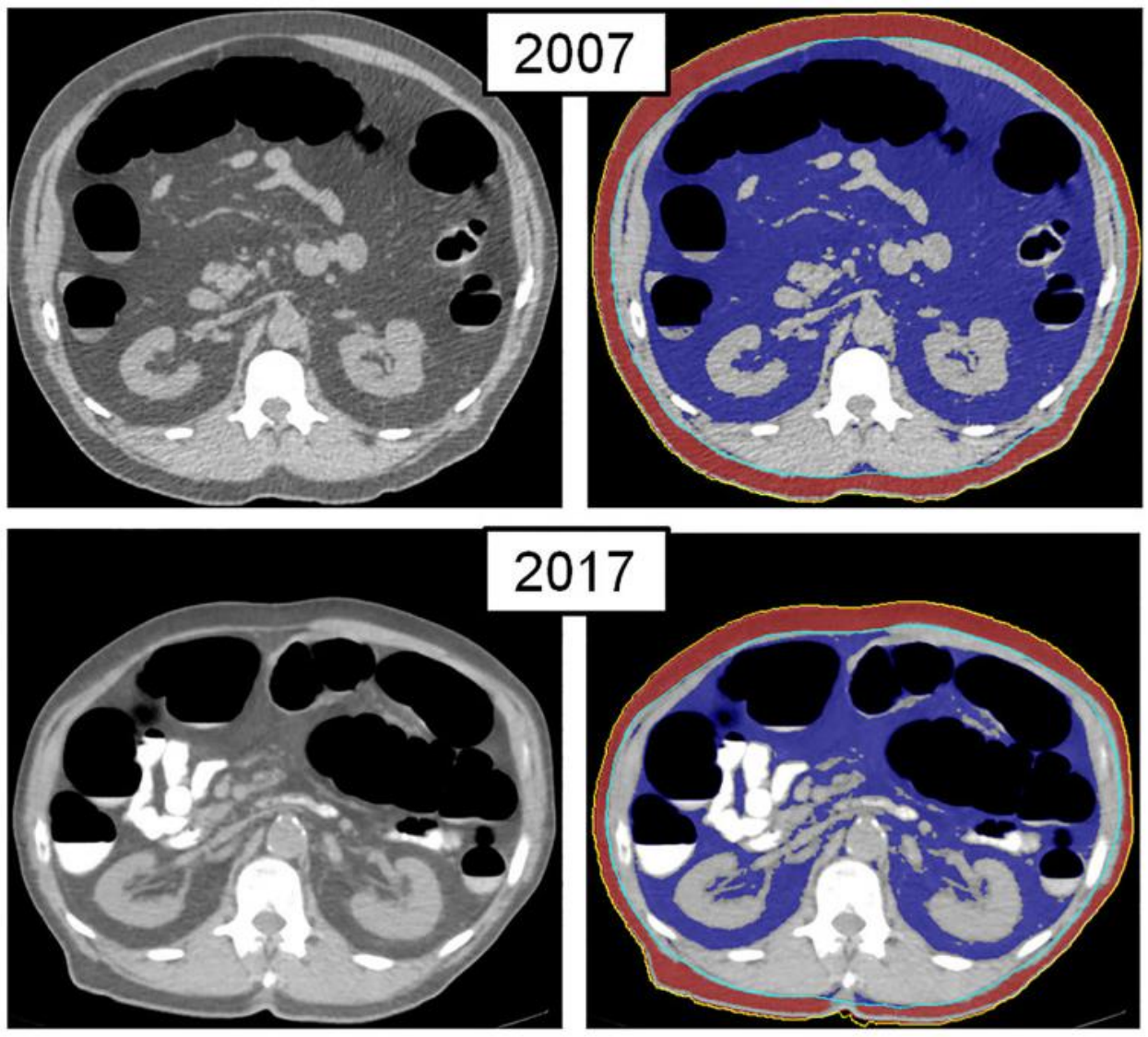
1. Bone Mineral Densitometry

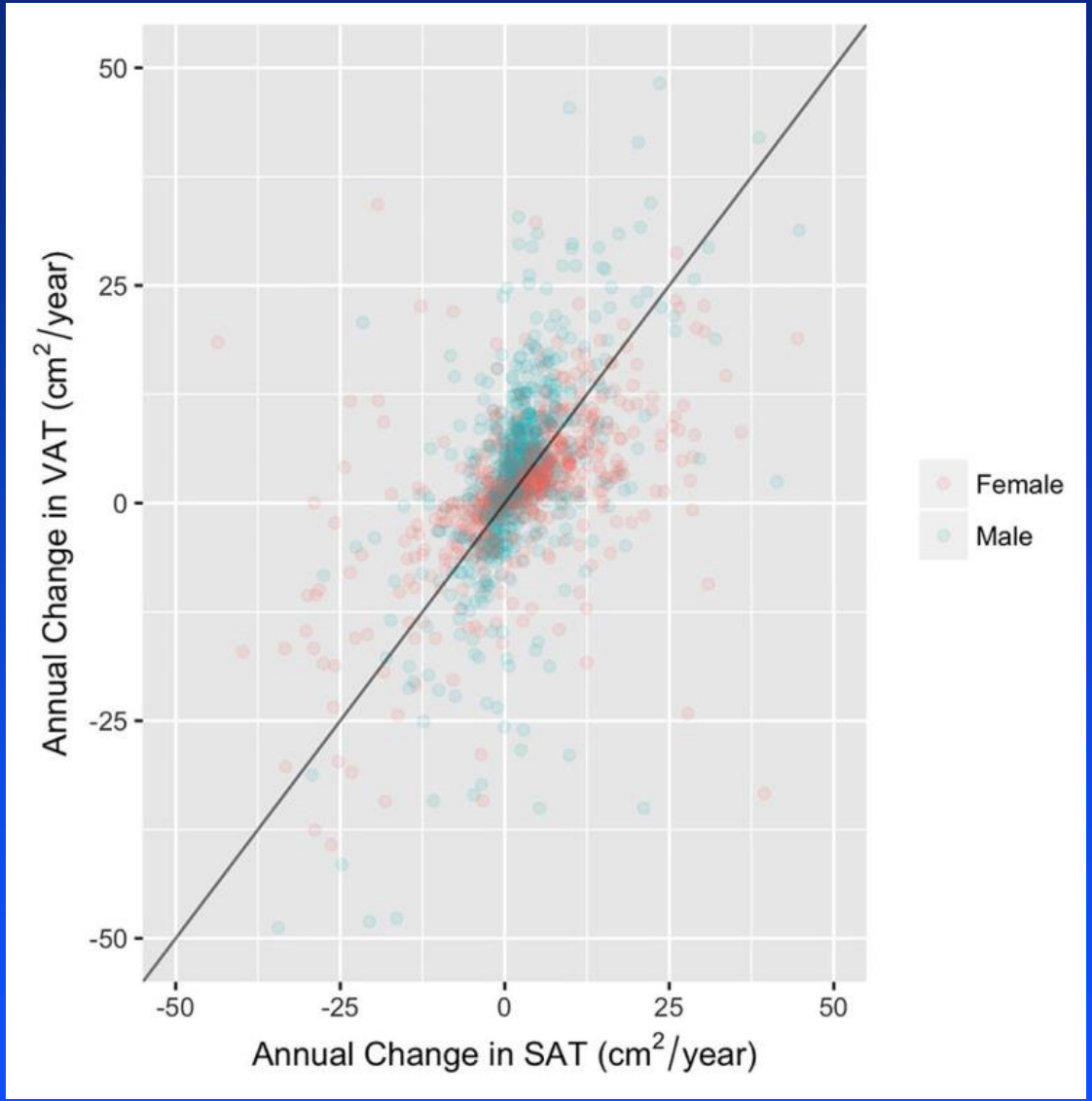




Jang et al. Radiology 2019

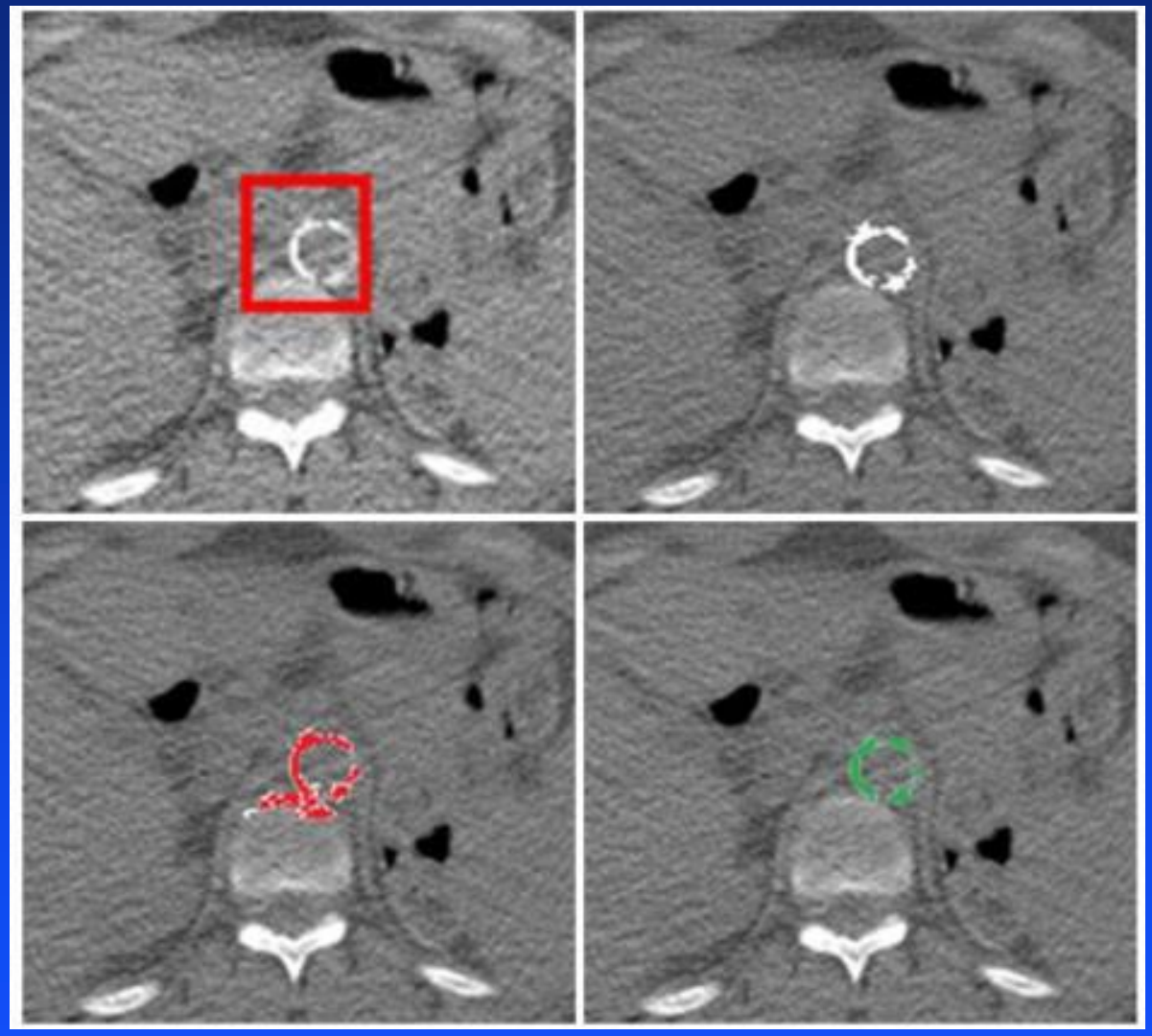
2. Visceral and SQ Fat





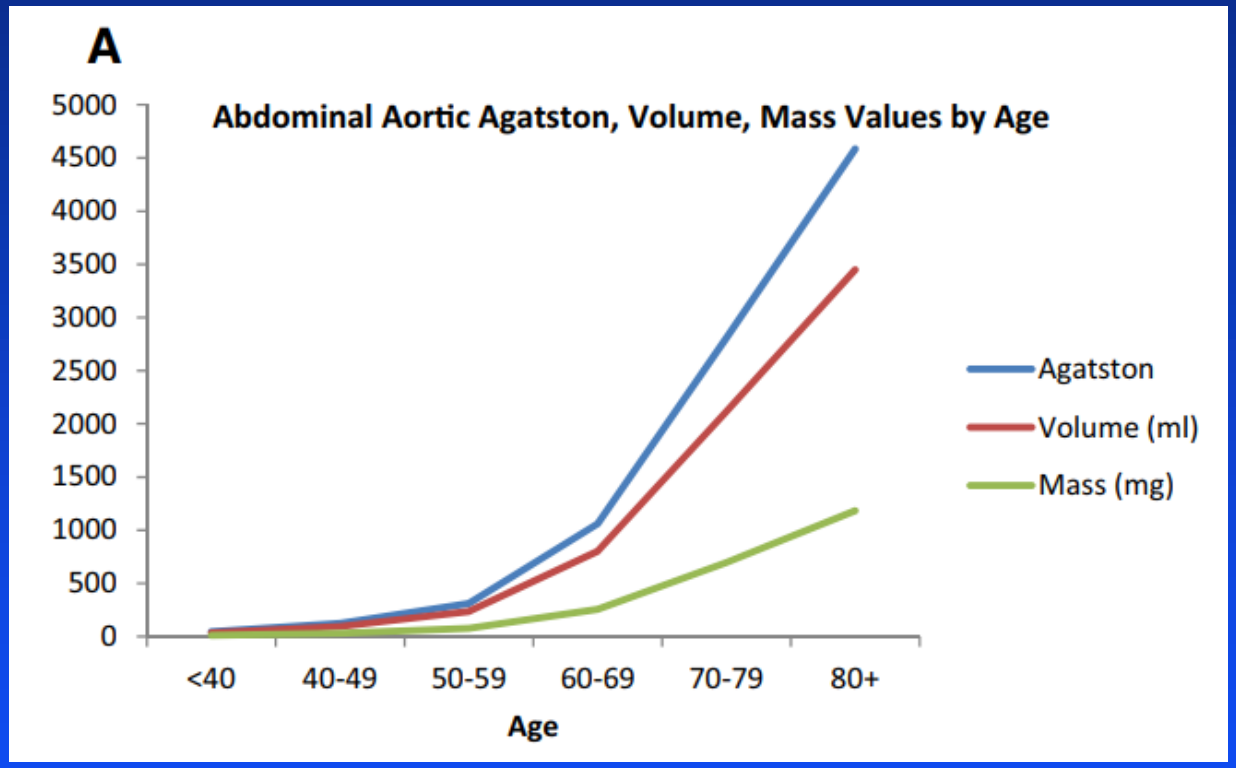
Lee et al. BJR 2018

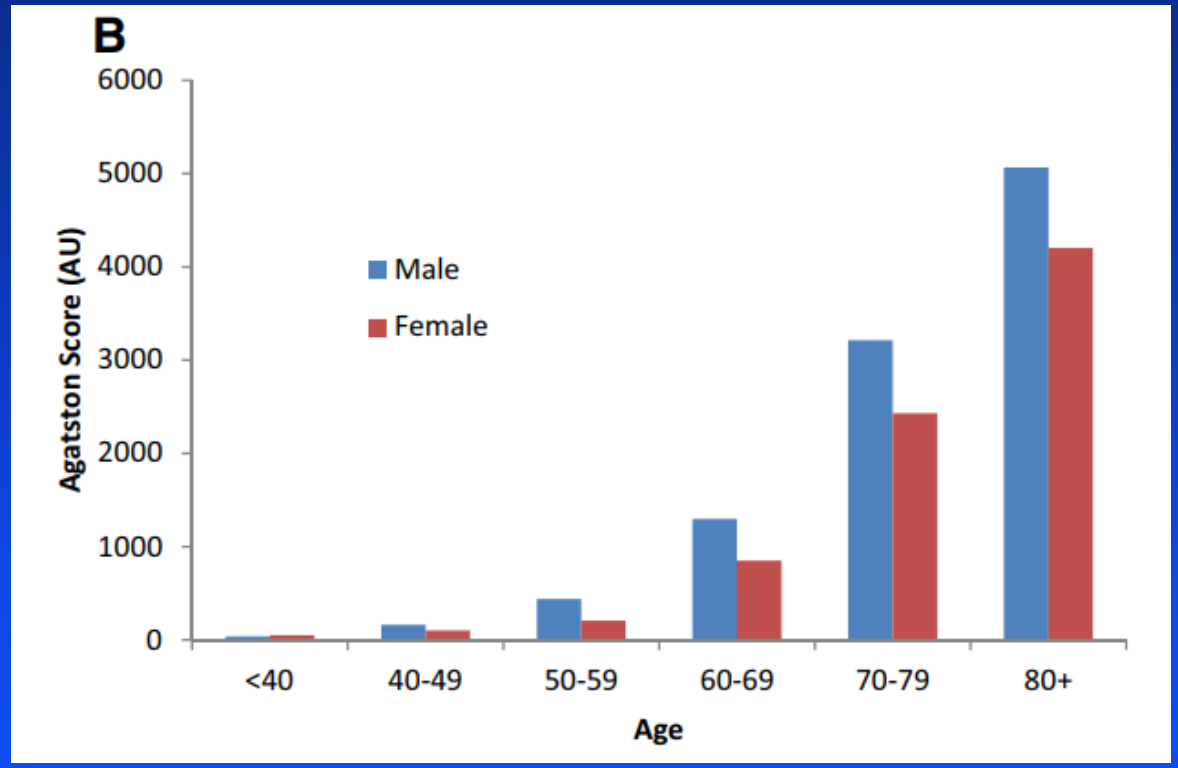
3. Abdominal Aortic Plaque



Chellamuthu et al. IEEE ISBI 2017

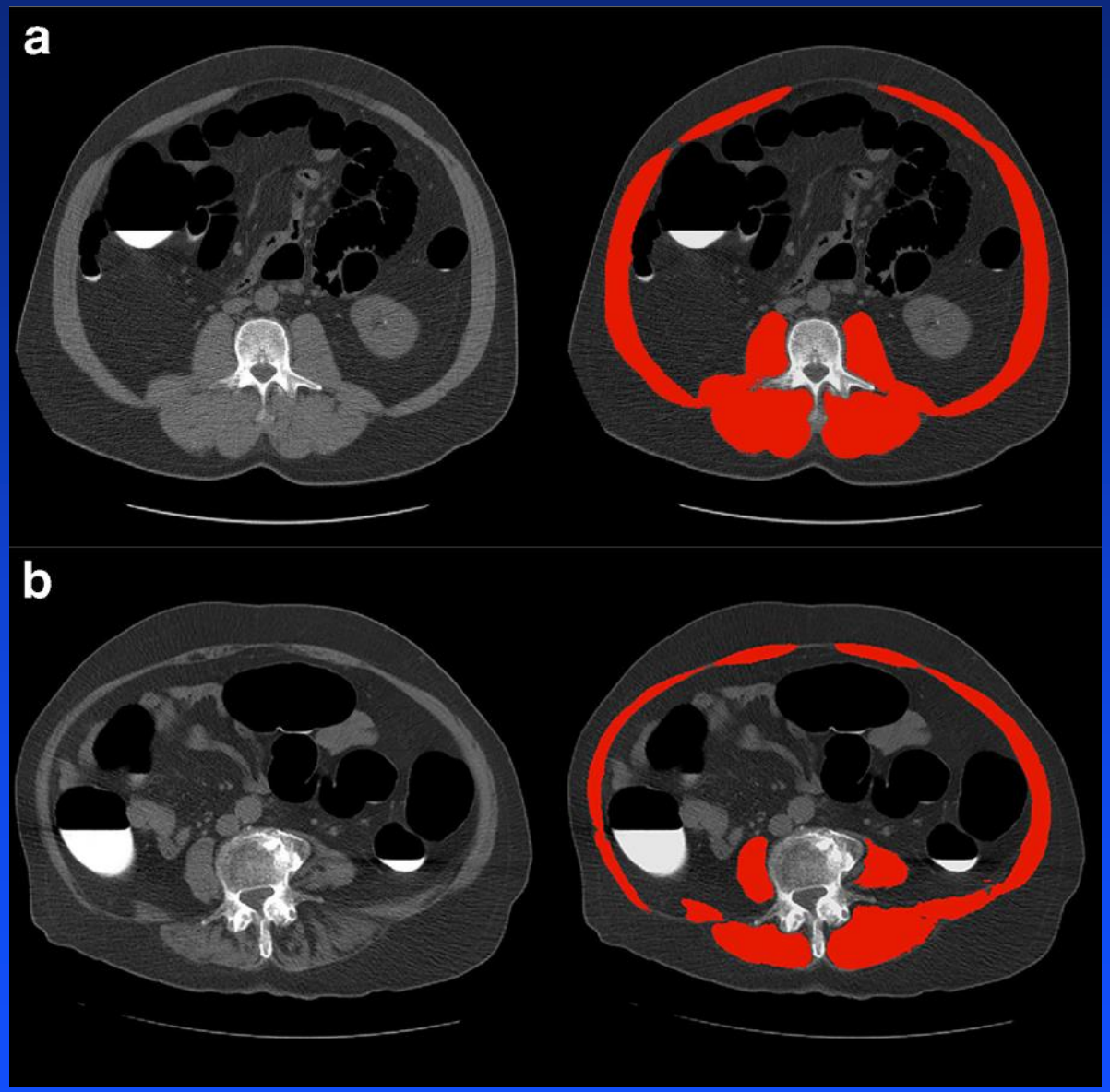
3. Abdominal Aortic Plaque





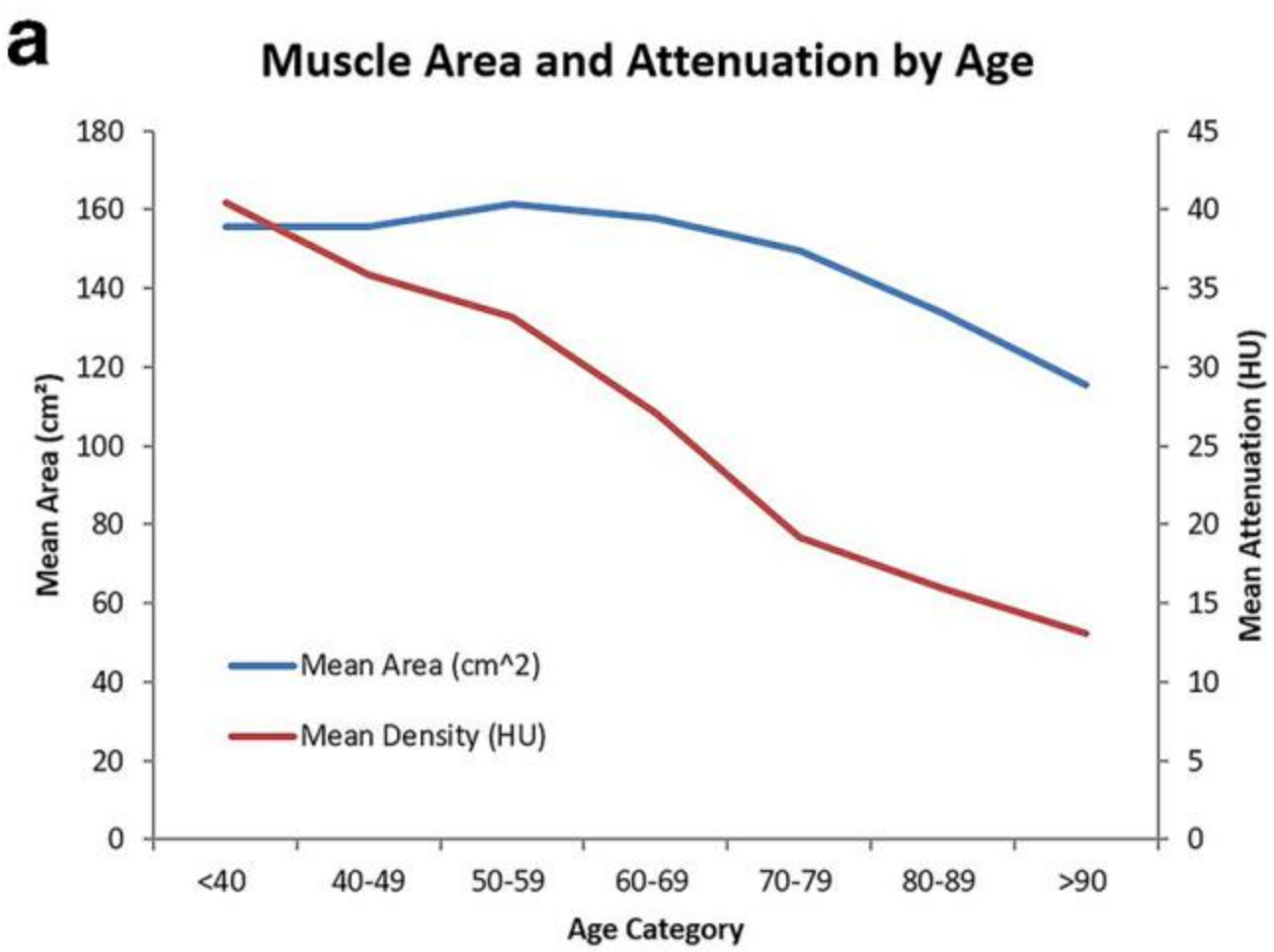
Graffy et al. Abdominal Radiology 2019

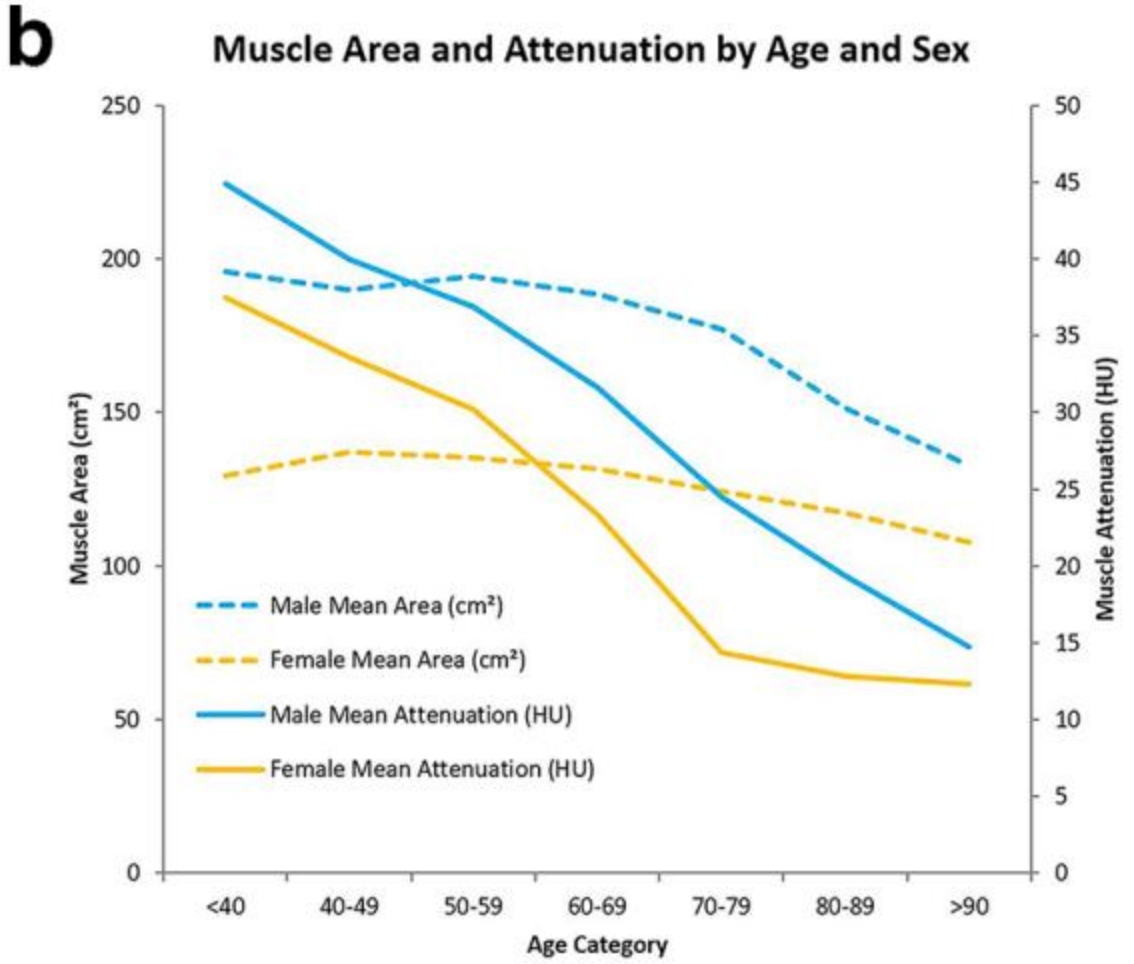
4. Abdominal Skeletal Muscle



Graffy et al. BJR 2019

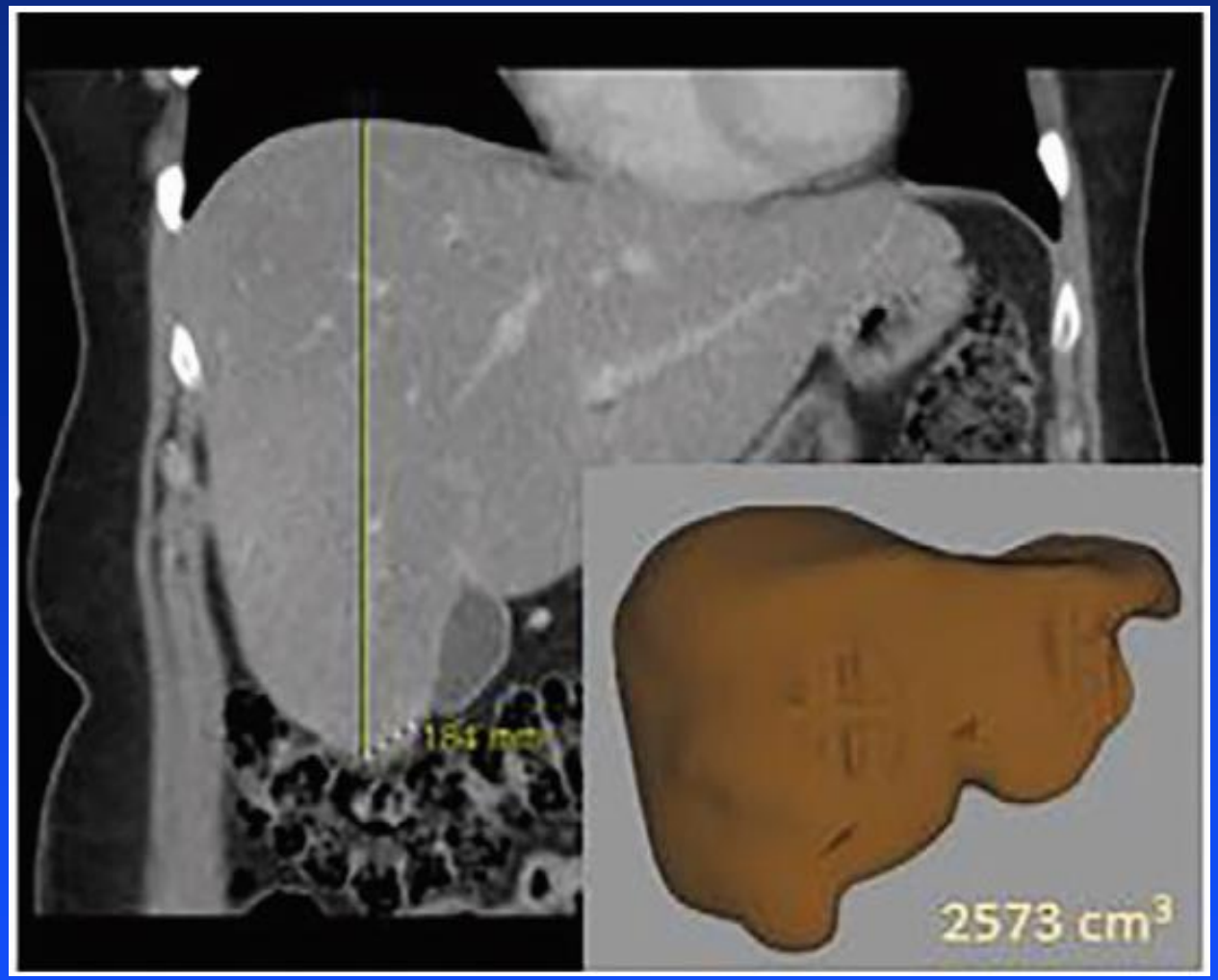
4. Abdominal Skeletal Muscle





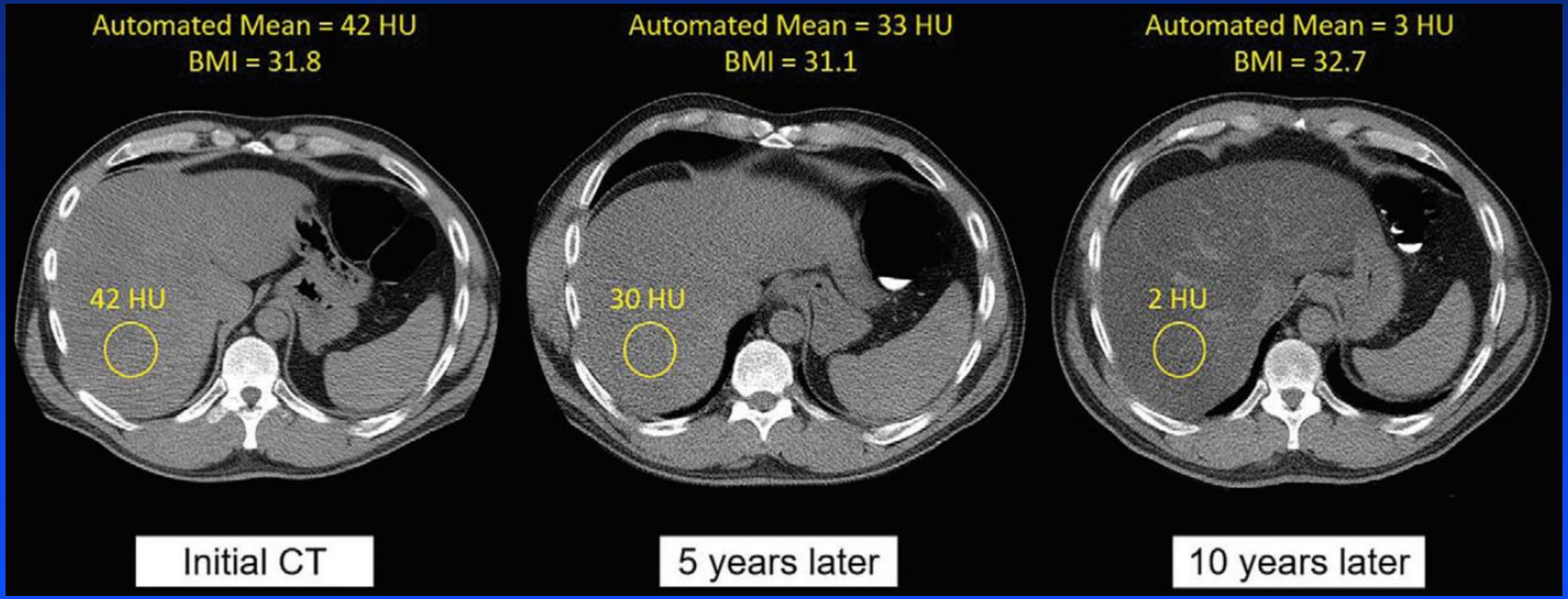
Graffy et al. BJR 2019

5. Liver CT Volume & Attenuation



Pickhardt et al. Radiographics 2021

5. Liver CT Volume & Attenuation



Graffy et al. Radiology 2019

5. Liver CT Volume & Attenuation

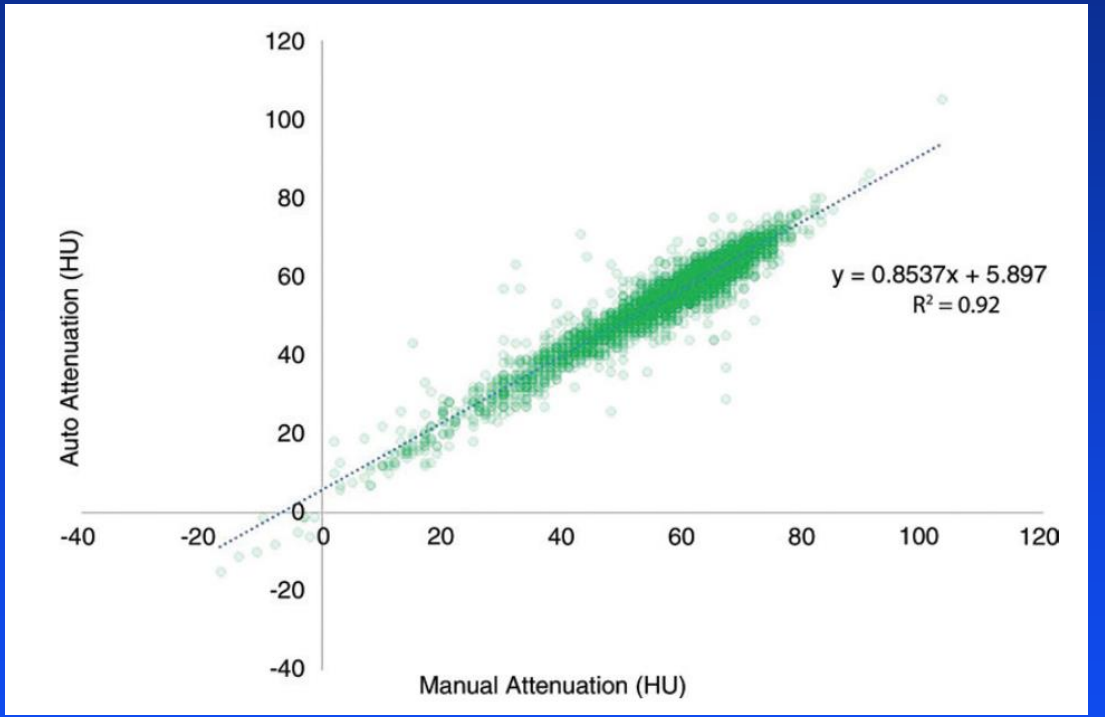


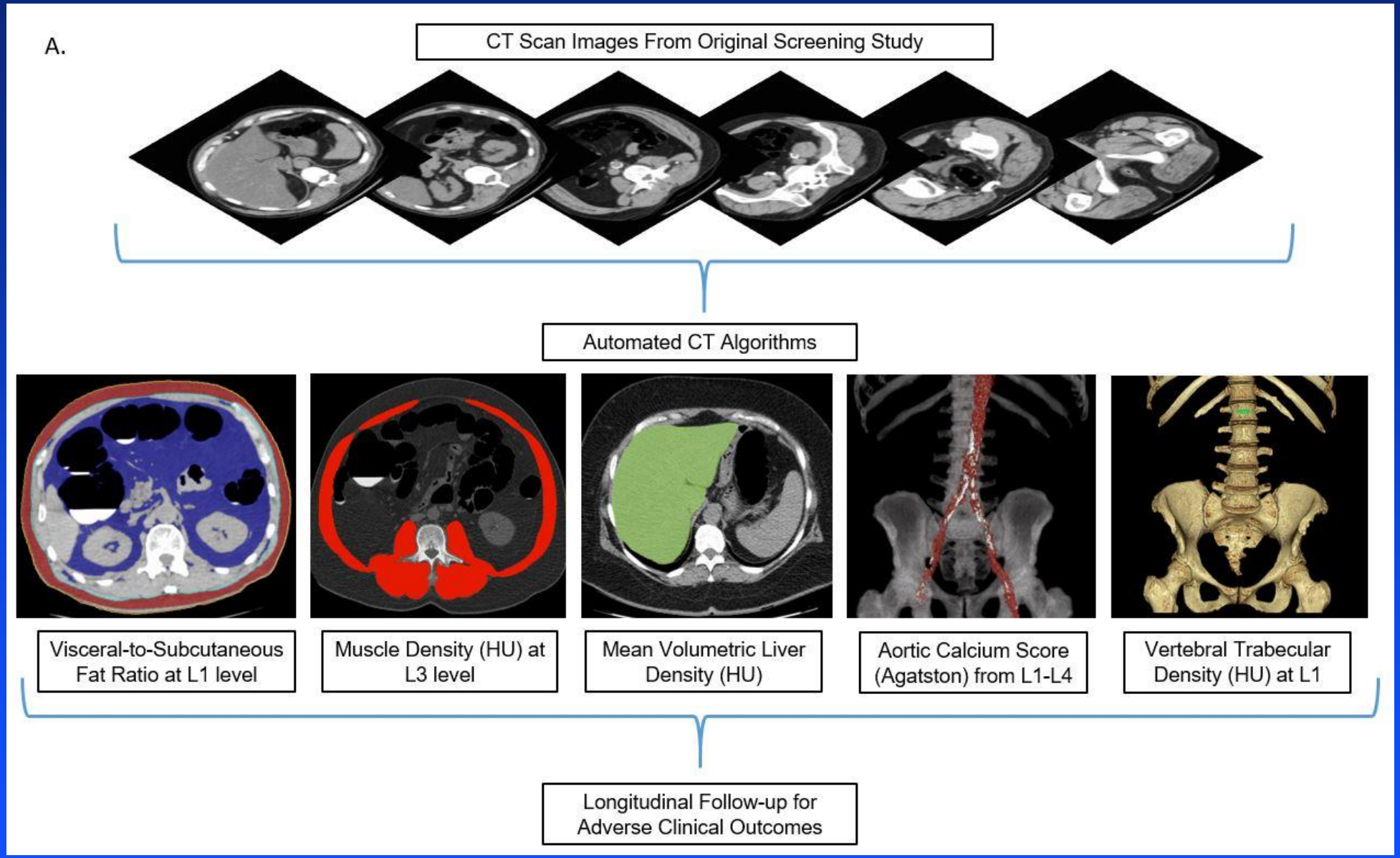
Table 3: 4 × 4 Confusion Matrix for Steatosis Categorization by Using Automated versus Manual Measures (*n* = 5265)

	Automated Steatosis Categorization			
Manual Steatosis Categorization	Normal	Mild	Moderate	Severe
Normal	3137	661	2	0
Mild	51	892	63	0
Moderate	3	41	343	5
Severe	0	1	16	50

Note.—Data are CT scans of patients.

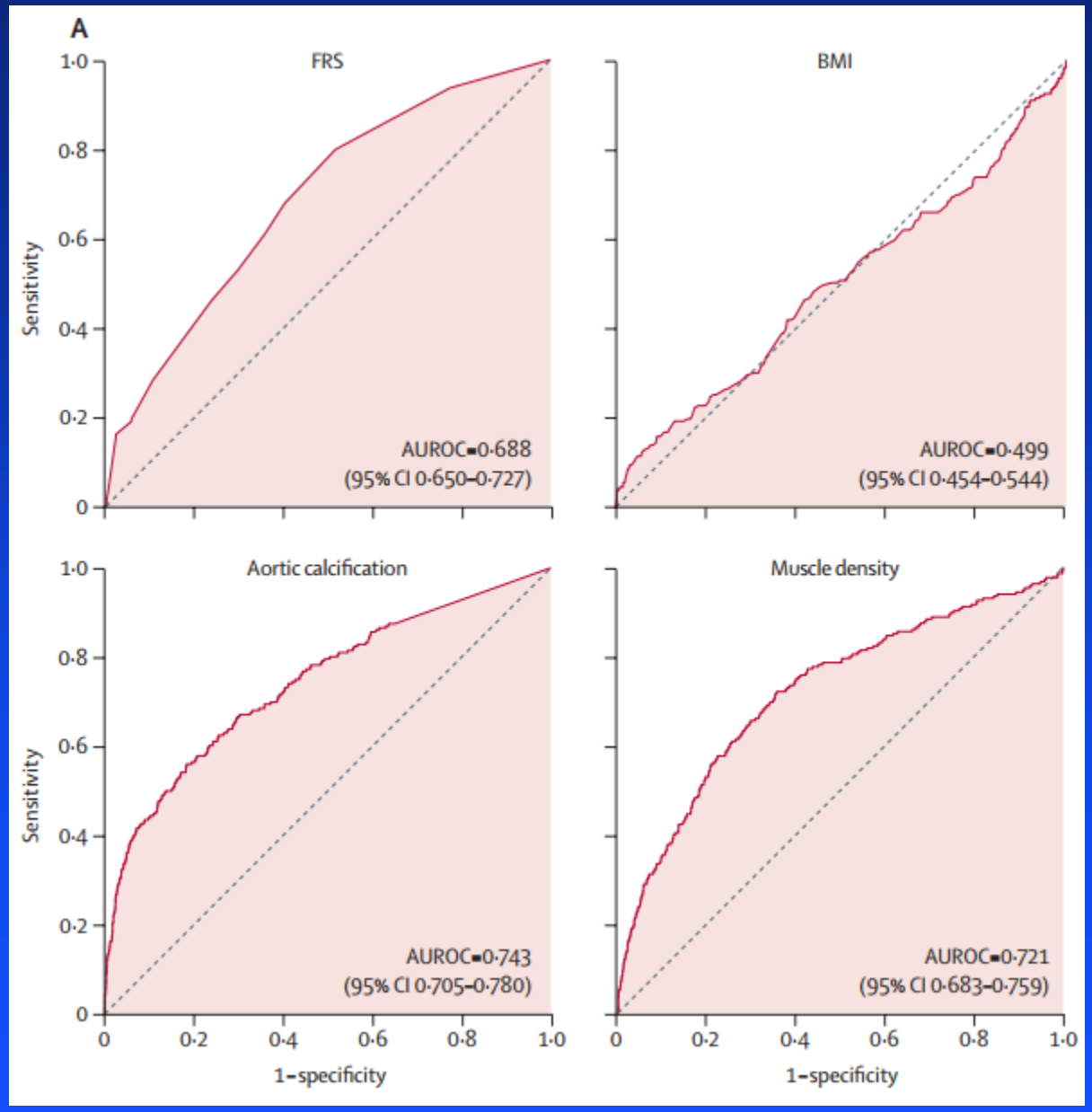
Graffy et al. Radiology 2019

Cardiovascular Risk and Overall Survival

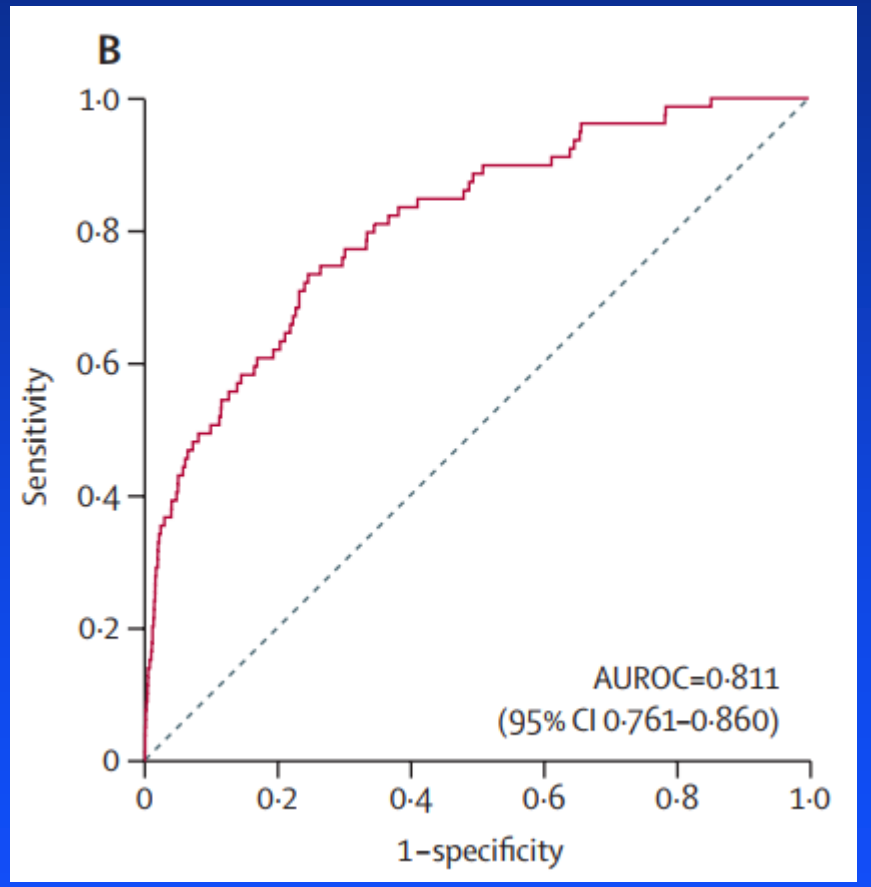


Pickhardt et al. Lancet Digital Health 2020

Cardiovascular Risk and Overall Survival



Multivariate Ao Ca++, Muscle Density, Liver Density

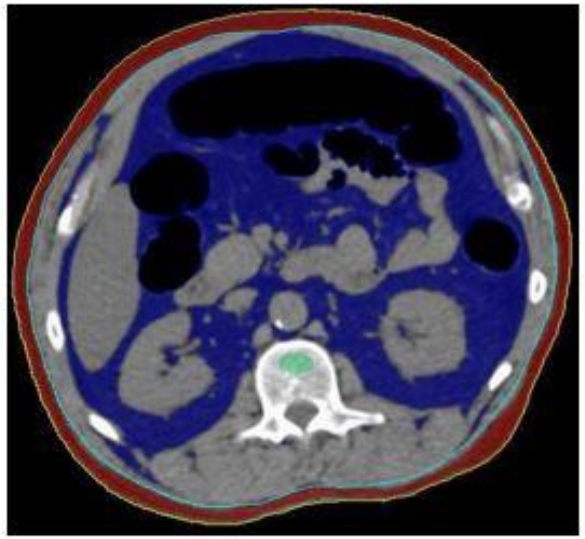


Pickhardt et al. Lancet Digital Health 2020

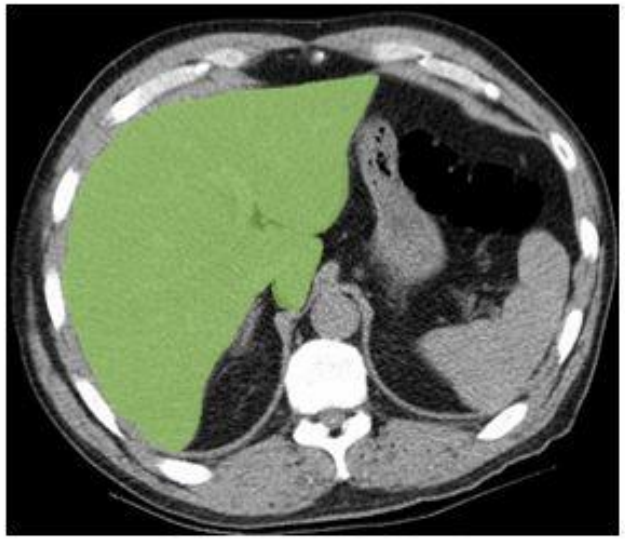
Cardiovascular Risk and Overall Survival

B.

Asymptomatic 57-year-old Man
Undergoing Colorectal Cancer Screening



V/S Ratio = 3.1
(99th percentile)



Liver Density = 28 HU
(97th percentile)

Agatston Score = 5070
(97th percentile)

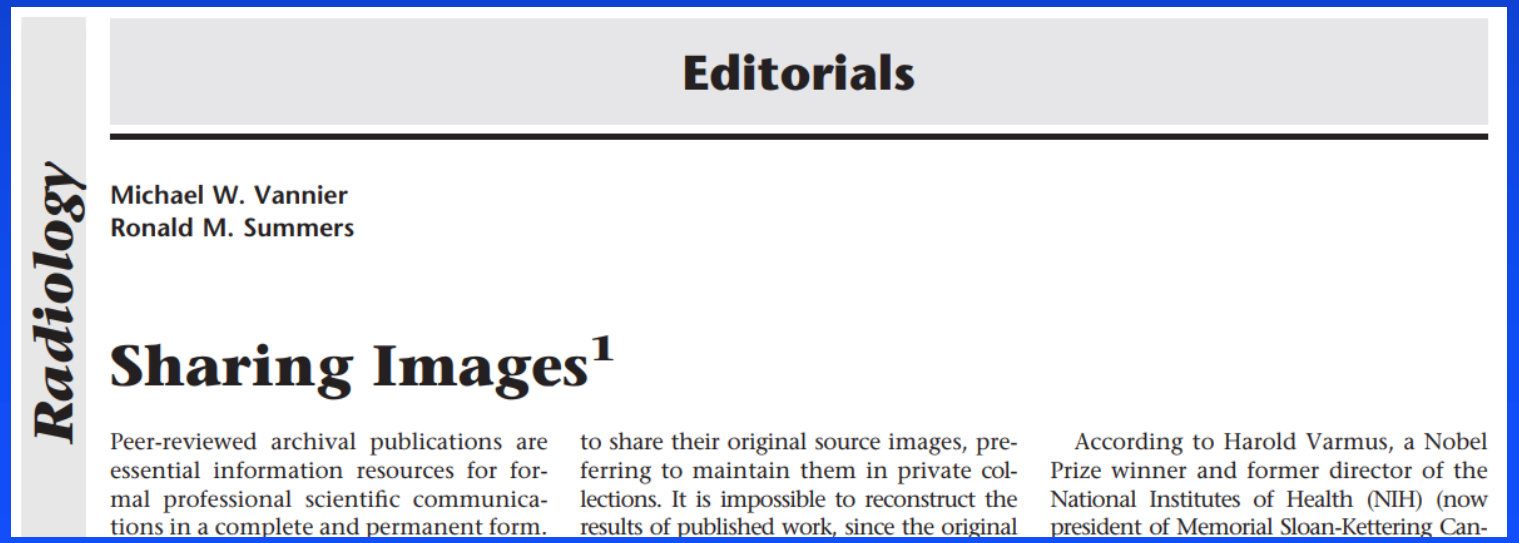
C.

CT scan performed 12
years after original study

Pickhardt et al. Lancet Digital Health 2020

Public Datasets for AI & ML

- Permit reproducible research
- Permit others to improve algorithms
- Encourage interest in medical imaging AI
- Provide recognition to dataset authors



Radiology

Editorials

Michael W. Vannier
Ronald M. Summers

Sharing Images[1]

Peer-reviewed archival publications are essential information resources for formal professional scientific communications in a complete and permanent form. to share their original source images, preferring to maintain them in private collections. It is impossible to reconstruct the results of published work, since the original

According to Harold Varmus, a Nobel Prize winner and former director of the National Institutes of Health (NIH) (now president of Memorial Sloan-Kettering Can-

Radiology 2003

NEWS RELEASES

Media Advisory

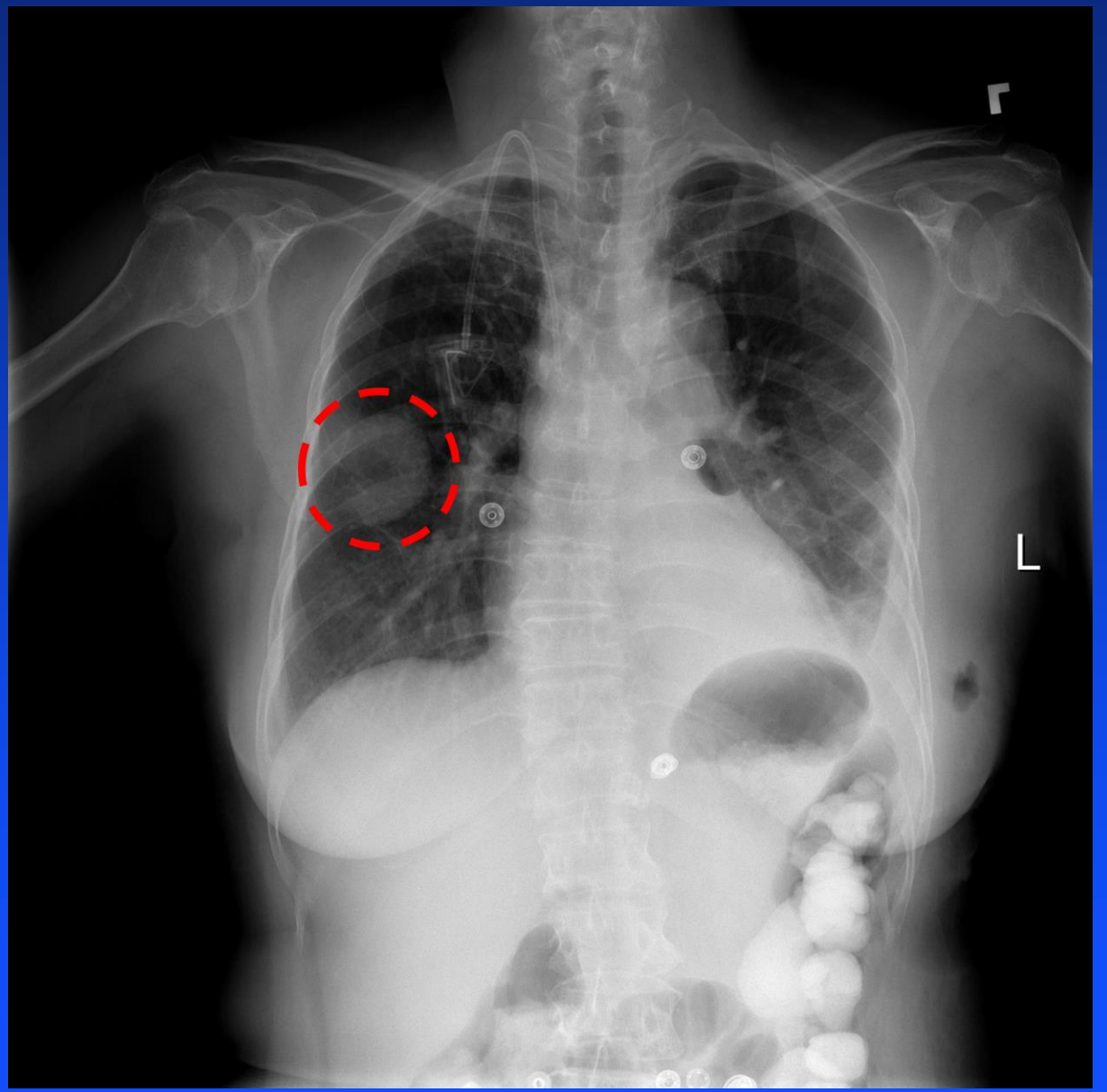
Wednesday, September 27, 2017

NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community

The dataset of scans is from more than 30,000 patients, including many with advanced lung disease.

ChestX-ray8 Dataset

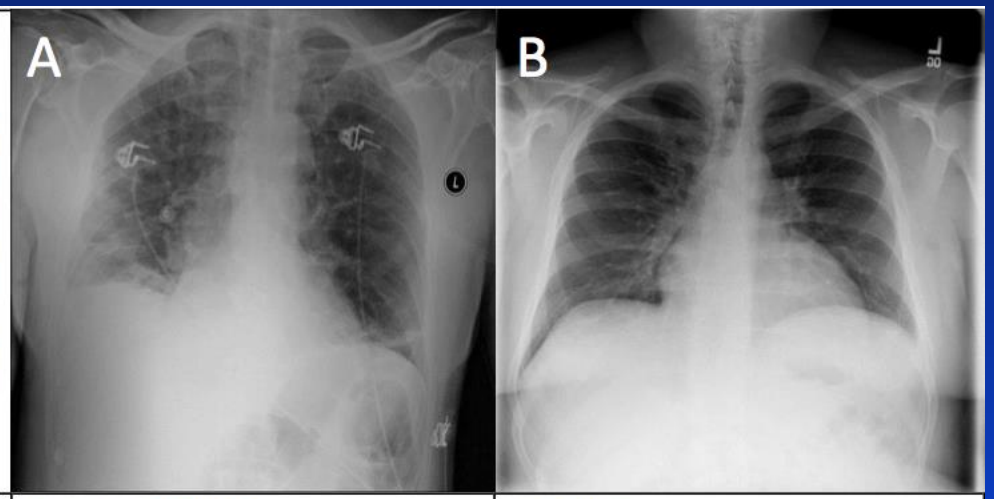
- https://nihcc.app.box.com/v/ChestXray-NIHCC
- “ChestX-ray8 Dataset”
- 112,120 frontal-view chest radiographs, 30,805 unique patients
- 42 GB
- Metadata for all images
- Bounding boxes for 1000 images

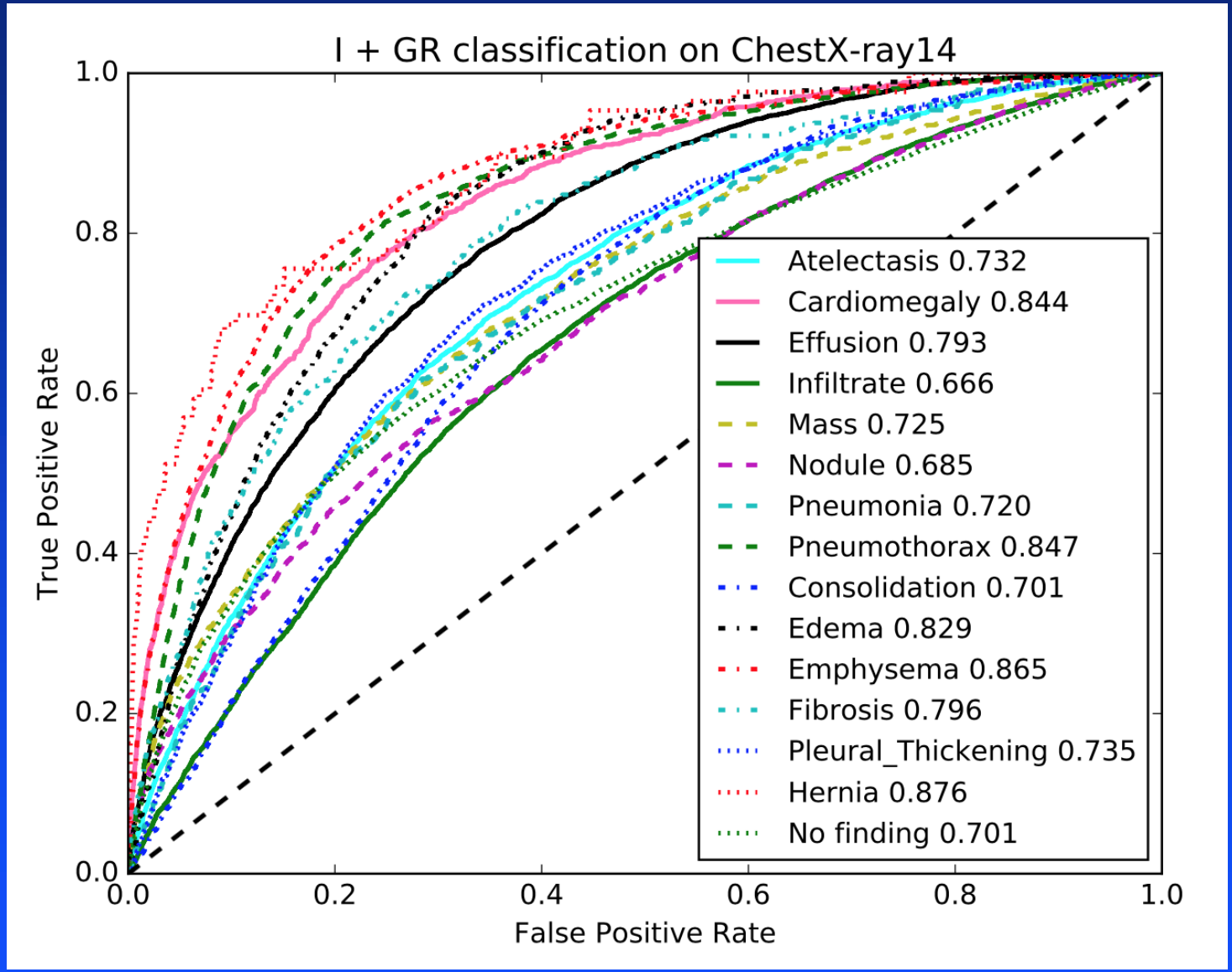


Challenges to Creating Large Datasets

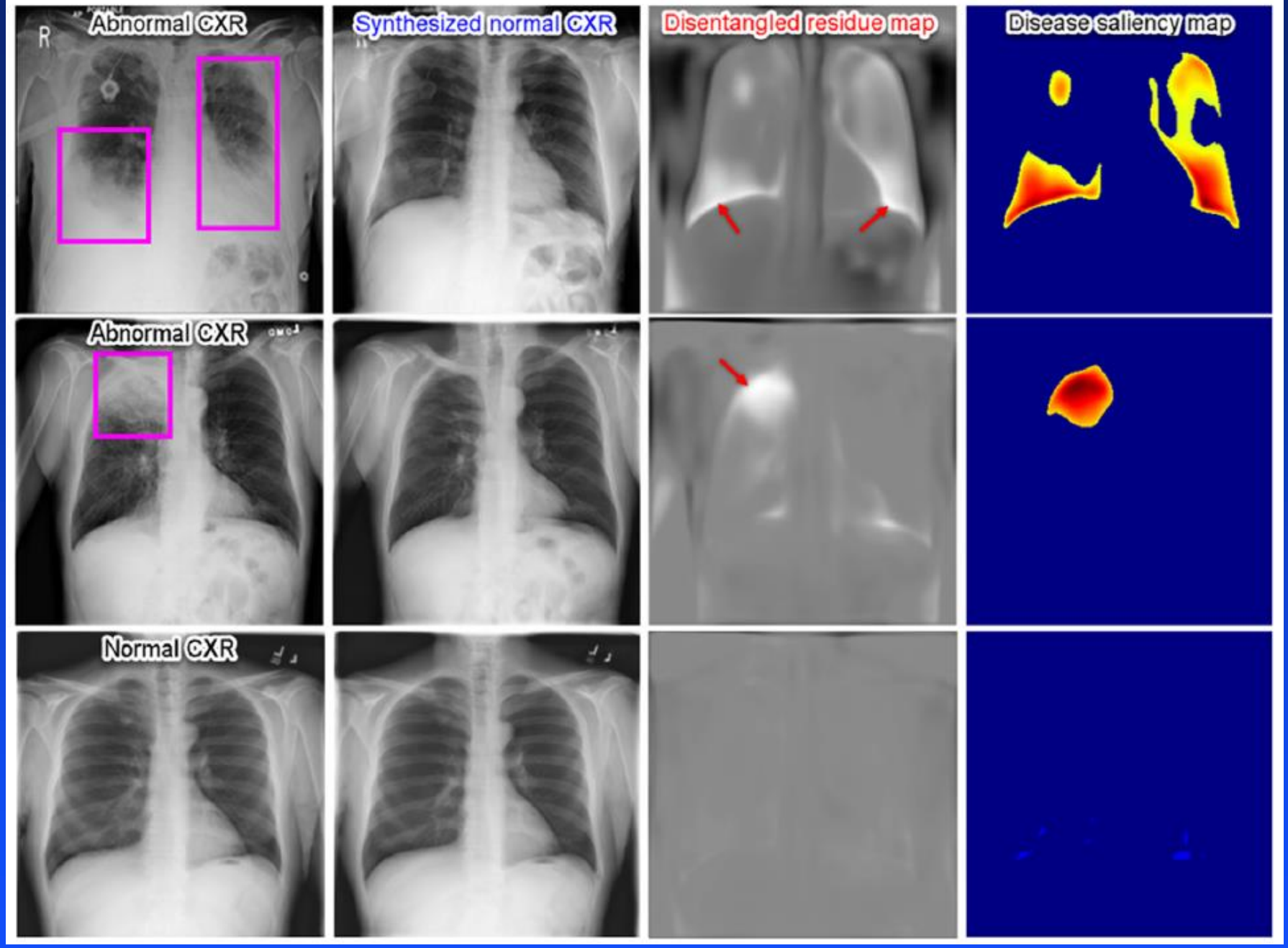
- IRB approval
- PACS support
- Good annotation requires gold standard, efficient tools and domain expertise
- Data wrangling is time consuming and costly

Automated Report Generation

Image Sample cases		
P	Atelectasis Effusion	No finding
Original report	findings : a single ap view of the chest demonstrates increasing bibasilar interstitial opacities with decreased overall aeration . increasing blunting of right costophrenic angle. … impression : increasing bibasilar atelectasis with possible development of right pleural effusion .	Normal no evidence of lung infiltrate .
Generated Report	findings : a single ap view of the chest demonstrates unchanged bilateral reticular opacities , consider atelectasis . continued left basilar atelectasis . no evidence of developing infiltrate . the cardiac and mediastinal contours are stable . impression : no evidence of developing infiltrate .	findings : pa and lateral views of the chest demonstrate lungs that are clear without focal mass , infiltrate or effusion . cardiomediastinal silhouette is normal size and contour . pulmonary vascularity is normal in caliber and distribution . impression : no evidence of acute pulmonary pathology

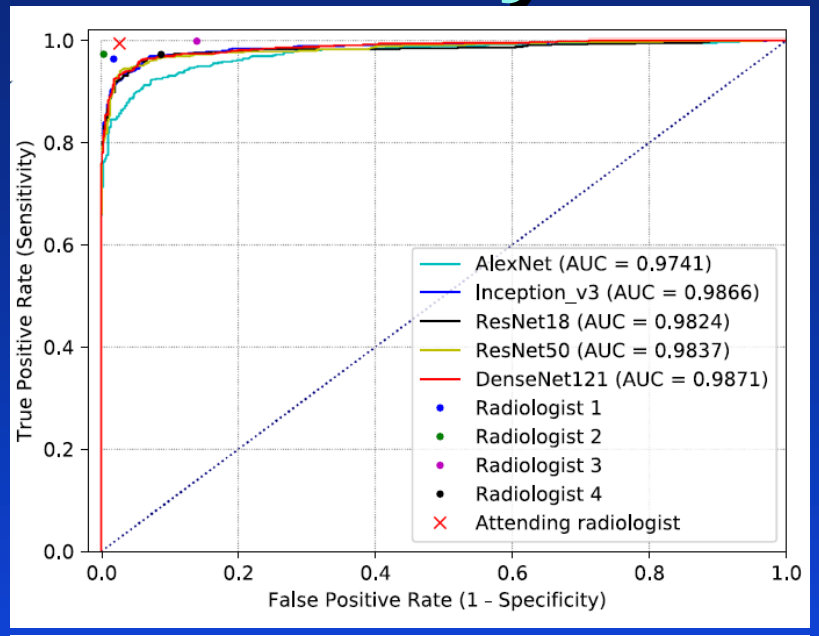


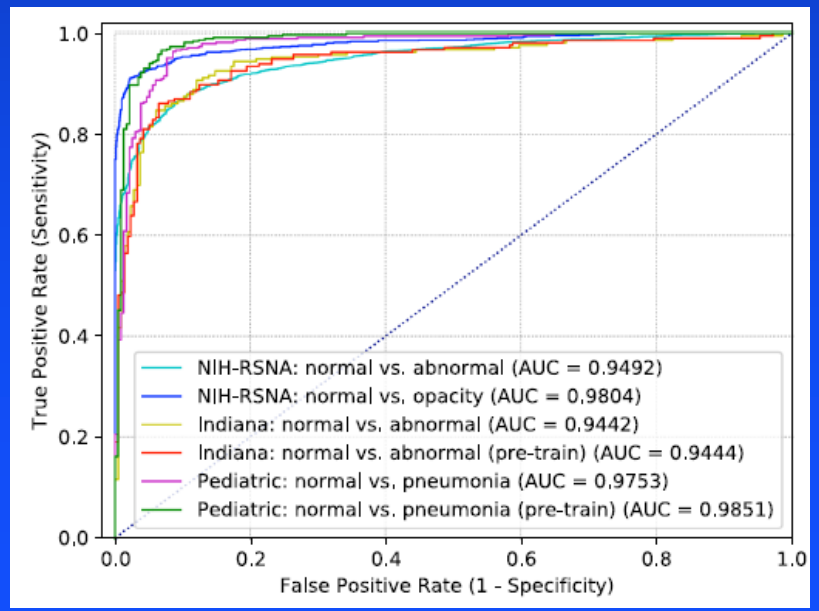
X Wang et al. CVPR 2018

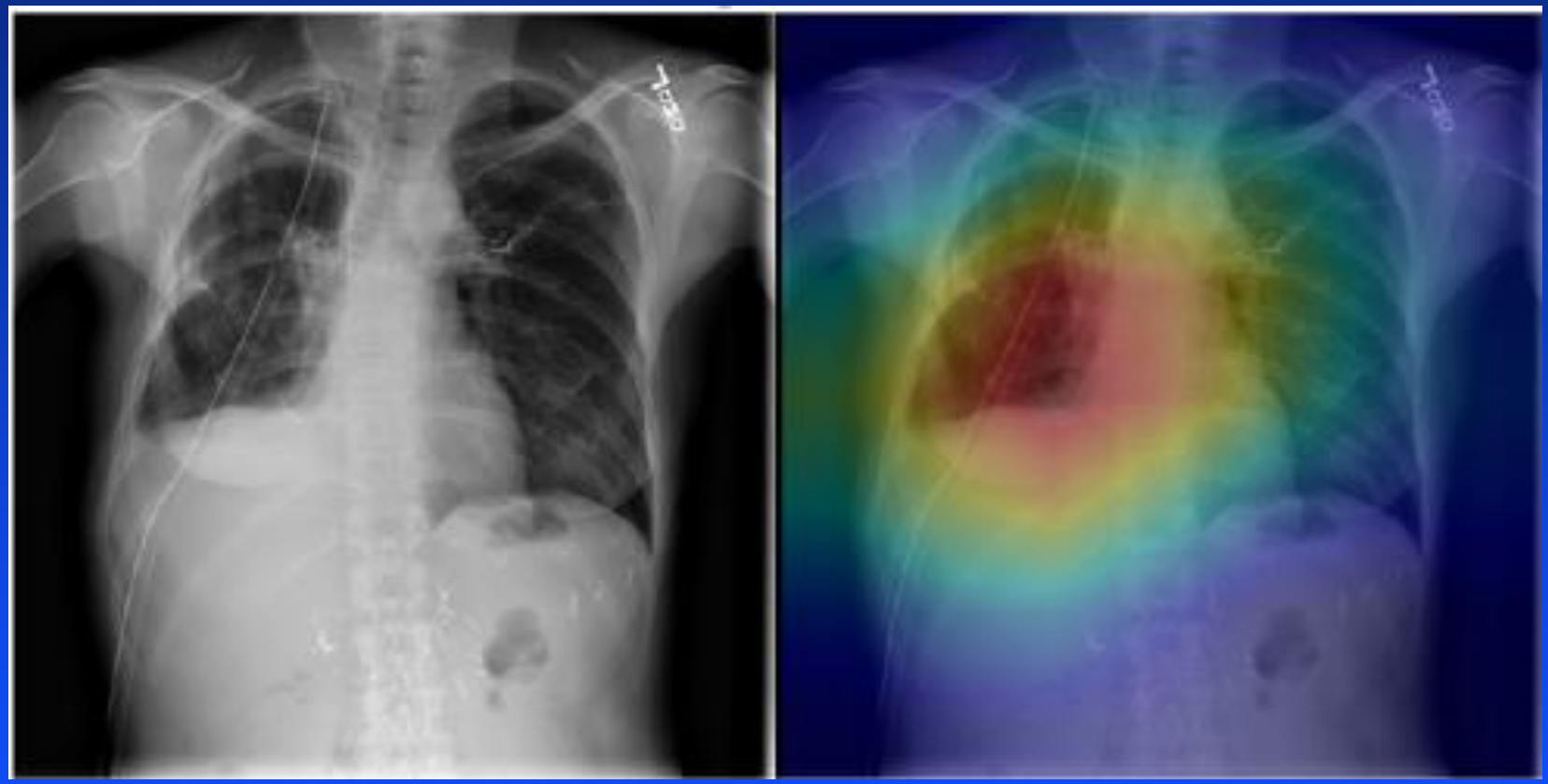


Youbao Tang et al. Medical Image Analysis 2021

Binary Chest Radiograph Classification

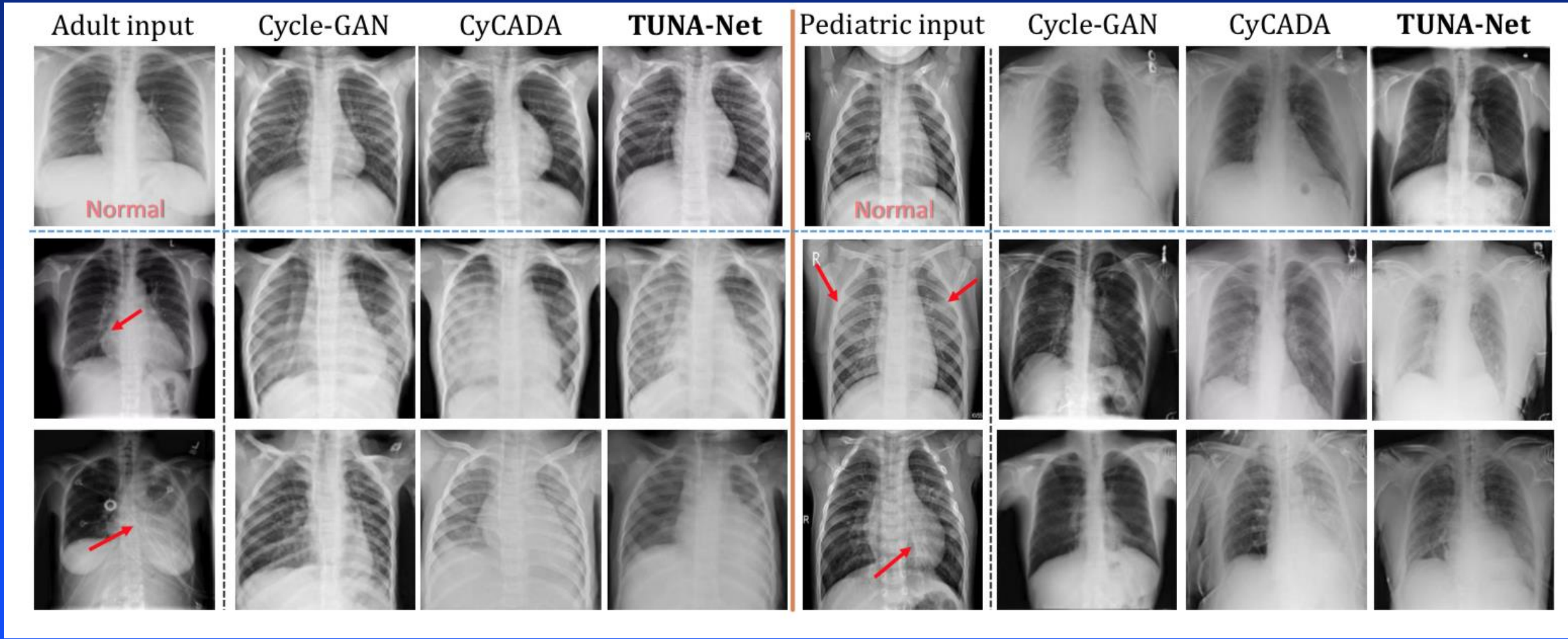






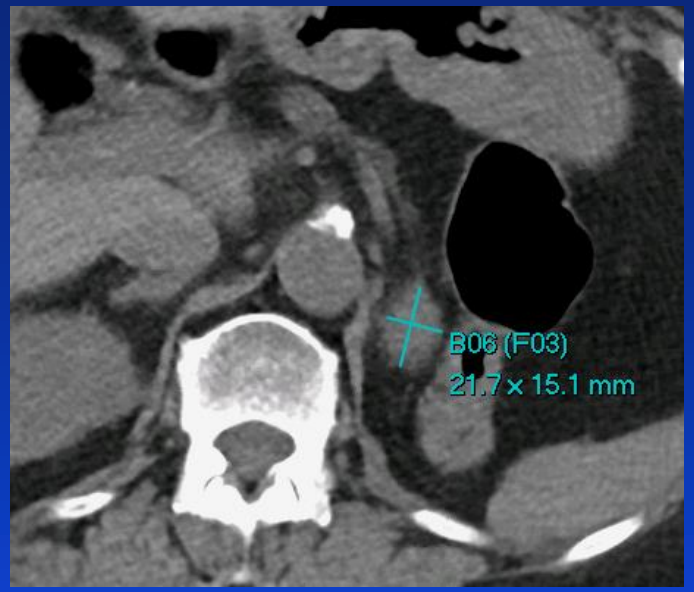
Yuxing Tang et al. npj Digital Medicine 2020

Domain Adaptation with Adversarial Networks

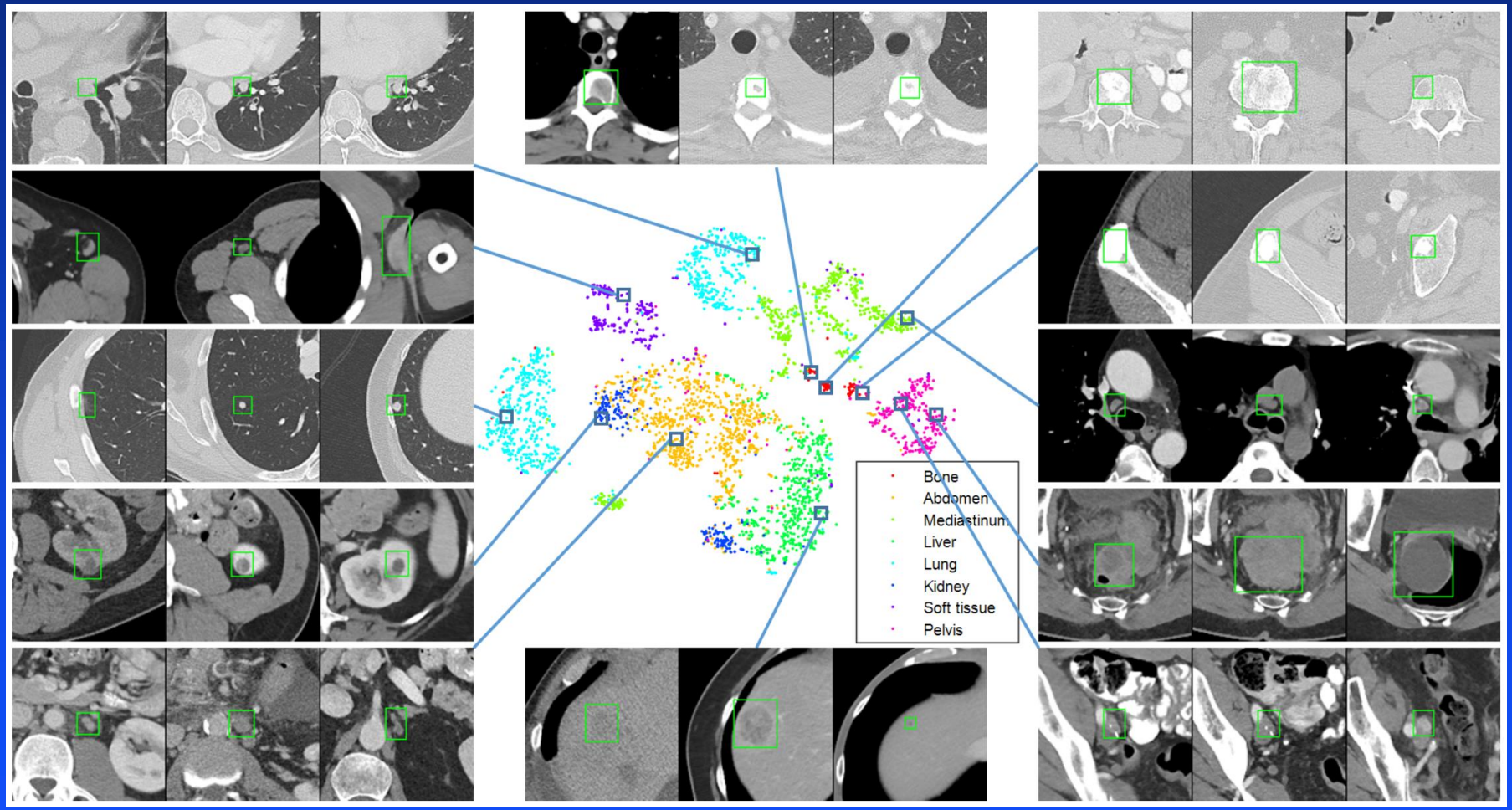


Yuxing Tang et al. MICCAI 2019

Deep Lesion Dataset

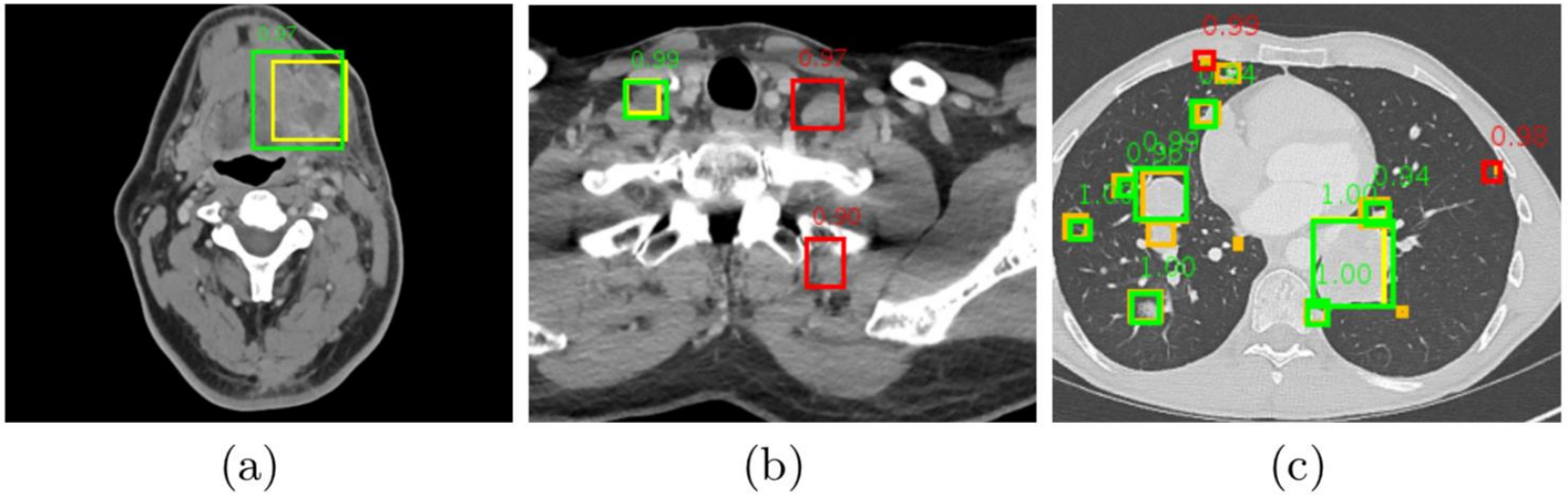


- 32,735 lesions
- 32,120 CT slices
- 10,594 studies
- 4,427 unique patients
- https://nihcc.box.com/v/DeepLesion



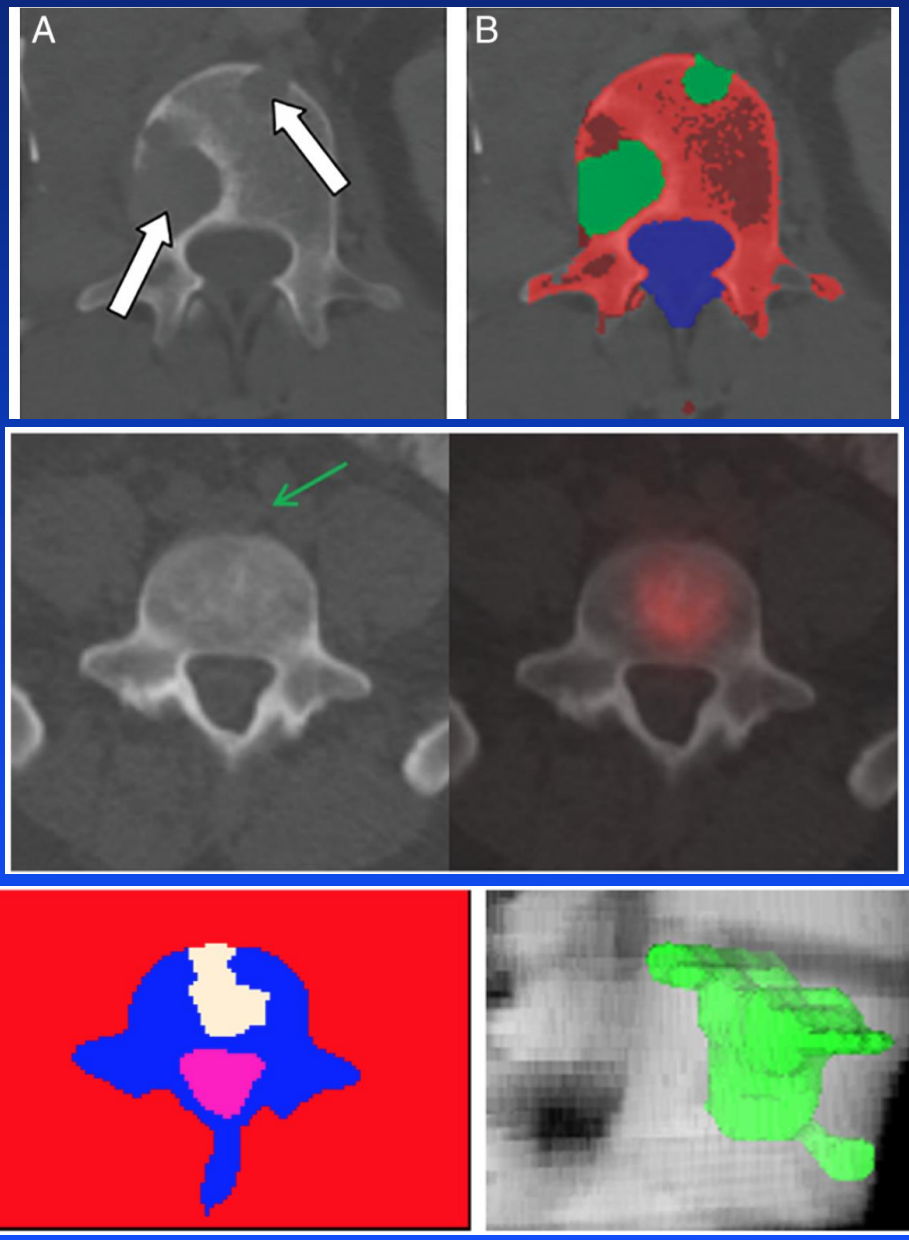
Yan et al. JMI 2018; CVPR 2018; RSNA 2018

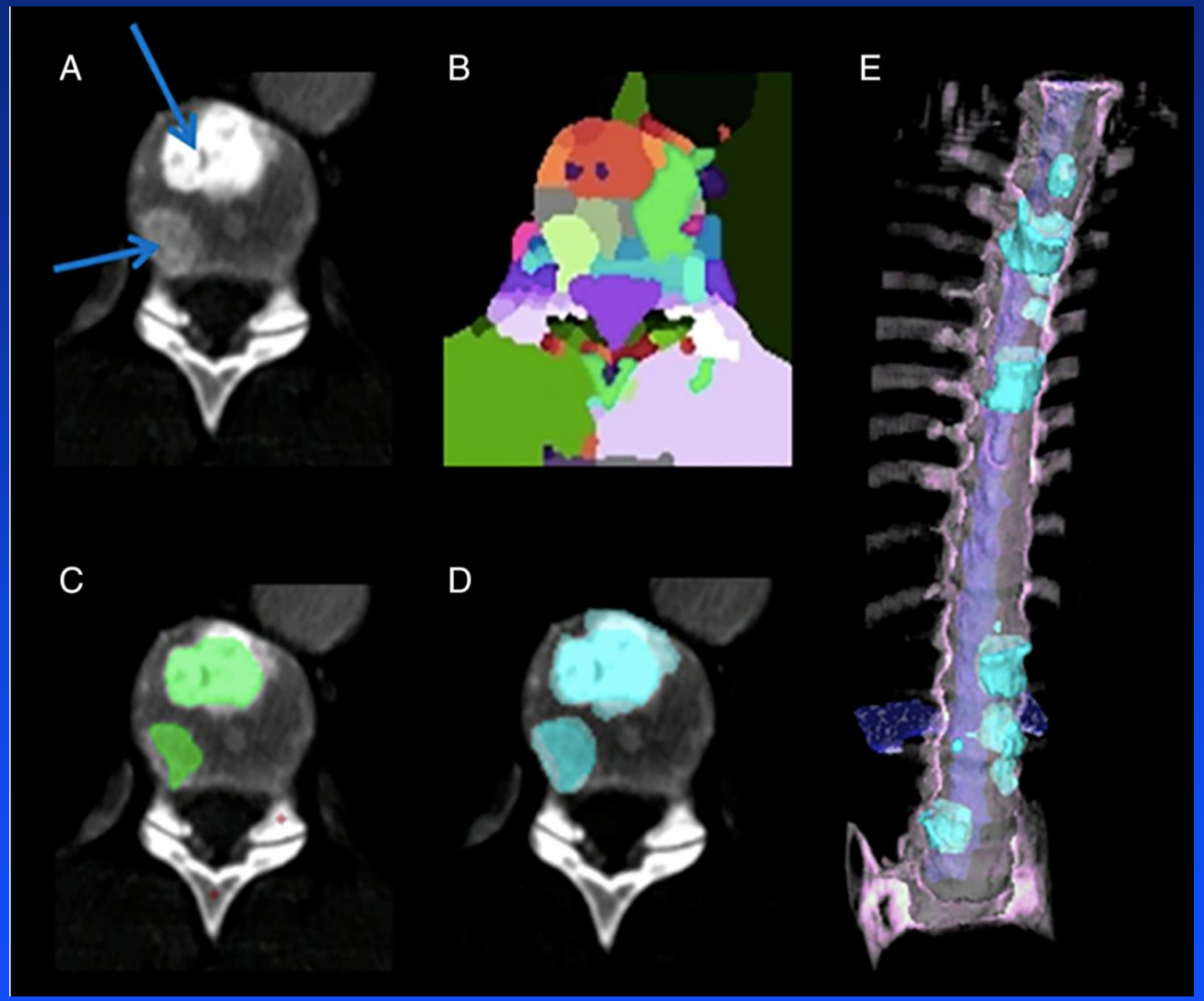
Universal Lesion Detector



Yan et al. MICCAI 2018

Comprehensive Spine Oncology Analysis

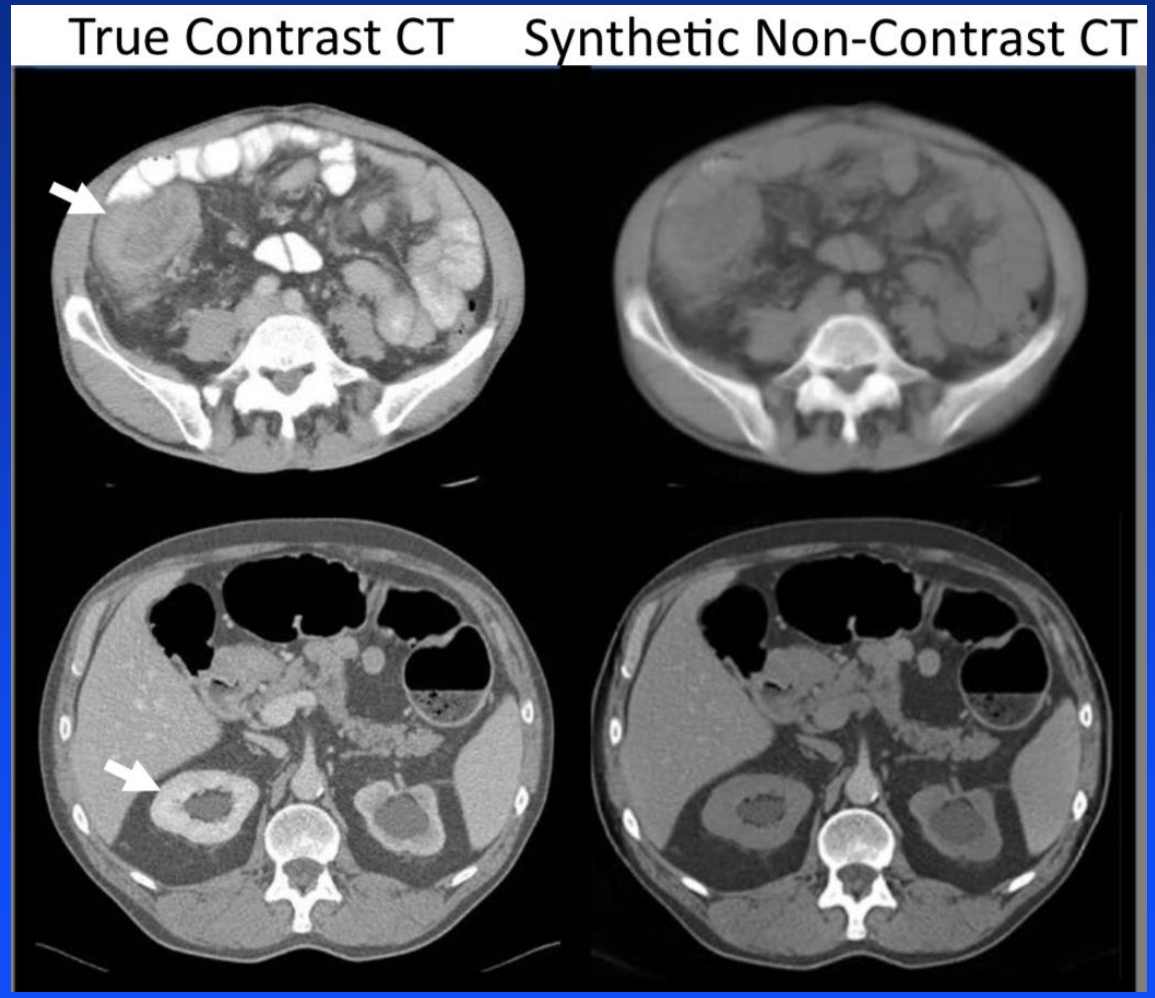




O'Connor et al. Radiology 2007; Yao et al. JMI 2017; Burns et al. JBMR 2020

Challenges & Questions

- Interpretability / explainability
- Brittleness
- Domain shift
- Ethics / Trustworthy AI



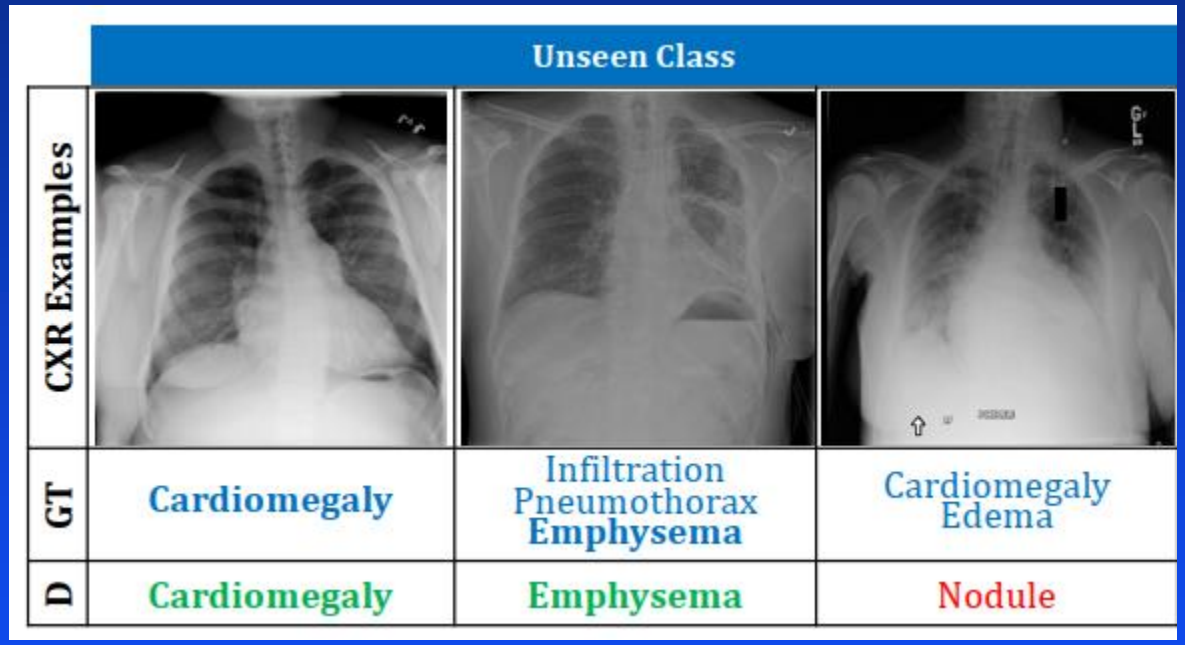
Sandfort et al. Sci Reports 2019

Challenges & Questions

- Dataset annotation is expensive; how to do it much more cost-effectively?
- Multi-institutional data; how to get it?
- Radiologists can diagnose 1000’s of diseases; how to do this with ML?
- Radiologists can do “one-shot” learning, e.g., for rare diseases; how to do this with ML?

Approaches to Using Less Labels

- Zero shot learning
- Few shot learning
- Transfer learning
- Unsupervised learning
- Semi-supervised learning
- Federated learning



A. Paul et al. MICCAI MI3LID 2020

See also A. Paul et al. SPIE MI 2020, MedIA 2021, IEEE TMI 2021, IEEE ISBI 2021

Conclusions

- Rapid developments in AI ➡ Exciting time for medical imaging research and patient care
- Practical clinical benefits in radiology expected

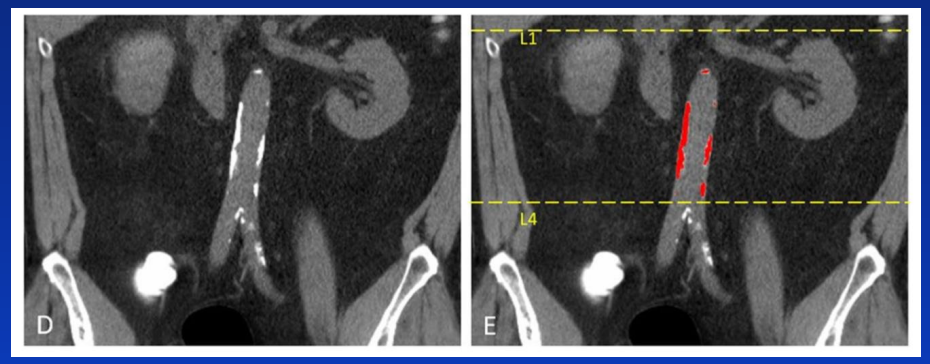
Acknowledgments

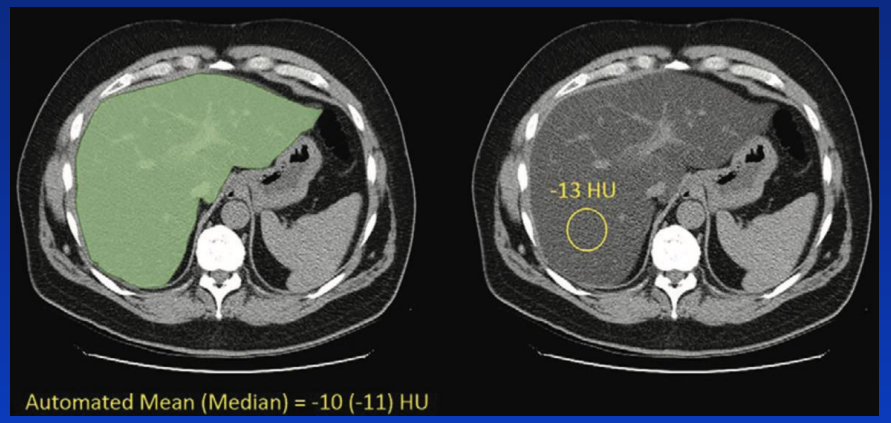
- Le Lu
- Jack Yao
- Jiamin Liu
- Nathan Lay
- Hadi Bagheri
- Holger Roth
- Hoo-Chang Shin
- Xiaosong Wang
- Adam Harrison
- Ke Yan
- Ling Zhang
- Isabella Nogues
- Youbao Tang
- Yuxing Tang
- Veit Sandfort
- Dan Elton
- Nicholas Petrick
- Berkman Sahiner
- Joseph Burns
- Perry Pickhardt
- Mingchen Gao
- Daniel Mollura
- Baris Turkbey
- Peter Choyke
- Matthew Greer
- Brad Wood
- Peter Pinto
- Jin Tae Kwak
- Ruida Cheng
- Jing Xiao
- Yifan Peng
- Zhiyong Lu

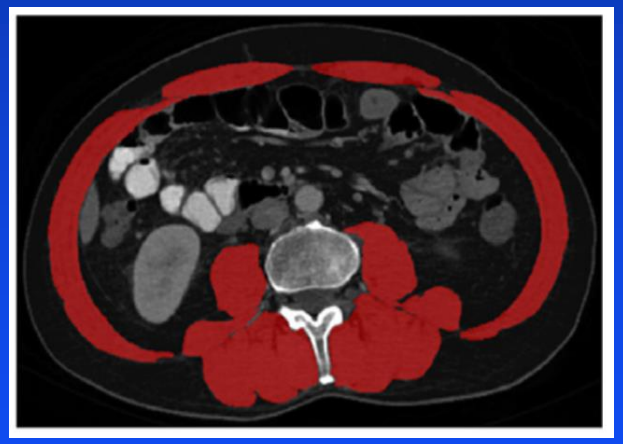
- CC
- NCI
- NLM
- NHLBI
- NIDDK
- NIAID
- FDA
- Mayo Clinic
- DOD
- U. Wisconsin

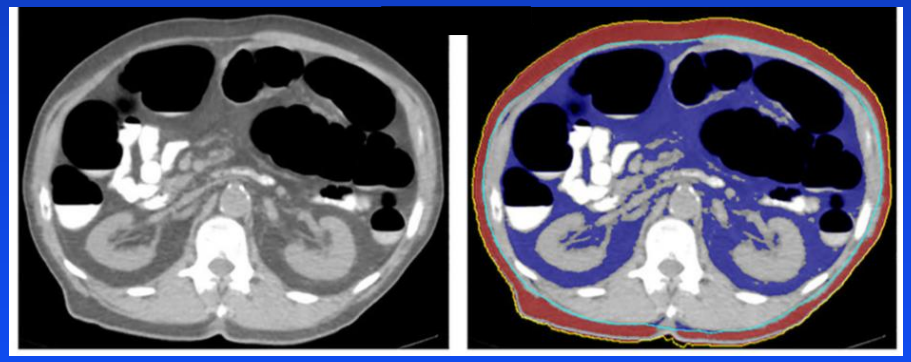
- NIH Fellowship Programs:
  - Fogarty
  - ISTP
  - IRTA
  - BESIP
  - CRTP
  - SIP
- Nvidia for GPU card donations
- CRADA with Ping An

To Learn More …









rms@nih.gov

www.cc.nih.gov/drd/summers.html

github.com/rsummers11

Lee et al. BJR 2018; Burns et al. Acad Radiol 2019; Graffy et al. Abd Radiol 2019; Jang et al. Radiology 2019; Graffy et al. Radiology 2019