Achieving Excellence in Cancer Diagnosis: A Workshop
October 6, 2021
Session 4: Novel Diagnostic Strategies and Tools for Cancer Diagnosis

Novel Technologies and Strategies that Aim to Optimize Diagnosis via Imaging

Maryellen L. Giger, Ph.D.
A. N. Pritzker Distinguished Service Professor of Radiology, Committee on Medical Physics, and the College
The University of Chicago
m-giger@uchicago.edu

Funding and COIs

- Supported in parts by NIH grants CA 195564, CA 166945, and CA 189240; NIH S10 OD025081 Shared Instrument Grant; and The University of Chicago CTSA UL1 TR000430 pilot awards; UChicago Cancer Center Koleseiki Funding and Dancing with Chicago Celebrities Funding; CDAC Grant; c3.ai Grant; NIBIB COVID-19 Contract 75N92020D00021
- MLG is a stockholder in R2/Hologic, shareholder in Qview, and receives royalties from Hologic, GE Medical Systems, MEDIAN Technologies, Riverain Medical, Mitsubishi, and Toshiba.
- MLG is scientific advisor, co-founder, and equity holder in Quantitative Insights, [now Qlarity Imaging] makers of QuantX -- the first FDA-cleared machine learning system for aiding in cancer diagnosis.
- It is the University of Chicago Conflict of Interest Policy that investigators disclose publicly actual or potential significant financial interest that would reasonably appear to be directly and significantly affected by the research activities.

The benefit of a medical imaging examination in terms of its ability to yield an accurate diagnosis depends on:

- Quality of the imaging technology
  - Improvement in standardizing screening mammography (FFDM/Breast Tomosynthesis)
  - New tomographic imaging systems (e.g., CT, MRI, PET, MRI/PET)

Need a good image

- Quality of the interpretation
  - Mainly performed by radiologist
  - Incorporate a computer analysis (AI)

Need a good reader

Medical Image Interpretation

Medical images are meaningless grayscale/colorscale patterns unless “viewed and analyzed” by an intelligent observer

- Radiologist, Computer (AI), or Combination of human & computer (AI-aided)

Tasks of the Human eye-brain system

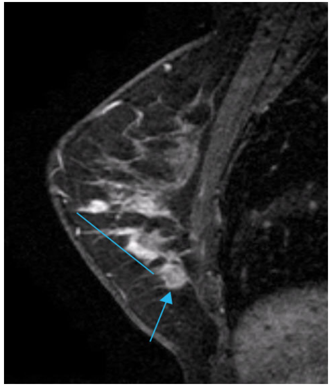
- Finding/locating a signal in an image
- Characterizing/classifying/diagnosing the signal as disease or non-disease
- Clinical decision making on patient management through integrated diagnostics (monitoring)

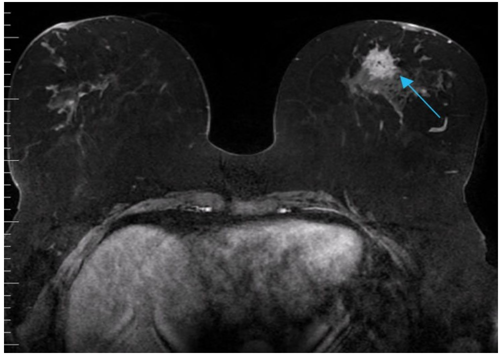
Tasks of AI (computer vision, radiomics, machine learning, deep learning)

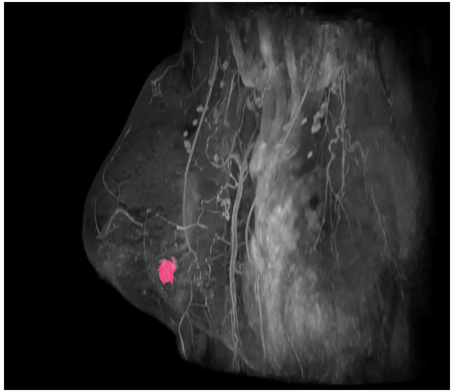
- Similar – converting images to quantitative values
- Need to know the clinical task!!!!

Medical Imaging & AI in Precision Medicine & Oncology

- The focus is AI for images that are "**clinically & routinely**" obtained on the population.
- We want to ask questions about the relationships between features "seen" in medical images and the biology of **cancer** so that eventually we can **detect/diagnose cancer early** and **give the right patient the right treatment at the right time**.
- And to **improve the efficiency** and workflow of medical imaging interpretation.







AI can be applied at many Stages along the Medical Imaging Chain

- Imaging Source
- Subject (patient, animal, tissue, cells)
- Imaging detector system (e.g., optimizing detector parameters)
- Contrast media & Imaging probes
- Image presentation (e.g., hanging protocols)
- Image reconstruction (e.g., tomosynthesis)
- Image processing (e.g., image denoising)
- Quantitative image analysis/CADx/Radiomics (e.g., Density estimation, detection, diagnosis, prognosis, therapeutic response)
- Image/Data integration
- Image/Data output display/interface/GUI

AI can be applied at many Stages along the Medical Imaging Chain

- Imaging Source
- Subject (patient, animal, tissue, cells)
- Imaging detector system (e.g., optimizing detector parameters)
- Contrast media & Imaging probes
- Image presentation (e.g., hanging protocols)
- Image reconstruction (e.g., tomosynthesis)
- Image processing (e.g., image denoising)
- **Quantitative image analysis/CADx/Radiomics (e.g., Density estimation, detection, diagnosis, prognosis, therapeutic response)**
- Image/Data integration
- Image/Data output display/interface/GUI

Overall Considerations for AI for Cancer Diagnosis in Medical Imaging

Answer to some Medical Question
(e.g., risk assessment, detection, diagnosis, prognosis, therapy response)

Data
(images & clinical, demographics)

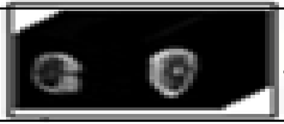


Image Acquisition
(physical parameters, variations and harmonization needs)

AI
Algorithm (human-engineered radiomics or deep learning)

AI use as an **aid** by radiologist
CADe, CADx. AI-aided
(secondary or concurrent reader)

AI use as a **primary** reader
triage (CADt), "rule out"

Autonomous AI
replaces human

Appropriate Metrology and Evaluation Methods
(e.g., standalone evaluation and evaluation of the performance of the enduser)

Overall Considerations for AI for Cancer Diagnosis in Medical Imaging

Answer to some Medical Question
(e.g., risk assessment, detection, diagnosis, prognosis, therapy response)

Data
(images & clinical, demographics)

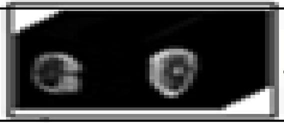


Image Acquisition
(physical parameters, variations and harmonization needs)

AI
Algorithm (human-engineered radiomics or deep learning)

AI use as an **aid** by radiologist
CADe, CADx. AI-aided
(secondary or concurrent reader)

All require continuing resources & research

Autonomous AI
replaces human

Appropriate Metrology and Evaluation Methods
(e.g., standalone evaluation and evaluation of the performance of the enduser)

Overall Considerations for AI for Cancer Diagnosis in Medical Imaging

Answer to some Medical Question
(e.g., risk assessment, detection, diagnosis, prognosis, therapy response)

Data
(images & clinical, demographics)

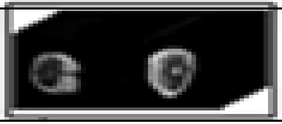


Image Acquisition
(physical parameters, variations and harmonization needs)

AI
Algorithm (human-engineered radiomics or deep learning)

AI use as an **aid** by radiologist
CADe, CADx. AI-aided
(secondary or concurrent reader)

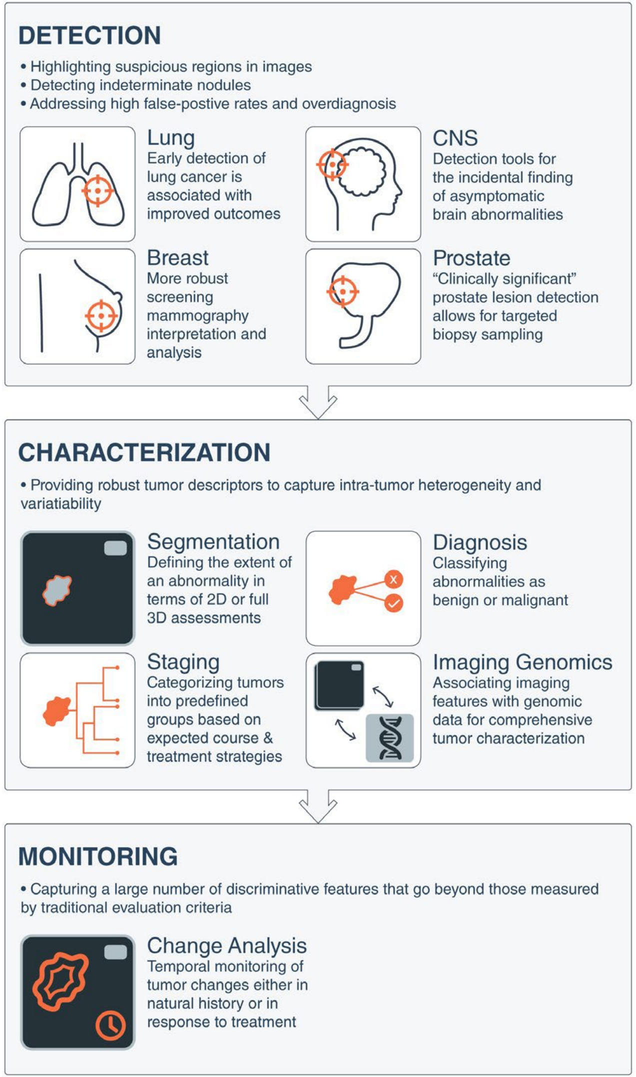
AI use as a **primary** reader
triage (CADt), "rule out"

Autonomous AI
replaces human

Appropriate Metrology and Evaluation Methods
(e.g., standalone evaluation and evaluation of the performance of the enduser)

AI can be applied in many Cancer Imaging Decision-Making Tasks

- Risk Assessment
- Screening – CADe, CADt
- Diagnosis -- CADx
- Prognosis (subtyping)
- Treatment planning
- Assessing treatment response
- Patient management and monitoring
- Disease Discovery (radiogenomics; multi-omics)



Bi WL, Hosny A, Schabath MB, Birkbak NJ, Mehrtash A, Giger ML, Allison T, Arnaout O, Abbosh C, Dunn IF, Mak RH, Tamimi RM, Tempany CM, Swanton C, Hoffmann U, Schwartz LH, Gillies RJ, Huang R, Aerts HJWL: Artificial intelligence in cancer imaging: clinical challenges and applications. CA: A Cancer Journal for Clinicians 2019 Mar;69(2):127-157. doi: 10.3322/caac.21552. Epub Feb 5, 2019.

Overall Considerations for AI for Cancer Diagnosis in Medical Imaging

Answer to some Medical Question
(e.g., risk assessment, detection, diagnosis, prognosis, therapy response)

Data
(images & clinical, demographics)

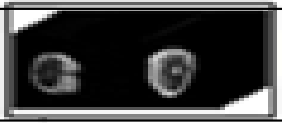


Image Acquisition
(physical parameters, variations and harmonization needs)

AI
Algorithm (human-engineered radiomics or deep learning)

AI use as an **aid** by radiologist
CADe, CADx. AI-aided (secondary or concurrent reader)

AI use as a **primary** reader
triage (CADt), "rule out"

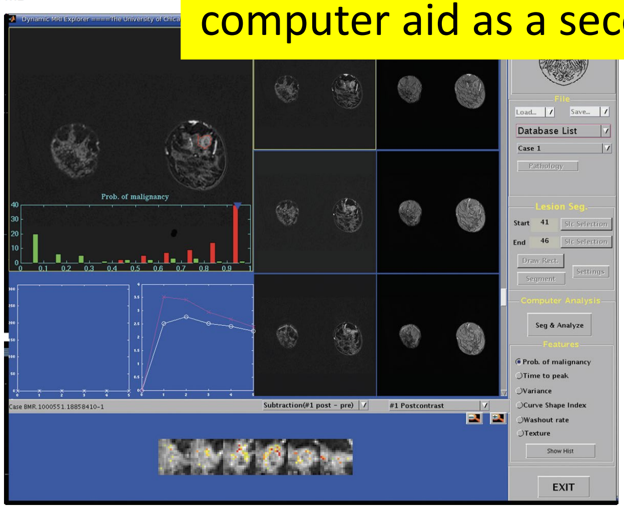
Autonomous AI
replaces human

Appropriate Metrology and Evaluation Methods
(e.g., standalone evaluation and evaluation of the performance of the enduser)

Evaluation of Clinical Breast MR Imaging Performed with Prototype Computer-aided Diagnosis Breast MR Imaging Workstation: Reader Study[1]

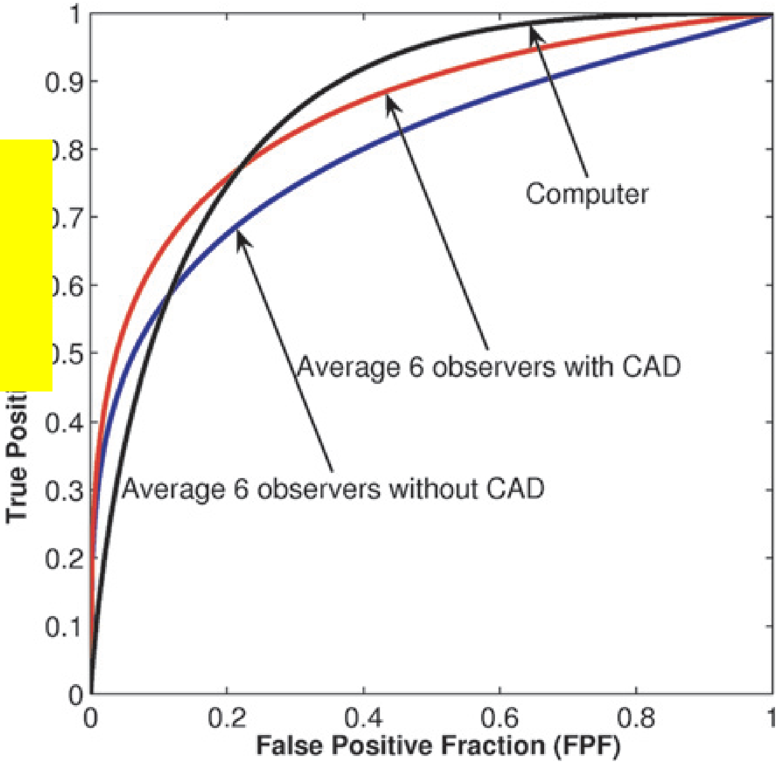
Akiko Shimauchi, MD, PhD
Maryellen L. Giger, PhD
Neha Bhooshan, PhD
Li Lan, MS
Lorenzo L. Pesce, PhD
John K. Lee, MD
Hiroyuki Abe, MD, PhD
Gillian M. Newstead, MD

2011



i.e., Radiologists improved in their performance of <u>characterizing and diagnosing</u> lesions when using the computer aid as a second reader

CADx: Task of Distinguishing between Malignant & Benign Lesions on **Breast MRI**



Computer-Aided Detection in Breast Cancer Screening

Change from Using the Computer Output as a Second Reader to a Potential Independent Reader

Support from Computer as a Second Reader

Rodriguez-Ruiz A, et al Radiology, 10.1148/radiol.2018181371, 2019

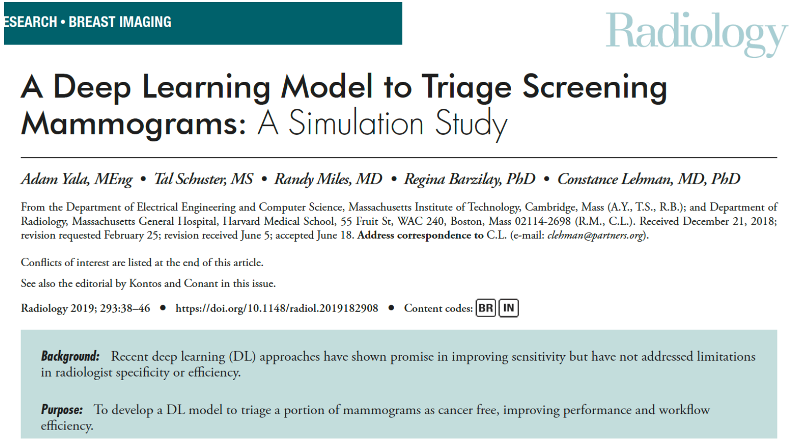
- The system uses deep learning convolutional neural networks and features classifiers / image analysis algorithms to indicate calcifications and soft-tissue lesions
- Trained, validated, & tested on 9000 cancers
- Radiologists can use an interactive decision support mode as well as traditional CAD
- Radiologists' unaided AUC = 0.866
- Radiologists' aided AUC = 0.886
- Statistically improved radiologists' performance

Potential future use as Standalone Independent Reader

Rodeiguez-Ruiz, A, et al, JNCI, 10.1093/jnci/djy222, 2019

- Comparison with 101 (unaided) radiologists vs. computer alone
- 2652 mammographic exams (653 malignant)
- AI system was statistically noninferior to that of the average of the 101 radiologist
- Radiologists' unaided AUC = 0.814
- Computer alone AUC = 0.840

CADe in Breast Cancer Screening



ESEARCH • BREAST IMAGING

Radiology

A Deep Learning Model to Triage Screening Mammograms: A Simulation Study

Adam Yala, MEng • Tal Schuster, MS • Randy Miles, MD • Regina Barzilay, PhD • Constance Lehman, MD, PhD

From the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Mass (A.Y., T.S., R.B.); and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, 55 Fruit St, WAC 240, Boston, Mass 02114-2698 (R.M., C.L.). Received December 21, 2018; revision requested February 25; revision received June 5; accepted June 18. **Address correspondence to** C.L. (e-mail: *clehman@partners.org*).

Conflicts of interest are listed at the end of this article.

See also the editorial by Kontos and Conant in this issue.

Radiology 2019; 293:38–46 • https://doi.org/10.1148/radiol.2019182908 • Content codes: BR IN

Background: Recent deep learning (DL) approaches have shown promise in improving sensitivity but have not addressed limitations in radiologist specificity or efficiency.

Purpose: To develop a DL model to triage a portion of mammograms as cancer free, improving performance and workflow efficiency.

Change from

Second Reader to Independent Reader

- Role in **triaging** **(CADt) – “rule out”**
- Role in improving efficiency

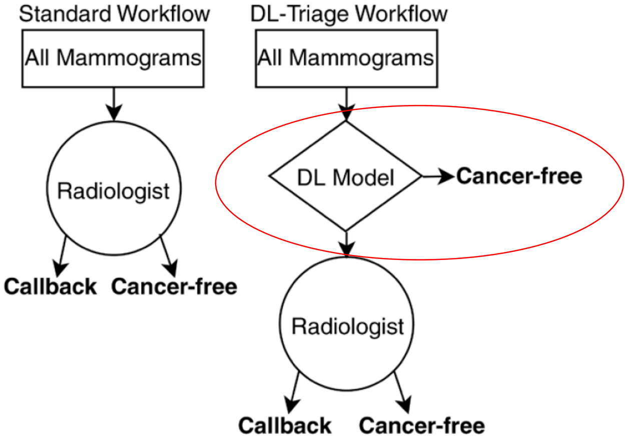


Figure 2: Diagram illustrates experimental setup for triage analysis. In standard scenario, radiologists read all mammograms. In deep learning (DL)–triage scenario, radiologists only read mammograms above model cancer-free threshold. To simulate both scenarios, original interpreting radiologist’s assessment on test set was used for radiologist read.

- Trained a deep learning model to triage mammograms as cancer free and showed that their model could
- In the **simulation study**, 20% of mammograms were not need to be send by humans.
- Showed improvement in radiologist efficiency and specificity without harming sensitivity.

Overall Considerations for AI for Cancer Diagnosis in Medical Imaging

Answer to some Medical Question
(e.g., risk assessment, detection, diagnosis, prognosis, therapy response)

Data
(images & clinical, demographics)

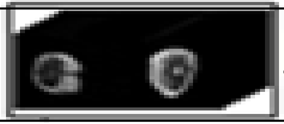


Image Acquisition
(physical parameters, variations and harmonization needs)

AI
Algorithm (human-engineered radiomics or deep learning)

AI use as an **aid** by radiologist
CADe, CADx. AI-aided
(secondary or concurrent reader)

AI use as a **primary** reader
triage (CADt), “rule out”

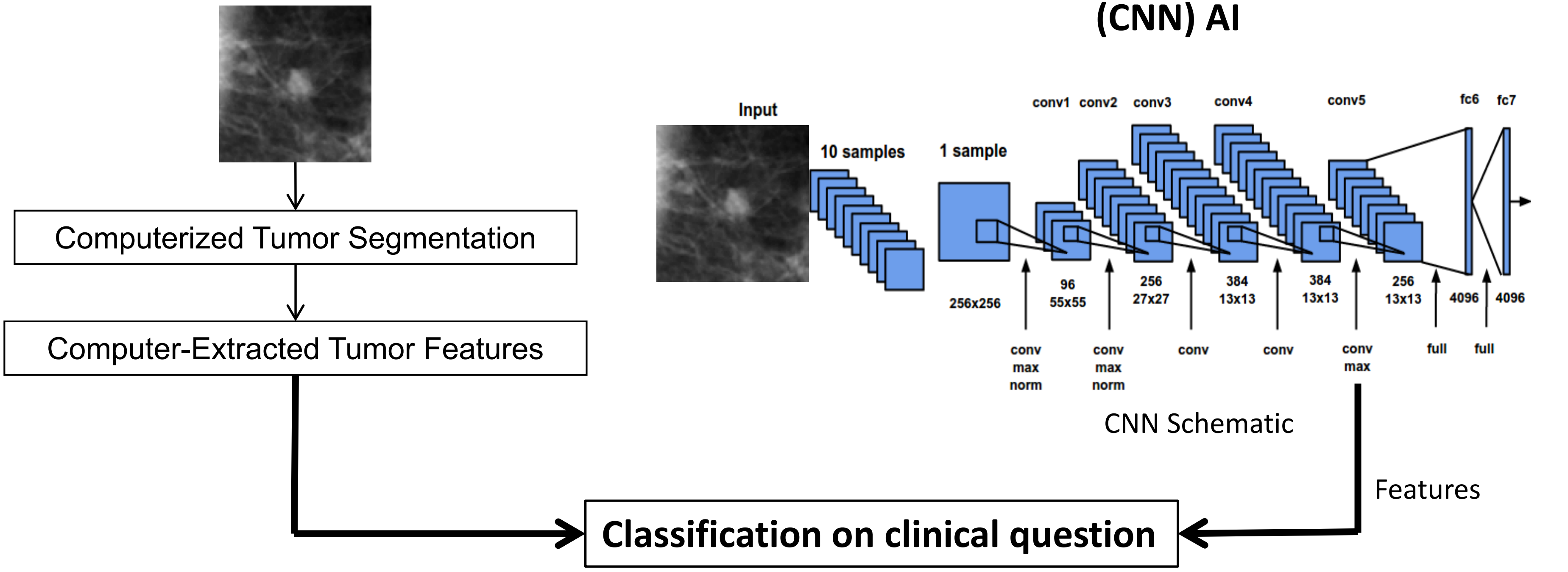
Autonomous AI
replaces human

Appropriate Metrology and Evaluation Methods
(e.g., standalone evaluation and evaluation of the performance of the enduser)

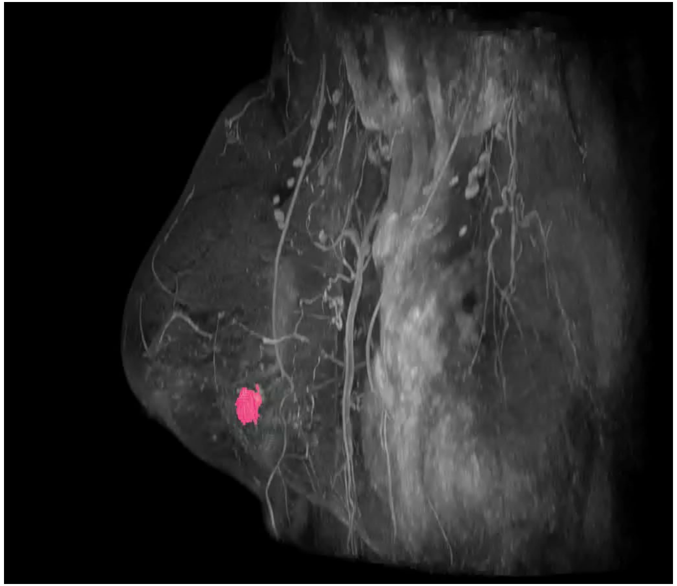
Comparison of Human-Engineered AI and Deep Transfer Learning in distinguishing between malignant and benign breast lesions

Human-Engineered AI

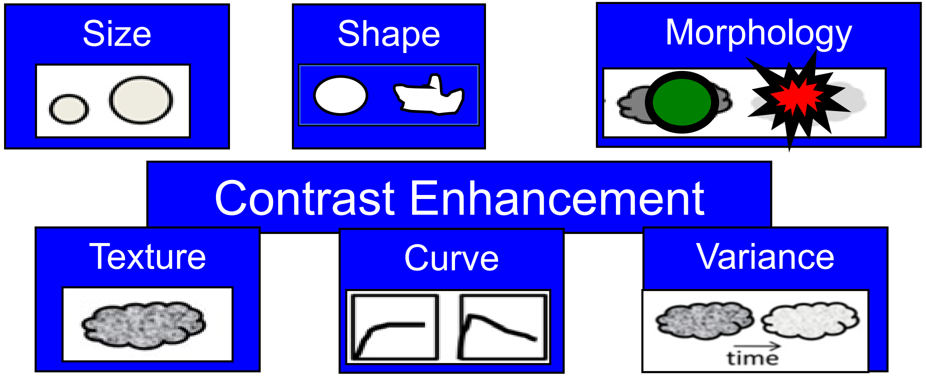
Deep Learning: Convolutional Neural Networks (CNN) AI



Use case example: Analysis of Breast Cancer on MRI using **Human-Engineered Features**



- After the lesion is **automatically segmented**, image features (i.e., mathematical descriptors; radiomics) are **extracted** from the lesion.

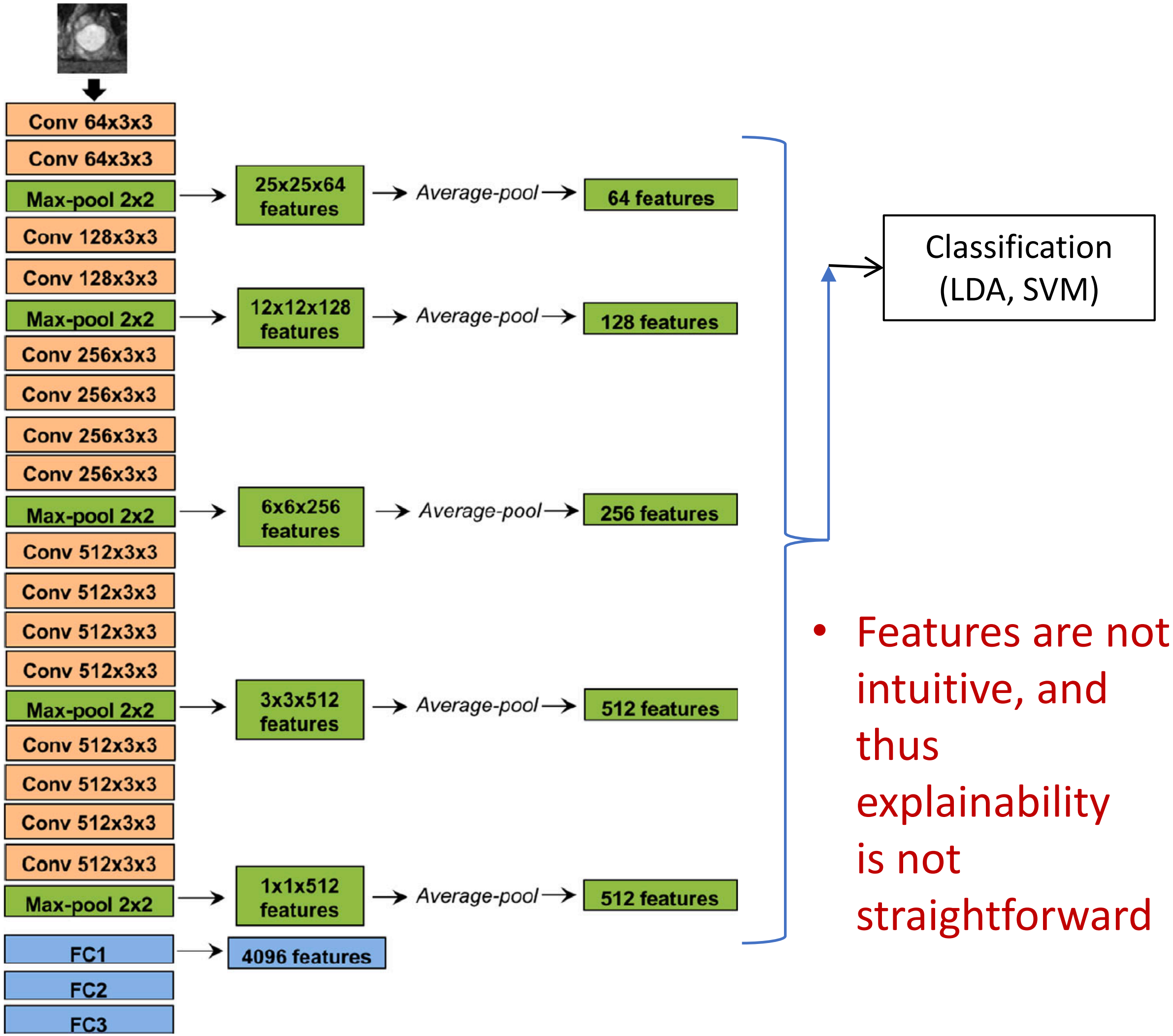


- Features are intuitive, and thus output is more **explainable**
- 3D/4D features then merged by a **classifier** (e.g., LDA, SVM) to yield a **signature** indicating an estimate of the **likelihood of malignancy, estimate of the severity of disease, or predicted response**

Use case example: Analysis of Breast Cancer on MRI using **Transfer Deep Learning**

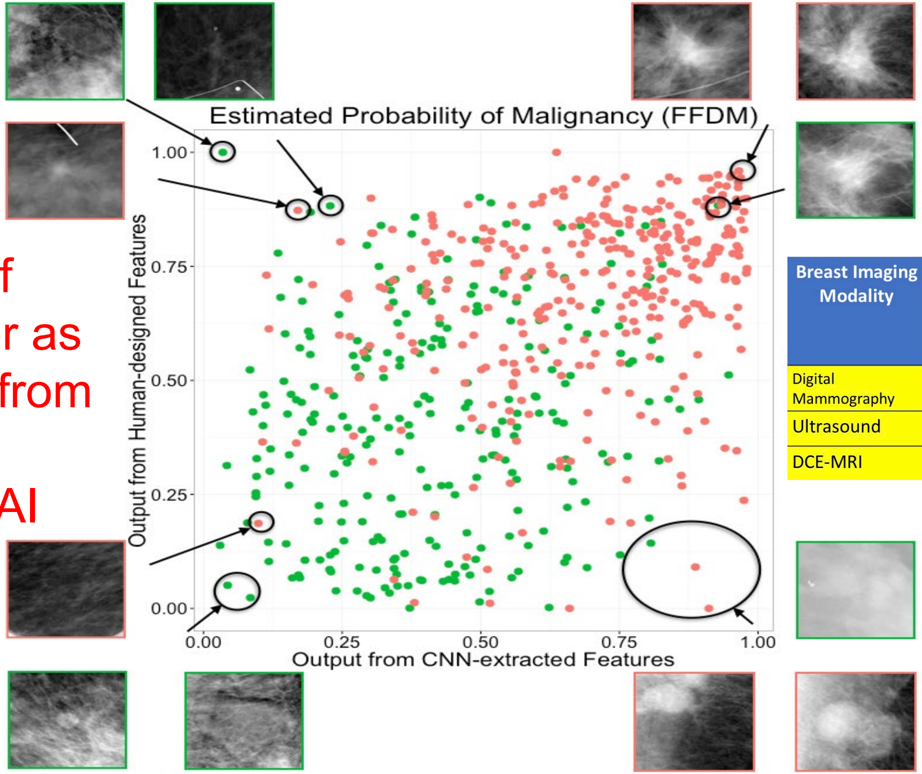
- Task of distinguishing between cancers and non cancers
- Transfer learning reduces the number of cases required

Antropova N, Huynh BQ, Giger ML: A deep fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Medical Physics online doi.org/10.1002/mp.12453, 2017.



- Features are not intuitive, and thus explainability is not straightforward

Human-Engineered CADx/Radiomics & Deep Learning CADx/Radiomics
(task of distinguishing between cancers and non cancers)



RED = CANCER

GREEN = Non-CANCER

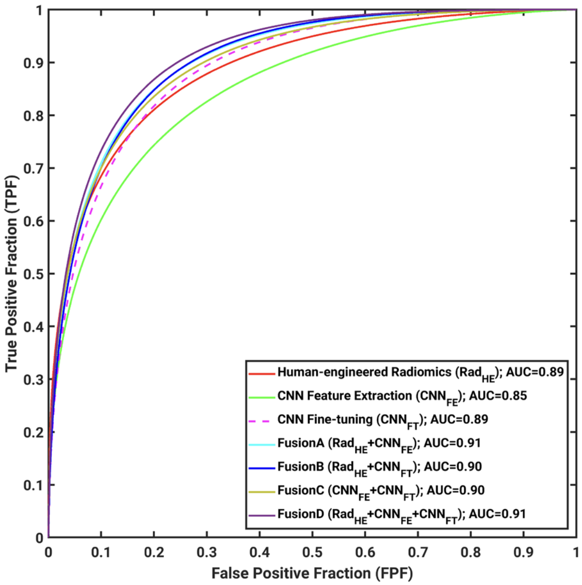
Likelihood of being cancer as determined from human-engineered AI

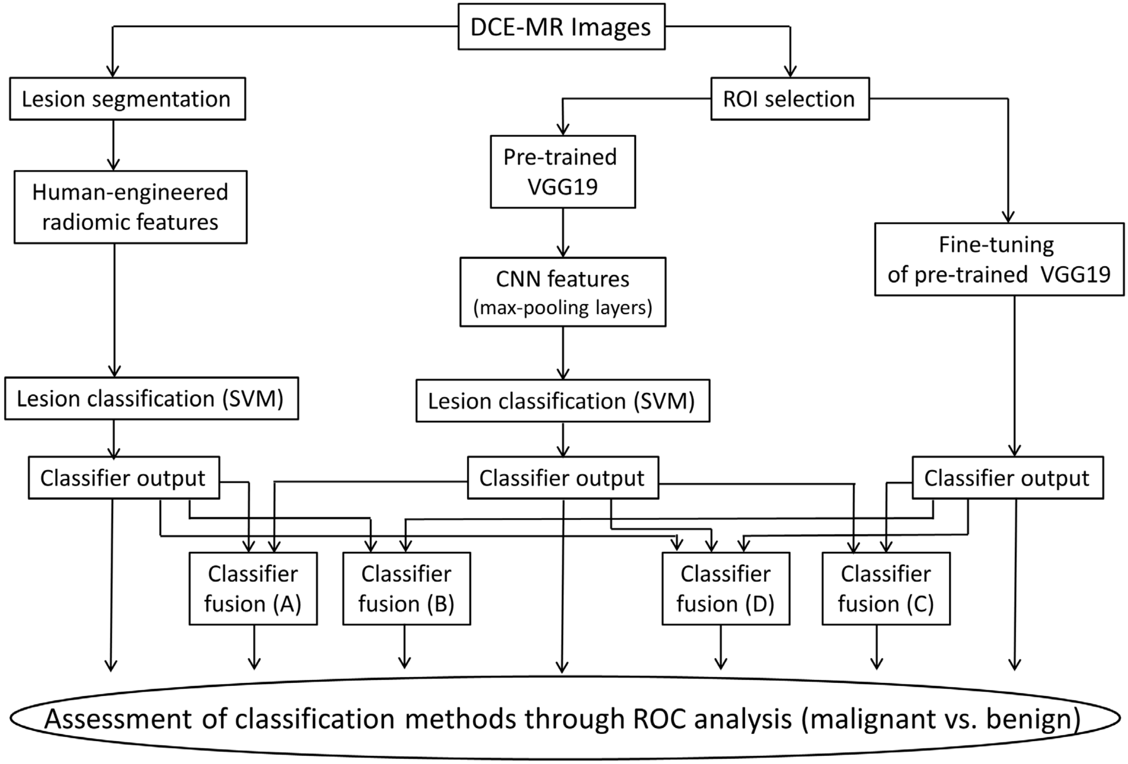
Breast Imaging Modality	Number of Cases	Human-Engineered CADx (AUC)	Deep Transfer Learning CNN (AUC)	Fusion of Human-Engineered & CNN (AUC)
Digital Mammography	245	0.79	0.81	**0.86**
Ultrasound	1125	0.84	0.87	**0.90**
DCE-MRI	690	0.86	0.87	**0.89**

Likelihood of being cancer as determined from deep learning AI

Antropova N, Huynh BQ, Giger ML: A deep fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Medical Physics online doi.org/10.1002/mp.12453, 2017.

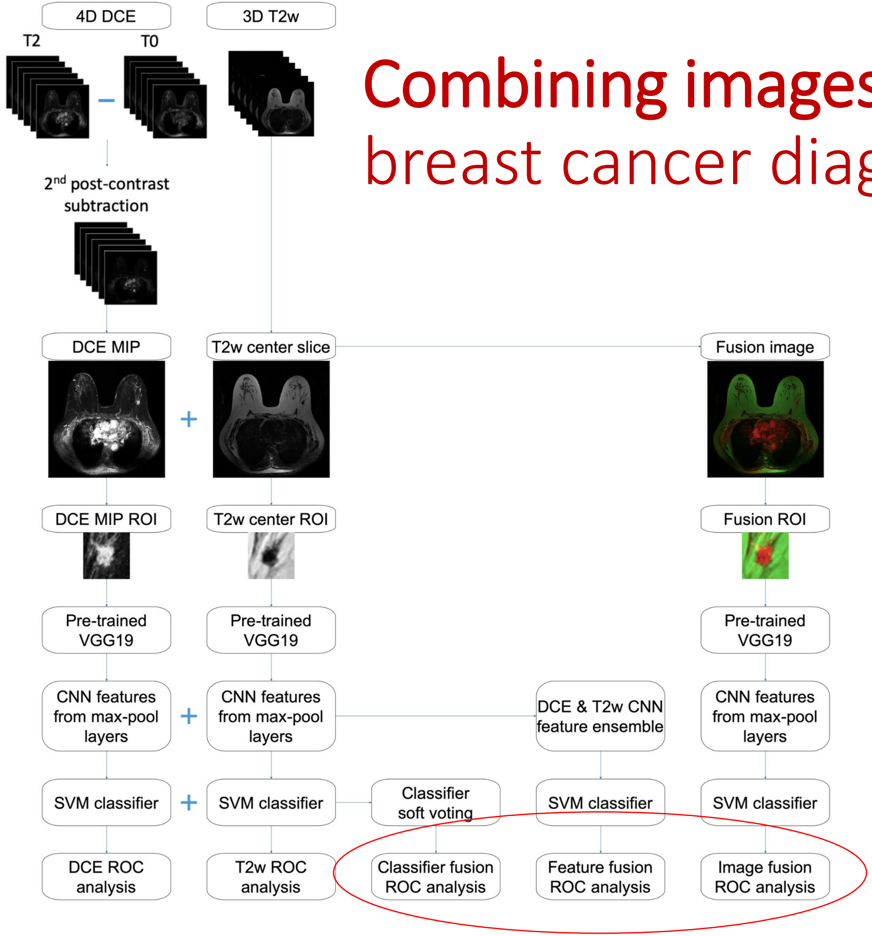
Combining Multiple AI Methods for Breast Cancer Diagnosis

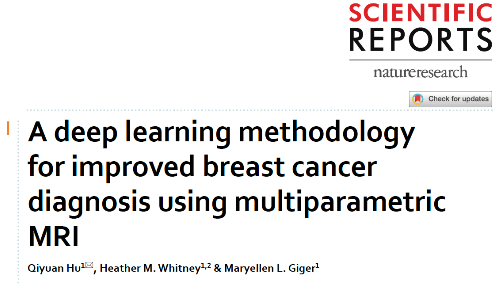




Whitney H*, Li H*, Ji Y, Liu P, Giger ML: Comparison of breast MRI tumor classification using human-engineered radiomics, transfer learning from deep convolutional neural networks, and fusion methods. Proceedings of the IEEE, DOI: [10.1109/JPROC.2019.2950187](https://doi.org/10.1109/JPROC.2019.2950187), 2019.

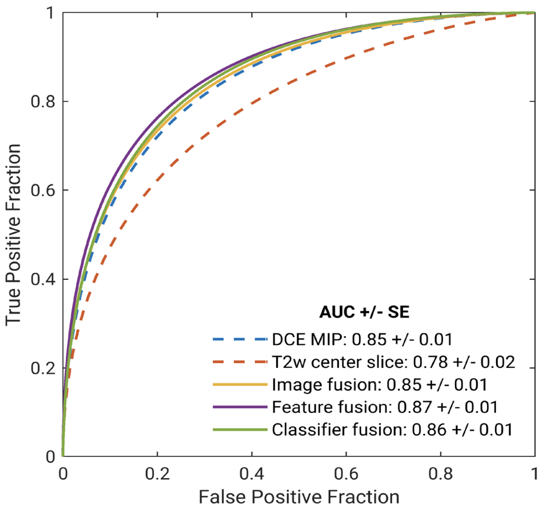
Combining images in deep learning AI for improved breast cancer diagnosis using multiparametric MRI




SCIENTIFIC REPORTS
natureresearch
Check for updates

A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI

Qiyuan Hu[1], Heather M. Whitney[1,2] & Maryellen L. Giger[1]



Hu Q, Whitney HM, Giger ML. A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI. Sci Rep. 2020 Jun 29;10(1):10536. doi: 10.1038/s41598-020-67441-4.

Drukker *et al. Cancer Imaging* (2018) 18:12
https://doi.org/10.1186/s40644-018-0145-9

Cancer Imaging

RESEARCH ARTICLE Open Access

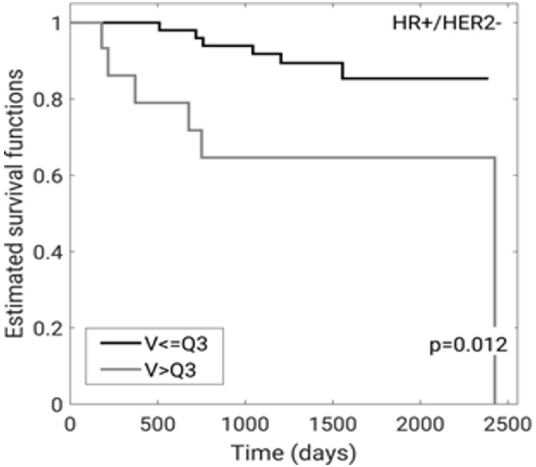
Most-enhancing tumor volume by MRI radiomics predicts recurrence-free survival "early on" in neoadjuvant treatment of breast cancer

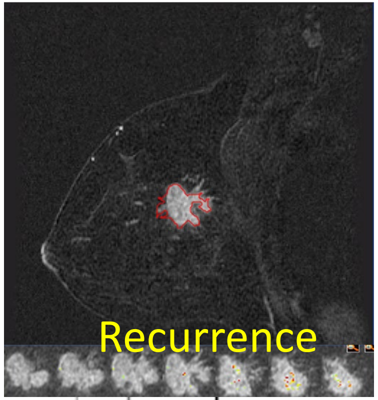
Karen Drukker*, Hui Li, Natalia Antropova, Alexandra Edwards, John Papaioannou and Maryellen L. Giger

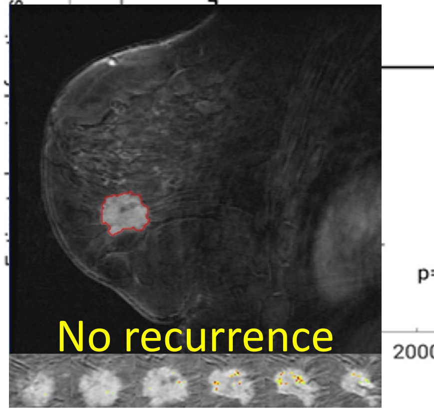
Transferring AI developments from use in diagnosis to use in prognosis and treatment response

Most-enhancing Tumor Volume Predicts Recurrence-free Survival

- Applied automatic calculation of quantitative radiomics to cases from the I-SPY 1 (ACRIN 6657) study of dynamic contrast-enhanced MR images.
- AI in pretreatment prediction of response to neoadjuvant chemotherapy; risk of recurrence, recurrence-free survival



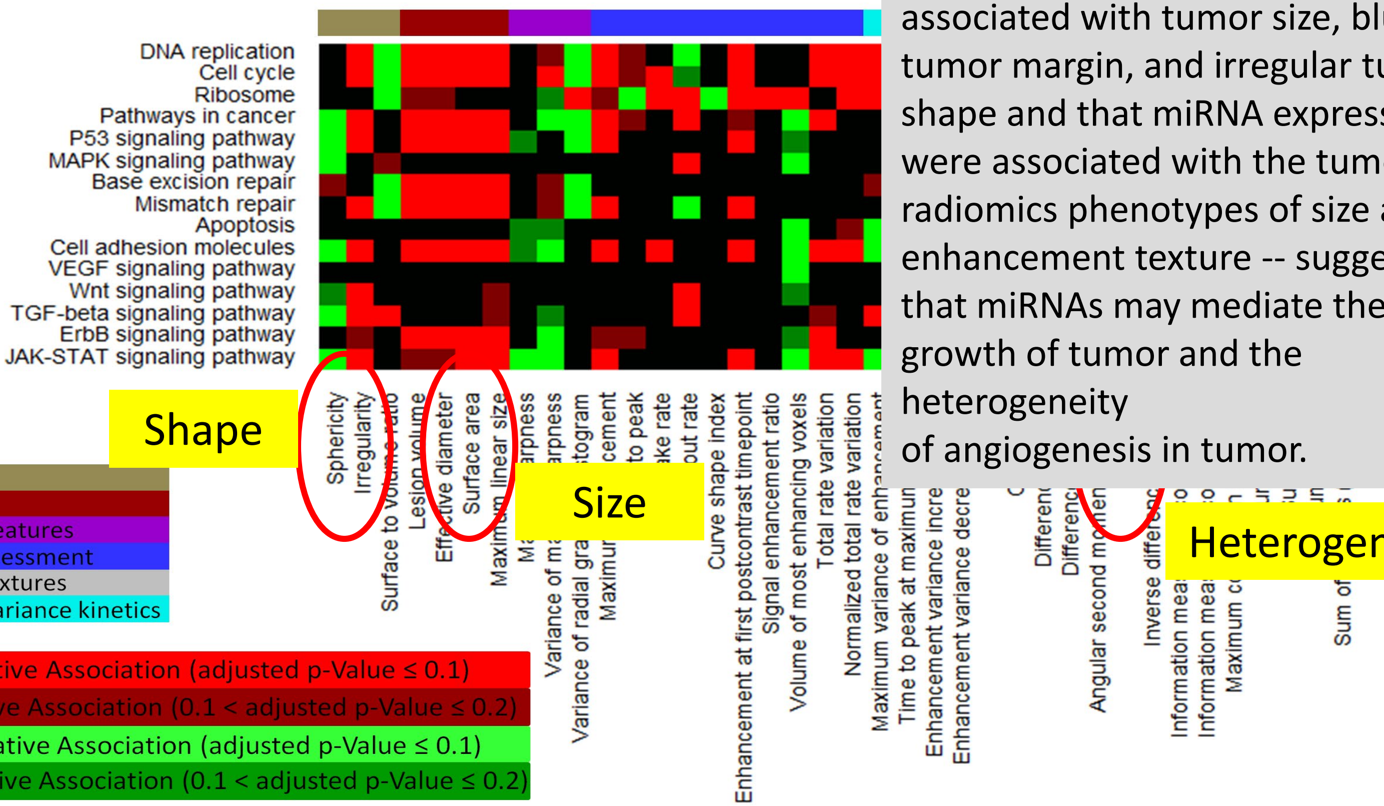




Imaging Genomics – using virtual biopsies

Pathway transcriptional activities associated with MRI quantitative r

Transcriptional activities of various genetic pathways were positively associated with tumor size, blurred tumor margin, and irregular tumor shape and that miRNA expressions were associated with the tumor radiomics phenotypes of size and enhancement texture -- suggesting that miRNAs may mediate the growth of tumor and the heterogeneity
of angiogenesis in tumor.



Shape features
Size features
Morphological features
Kinetic curve assessment
Enhancement textures
Enhancement-variance kinetics

Significant Positive Association (adjusted p-Value ≤ 0.1)
Marginal Positive Association (0.1 < adjusted p-Value ≤ 0.2)
Significant Negative Association (adjusted p-Value ≤ 0.1)
Marginal Negative Association (0.1 < adjusted p-Value ≤ 0.2)

Zhu Y, Li H, … Giger ML*, Ji Y*: Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma.
Nature – Scientific Reports 5:17787 (2015)

Why after decades, is research in cancer imaging AI still being conducted and papers are still being published for detection, diagnosis, prognosis, and assessing response to therapy?

Overall Considerations for AI for Cancer Diagnosis in Medical Imaging

Answer to some Medical Question
(e.g., risk assessment, detection, diagnosis, prognosis, therapy response)

Data
(images & clinical, demographics)

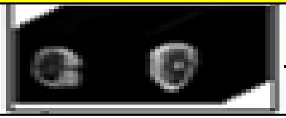


Image Acquisition
(physical parameters, variations and harmonization needs)

AI
Algorithm (human-engineered radiomics or deep learning)

AI use as an **aid** by radiologist
CADe, CADx. AI-aided
(secondary or concurrent reader)

AI use as a **primary** reader
triage (CADt), “rule out”

Autonomous AI
replaces human

Appropriate Metrology and Evaluation Methods
(e.g., standalone evaluation and evaluation of the performance of the enduser)

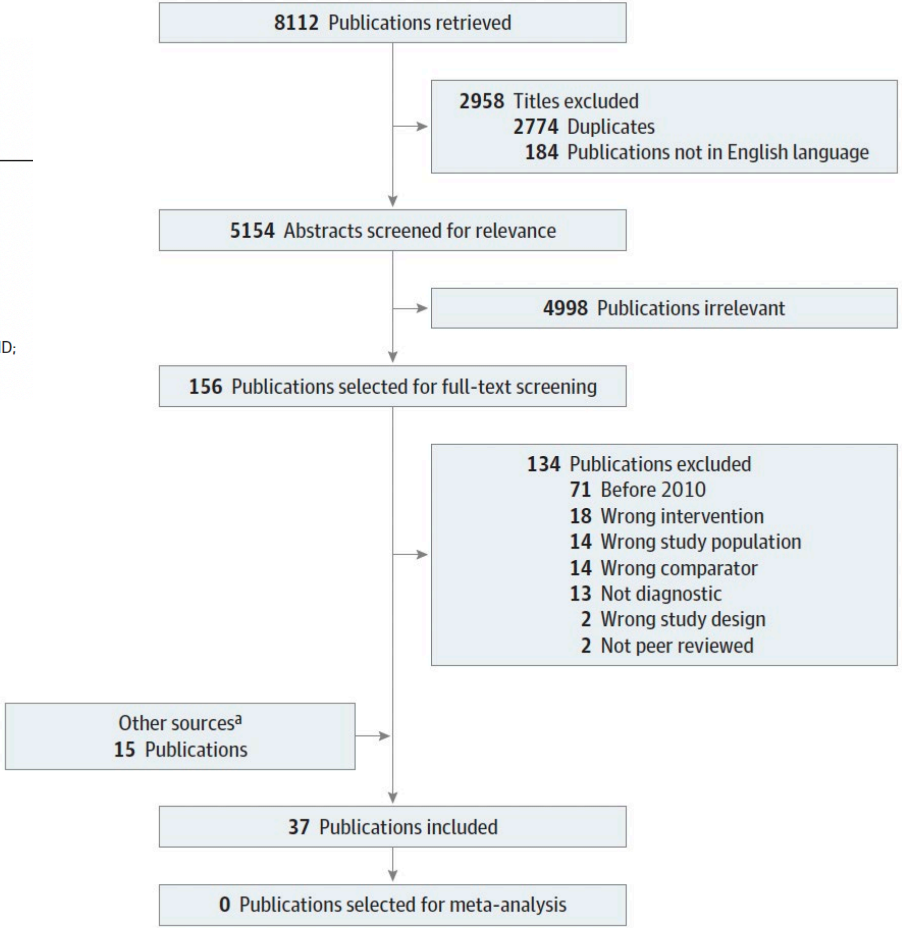
JAMA Network | Open

Original Investigation | Health Informatics

Association of Clinician Diagnostic Performance With Machine Learning–Based Decision Support Systems
A Systematic Review

Baptiste Vasey, MMed; Stephan Ursprung, MMed; Benjamin Beddoe, BSc; Elliott H. Taylor, BSc; Neale Marlow, MBBS; Nicole Bilbro, MD; Peter Watkinson, MD; Peter McCulloch, MD

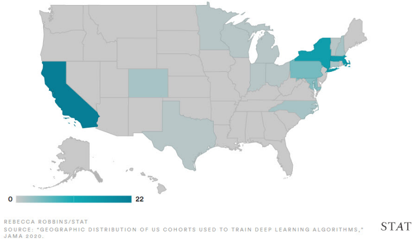
- This systematic review found no robust evidence that the use of ML-based algorithms was associated with better clinician diagnostic performance.
- The evidence for any conclusion was weak because of a high risk of bias in many of the studies and a low number of study participants
- Almost half of all results reported with statistical significance showed no significant difference in performance with or without the use of the computer aids.



JAMA Network Open. 2021;4(3):e211276

Critical gaps in AI/ML deployment
Lack of diverse and representative data

Geographic Distribution of Data to Train AI Algorithms



Kaushal A, Altman R, Langlotz C. JAMA. 2020;324: 1212–1213.

AuntMinnie.com

Is radiology AI technology racist?

August 6, 2021 -- Artificial intelligence (AI) models can recognize a patient's racial identity on medical images, even though radiologists can't, ...

3 weeks ago

Judy W. Gichoya

"…report all results by relevant clinical and demographic group… "

<u>Need for representative dataset</u>

Area	Current State of the Art
Data needs for machine learning research	Few public image data sets are available, mostly small in size and lacking real-world variation.

Langlotz CP, Allen B, Erickson BJ, et al. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. Radiology. 2019;291: 781–791. https://doi.org/10.1148/radiol.2019190613

Area	Current State
Software use cases for AI	AI algorithms are being created based on use cases developed at single institutions working with single developers, limiting diversity and generalizability to widespread clinical practice.

Allen B Jr, Seltzer SE, Langlotz CP, et al. A Road Map for Translational Research on Artificial Intelligence in Medical Imaging: From the 2018 National Institutes of Health/RSNA/ACR/The Academy Workshop. J Am Coll Radiol. 2019. https://doi.org/10.1016/j.jacr.2019.04.014

Understanding Limitations of AI Development in Cancer Imaging can be appreciated through the Limitations of AI Development in COVID-19 Imaging AI

MIT Technology Review

Featured Topics Newsletters Events Podcasts

Artificial intelligence / Machine learning

Hundreds of AI tools have been built to catch covid. None of them helped.

Some have been used in hospitals, despite not being properly tested. But the pandemic could help make medical AI better.

by **Will Douglas Heaven** July 30, 2021

nature machine intelligence

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › nature machine intelligence › analyses › article

Analysis | Open Access | Published: 15 March 2021

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 & Carola-Bibiane Schönlieb

Nature Machine Intelligence **3**, 199–217 (2021) | Cite this article

RESEARCH

Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal

Laure Wynants,[1,2] Ben Van Calster,[2,3] Gary S Collins,[4,5] Richard D Riley,[6] Georg Heinze,[7] Ewoud Schuit,[8,9] Marc M J Bonten,[8,10] Darren L Dahly,[11,12] Johanna A Damen,[8,9] Thomas P A Debray,[8,9] Valentijn M T de Jong,[8,9] Maarten De Vos,[2,13] Paula Dhiman,[4,5] Maria C Haller,[7,14] Michael O Harhay,[15,16] Liesbet Henckaerts,[17,18] Pauline Heus,[8,9] Michael Kammer,[7,19] Nina Kreuzberger,[20] Anna Lohmann,[21] Kim Luijken,[21] Jie Ma,[5] Glen P Martin,[22] David J McLernon,[23] Constanza L Andaur Navarro,[8,9] Johannes B Reitsma,[8,9] Jamie C Sergeant,[24,25] Chunhu Shi,[26] Nicole Skoetz,[19] Luc J M Smits,[1] Kym I E Snell,[6] Matthew Sperrin,[27] René Spijker,[8,9,28] Ewout W Steyerberg,[3] Toshihiko Takada,[8] Ioanna Tzoulaki,[29,30] Sander M J van Kuijk,[31] Bas C T van Bussel,[1,32] Iwan C C van der Horst,[32] Florien S van Royen,[8] Jan Y Verbakel,[33,34] Christine Wallisch,[7,35,36] Jack Wilkinson,[22] Robert Wolff,[37] Lotty Hooft,[8,9] Karel G M Moons,[8,9] Maarten van Smeden[8]

OPINION

When the World Needed It Most, Artificial Intelligence Failed: How COVID-19 Poked Holes in AI

We should be celebrating how AI improved pandemic responses, but the rollout was messy and the published papers littered with unusable material

Emil Walleser Aug 4 · 4 min read ★

MIT Technology Review
Featured Topics Newsletters Events Podcasts

Artificial intelligence / Machine learning

Hundreds of AI tools have been built to catch covid. None of them helped.

Some have been used in hospitals, despite not being properly tested. But the pandemic could help make medical AI better.

by **Will Douglas Heaven** July 30, 2021

nature machine intelligence

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › nature machine intelligence › analyses › article

Analysis | Open Access | Published: 15 March 2021

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 & Carola-Bibiane Schönlieb

Nature Machine Intelligence **3**, 199–217 (2021) | Cite this article

53k Accesses | **1064** Altmetric | Metrics

Understanding Limitations of AI Development in Cancer Imaging can be appreciated through the Limitations of AI Development in COVID-19 Imaging AI

What went wrong?

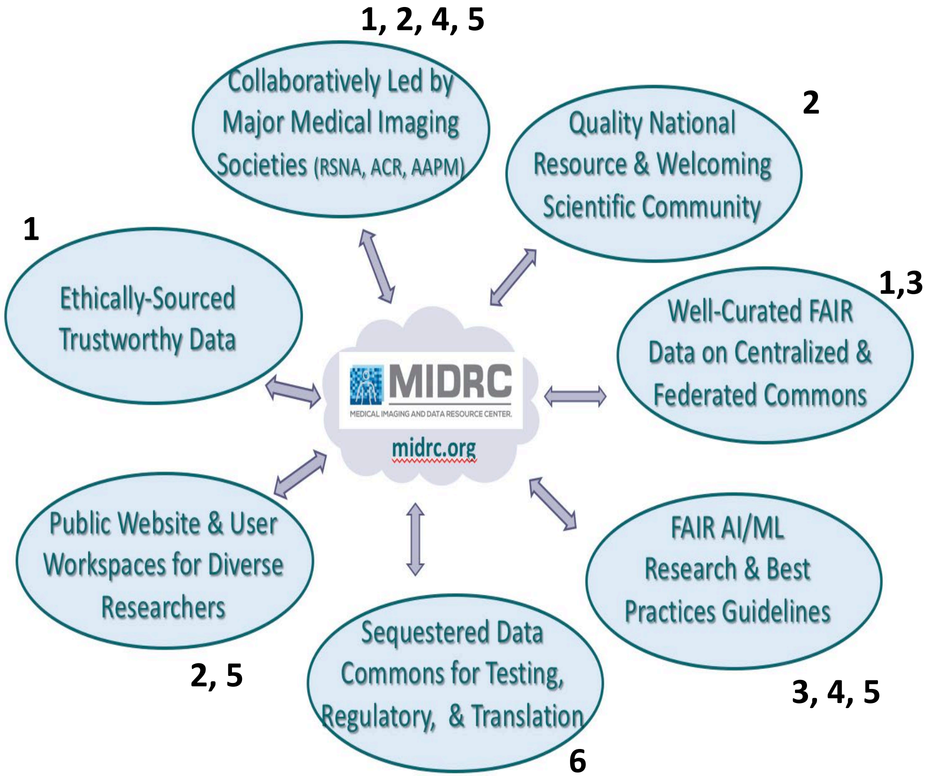
1) Poor quality of data, “Frankenstein data sets”
    - Mislabeled data
    - Multiple unknown sources
    - Duplicate data (resulting in leakage between training and testing)
    - No traceability, limited quality control
    - Lack of external validation

2) Lack of communications between AI/ML experts and Medical/ Biomedical experts; needed in this multidisciplinary field
    - Lack of valid ground truth

3) Bias
    - Collected often for a specific clinical question
    - Specific populations, lack of diversity
    - Single expert score, data sources correlated with ‘truth’, …

Suggest AI community embrace data commons (e.g., the Medical Imaging and Data Resource Center)

1. Focus on high quality data
   - **Trustworthy data**
2. Collaborative, community culture
   - **Bridge multiple expertise**
3. Promote **sharing & transparency**
   - **Data, models, limitations**
4. Create and promote **standards:**
   - **Data, Quality Control (QC)**
   - **Real world performance AI/ML**
5. Address bias:
   - **Representative data**
   - **Diversity of researchers**
   - **Lower barrier of access (FAIR, open)**
6. V**alue the "last mile"**
   - in "from bench to bedside"
   - Include post-market evaluations



midrc.org

Medical Imaging & AI in 10 years

1. Integrated multi-modality, multi-task AI
   - Currently, most algorithms are focused on one task, one cancer
   - Combinations of multiple human-engineered and deep learning AI algorithms
   - Increasing realization of the role of the end user in development and evaluation
2. AI as a means to improve access to healthcare (reduce health disparities)
   - AI when there are limited number of radiologists and other clinicians
   - Combine AI with inexpensive, portable imaging equipment
3. Multi-omics datasets for discovery and clinical “biomarkers” linked across patients over time
   - Ethically-sourced and trustworthy data and AI algorithms
   - Disease agnostic collection of datasets to Data Commons/Resource Centers
   - Open and sequestered data commons, such as MIDRC

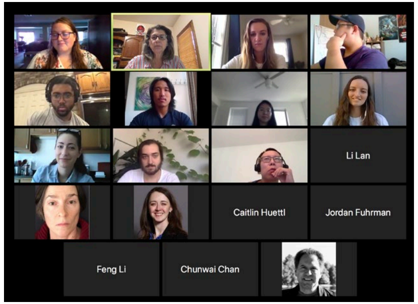
Recent & Current Graduate Students

Joel Wilkie, PhD
Martin King, PhD
Nick Gruszauskas, PhD
Yading Yuan, PhD
Robert Tomek, MS
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weiss, PhD
Chris Haddad, PhD
Natasha Antropova, PhD
Adam Sibley, PhD
Kayla Robinson, PhD
Jennie Crosby, PhD
Qiyuan (Isabelle) Hu
Jordan Fuhrman
Lindsay Douglas
Natalie Baughan

Thank you

Giger Lab

Pre-COVID



Research Lab

Karen Drukker, PhD
Hui Li, PhD
Heather Whitney, PhD
Yu Ji, MD
Chun Wai Chan, MS
Li Lan, MS
John Papaioannou, MS
Sasha (Alexandra) Edwards, MA
Madeleine Durkee, PhD
Summer medical students, undergraduates, and high school students

Collaborators

The MIDRC team

MIDRC
MEDICAL IMAGING AND DATA RESOURCE CENTER.

Extra Slides

Medical Imaging & AI in Precision Medicine

- **Artificial intelligence in medical imaging has been investigated for decades.** These investigations have included
- Understanding the changing role of AI in medical imaging in terms of:
  a. **Various medical decision-making tasks** -- from disease detection to diagnosis to therapeutic response & monitoring
  b. Methods of **how AI is used by the end-user** (e.g., the clinician) – from second reader to concurrent reader to autonomous reader
     i. Concerns of incorrect or off-label use, i.e. using a second reader AI system as a concurrent reader system
  c. **Human-engineered AI and deep learning AI**
  d. Need for **ethically-sourced imaging data & trustworthy AI**
  e. **Development of the AI algorithm** as well as the **evaluation of the radiologist when being aided by the AI** in order to assess translation to clinical practice.

- Image quality demands standardization
- Need to include segmentation to the list

- Also repositories vs. commons
  - Data from research effort with a focused question
    - Useful for Challenges but not representative of populations
  - Data from clinical trials with a strict acquisition protocols and also data release is embargoed
    - Data should be released by first publication or end of grant period at the latest
  - Data from large scale data from cooperative groups (like ACRIN etc)
  - While very useful, these are not representative data of the population
  - by collecting disease-agnostic data from diversity-directed collections one can slice and dice
    - Can then do discovery studies between imaging and nonimaging data – radiogenomics TCGA breast example
  - Avoid Frankenstein datasets
  - Coordination across case collections and clinical, demographics, and acquisition parameters data elements
    - If truth files of clinical data are not standardized, it could cause variations in how the truth is used and then different algorithms and performances.
  - Need non-imaging data standardized – somewhat like DICOM? LOINC?
  - Need sequestered datasets
  - Repeatability and reproducibility (variability)

What has changed with AI over the decades?

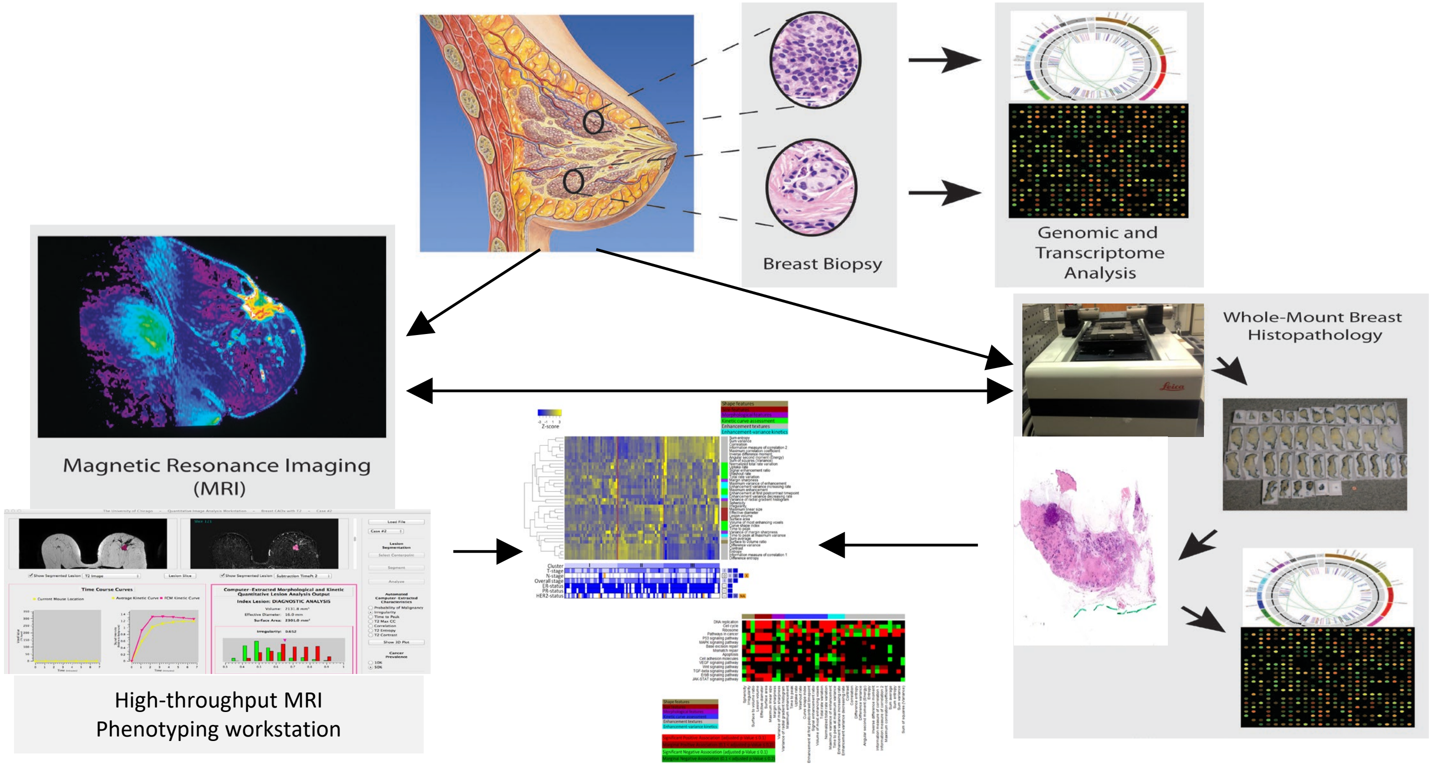
- Faster computers
- Larger datasets of images; although some datasets are limited or flawed
- More advanced algorithms including deep learning
- Realization of additional reasons & means to incorporate in clinical practice
- AI being developed for more clinical questions (modalities & disease sites)

<u>However</u>

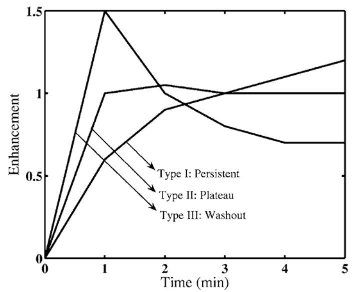
- Same clinical tasks of detection, diagnosis, response assessment
- Same concern for ”garbage in, garbage out”
- Same potential for misuse (i.e., off-label use)
- Same methods for statistical evaluations
- Same need for sufficient number of cases to span the distribution of disease and normal presentations
- Same need for imaging domain experts and computer domain experts

Spatially-correlated multi-omics analyses -- Advantages:

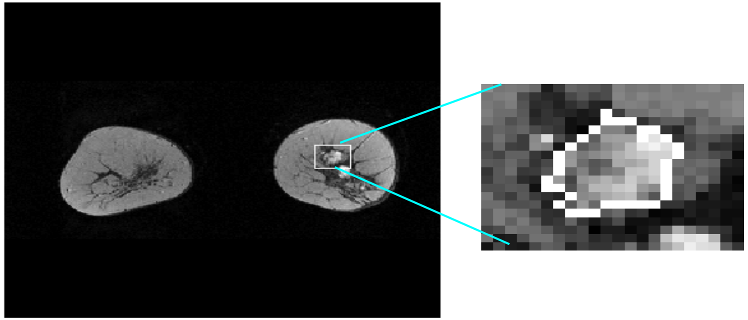
- Spatially registered MRI, histological, and genomics information – “virtual” digital biopsy
- Assesses tumor heterogeneity
- Accounts for tumor microenvironment.

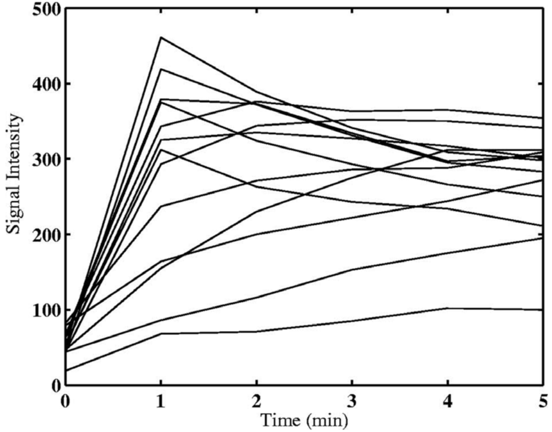


AI/Radiomics of intra-tumor heterogeneity e.g., Contrast Enhancement Heterogeneity in Breast DCE-MRI

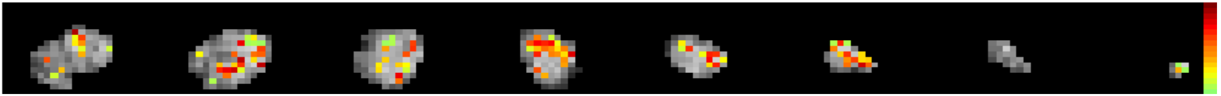


Heterogeneity of Tumors:





Regions of most enhancing voxels



Radiomics of texture giving a measure of the heterogeneity of contrast uptake.

Examine the heterogeneity of some features within the tumor to view regions (“habitats”), which may correspond to different biological states within the tumor.

Chen W, Giger ML, Bick U, Newstead G: Automatic identification and classification of characteristic kinetic curves of breast lesions on DCE-MRI. Medical Physics, 33: 2878-2887, 2006.

Human-engineered AI and deep learning AI

- Similar to radiology residency
  - Textbook learning of a limited number of specific examples
  - On the job training through reading cases and obtaining feedback from attending radiologists
- Advancements
  - Merging of human-engineered and deep learning AI algorithms
  - Multiple AI algorithms for a given task
  - Multiple image types (modalities) to mimic radiologists’ methods
  - Dependence on input images especially in a limited database situation

Radiology

ORIGINAL RESEARCH • SPECIAL REPORT

Ethics of Using and Sharing Clinical Imaging Data for Artificial Intelligence: A Proposed Framework

David B. Larson, MD, MBA • David C. Magnus, PhD • Matthew P. Lungren, MD, MPH • Nigam H. Shah, MBBS, PhD • Curtis P. Langlotz, MD, PhD

"After clinical data are used to provide care, the primary purpose for acquiring the data is fulfilled. At that point, clinical data should be treated as a form of public good, to be used for the benefit of future patients."

https://pubs.rsna.org/doi/pdf/10.1148/radiol.2020192536

Database vs. Repository vs. Resource Center

“Rough definitions”: Database vs. Repository vs. Resource Center

- Database – collection of data
- Repository – data stored and managed
- Resource center – repository with additional aspects including browse/search, analysis, metrology, evaluation, and user interfaces

Large Scale Medical Imaging Studies

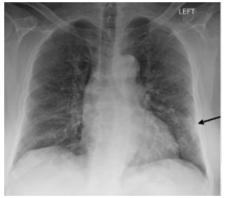
- Large is relative –

- Relative to the ML/AI task
- Relative to the prevalence of the disease in question
- Relative to the subtlety of the disease in question
- Relative to the difficulty of the clinical task
- Relative to sub-populations; bias and diversity aspects
- Relative to the type of ML/AI training, tuning, and testing
  - Merging of human-engineered features
  - Transfer learning (feature extraction or fine tuning)
  - Deep learning from scratch
- “Think like a human”

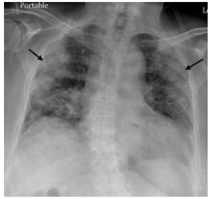
Use Case: Thoracic imaging in the COVID-19 Pandemic

While thoracic imaging, including **chest radiography (CXR) and computed tomography (CT)**, are being re-examined for their role in patient management, the limitations for improved interpretation are partially due to the qualitative interpretation of the images, and thus we aim to develop artificial intelligence (AI) methods to aid in the interrogation of medical images from COVID-19 patients, **eventually including cardiac, brain, and other images**.

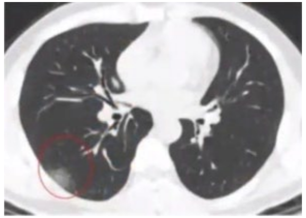
Bilateral lower lobe consolidations



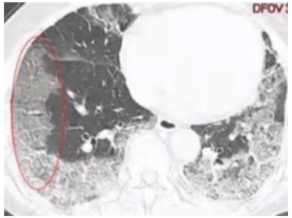
Patchy peripheral ground glass opacities



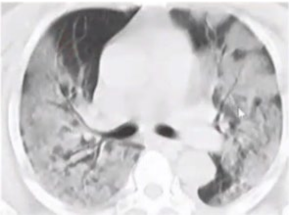
Peripheral air space opacities



Early Stage



Progressive Stage



Severe Stage

Rapid Response to COVID-19 Pandemic

MIDRC
MEDICAL IMAGING AND DATA RESOURCE CENTER.
Established August 21, 2020

MIDRC.org

University of Chicago NIBIB Contract PI: Maryellen Giger

American Association of Physicists in Medicine (AAPM) PIs:

- Maryellen Giger (University of Chicago & AAPM Data Science Committee Chair)
- Paul Kinahan (University of Washington & AAPM Research Committee Chair)

Radiological Society of North America (RSNA) PIs:

- Curtis Langlotz (Stanford University & RSNA Board Liaison for IT & Annual Meeting)
- Adam Flanders (Thomas Jefferson University & Member RSNA CDE Committee)

American College of Radiology (ACR) PIs:

- Etta Pisano (ACR Chief Research Officer & Harvard University)
- Michael Tilkin (ACR Chief Information Officer)

Gen3 PI: Robert Grossman

NIH National Institute of Biomedical Imaging and Bioengineering

THE UNIVERSITY OF CHICAGO

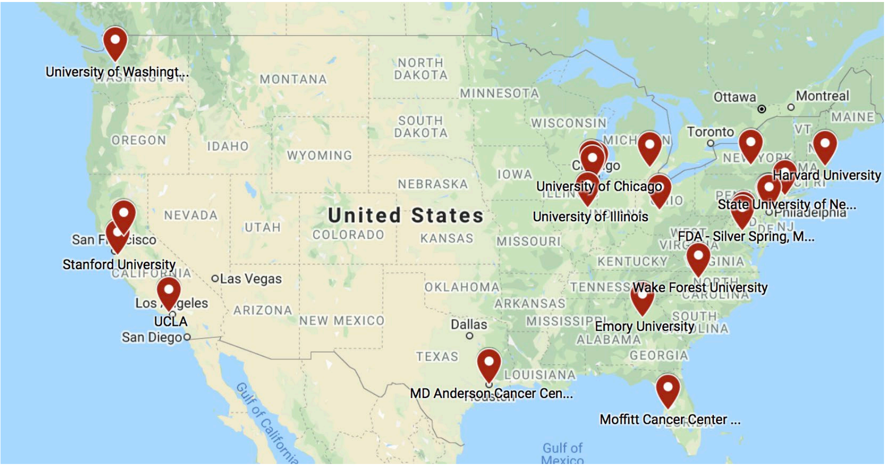
AMERICAN ASSOCIATION of PHYSICISTS IN MEDICINE

ACR® AMERICAN COLLEGE OF RADIOLOGY

RSNA® Radiological Society of North America

MIDRC

MEDICAL IMAGING AND DATA RESOURCE CENTER.



- MIDRC -- radiologists & medical physicists/imaging scientists across the nation
  - **Collaboration of 23 institutions from academia, community practices, FDA**
  - Expert collaboration with community engagement
- See website for listing of all investigators
  - https://www.midrc.org
- High-quality and diverse data commons enabling researchers to address topics no single archive could yield independently (including images/acquisition, clinical, demographic data)

MIDRC

MEDICAL IMAGING AND DATA RESOURCE CENTER.

midrc.org

Two Data Intake Portals

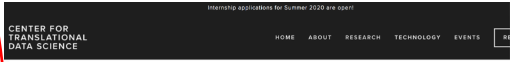
RSNA
COVID-19 Database
RICORD

ACR
COVID-19
Imaging Research
REGISTRY™

Image quality assurance and unbiased AI algorithm evaluation methods (AAPM)

One Output User Portal

University of Chicago



Technology
DATA COMMONS
GEN3 SOFTWARE
GUIDING PRINCIPLES

GEN3 DATA COMMONS

data.midrc.org for searching & downloading data

- Open Commons: Diverse commons to be accessed by AI researchers.
- Sequestered Commons: To expedite translation of AI through regulatory & to clinical care via sequestered datasets and task-based distribution.

MIDRC: Technology Development Projects

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

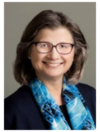
1. Creating an open discovery platform for COVID-19 imaging and associated data (**led by RSNA**).
2. Creating a real-world testing and implementation platform with direct real-time connections to health care delivery organizations (**led by ACR**).
3. Developing and implementing quality assurance and 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**).

Three MIDRC Data Science Subcommittees

- DSIT - Data Standards and Information Technology Subcommittee
  - led by RSNA
- DPP - Data Policy and Procedures Subcommittee
  - led by ACR
- DQH - Data Quality and Harmonization Subcommittee
  - led by AAPM

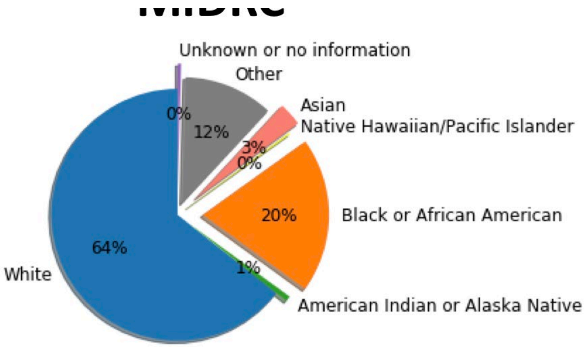
Diversity of the data is an essential component in the developing and testing of unbiased AI
-- MIDRC PIs aims in a pending NIH grant

1. Quantitative assessment of the diversity in imaging data within MIDRC, and establishment of fairness metrics and best practices to mitigate bias in AI development.
2. Development of algorithmic interventions to detect and reduce bias in designing and independently testing AI algorithms for medical imaging.
3. Investigation of de-biased AI algorithms at scale to understand root causes of health disparities.
4. Investigation of the clinical impact of our proposed AI interventions in clinical decision making with medical imaging.



Maryellen Giger — UChicago
Judy Gichoya — Emory
Jayashree Kalapathy-Cramer — MGH/Harvard
Sanmi Koyejo. — U Illinois
Berkman Sahiner — FDA

MIDRC data as of May 2021



MIDRC Data Dashboard **midrc.org**

MIDRC
MEDICAL IMAGING AND DATA RESOURCE CENTER.

Total ingested into MIDRC

# of Imaging Studies

41,071

Undergoing MIDRC Data Quality and Harmonization

# of Imaging Studies

38,927

Released by MIDRC

# of Imaging Studies

2,144

Including international datasets

Quality checked
Diversity assessed
Clinical Task AI

Goal of 60,000 curated imaging studies to be released by MIDRC by Sept 2021

Collaborative Research Projects – Investigators through the various Data Science Committees at ACR, RSNA, & AAPM

Project	Title
1	Natural Language Processing of Radiology Reports for COVID-19
2	Machine Intelligence Algorithms from Multi-Modal, Multi-institutional COVID-19 Data
3	Image Labeling and Annotation by a Crowd of Experts for COVID-19
4	Efficient Training and Explainability of Machine Learning Methods from Multi-Institutional Data
5	COVID Pneumonia Machine Learning Algorithm Validation and Visualization
6	Safe Public Training Dataset for COVID-19 Machine Learning Algorithms
7	Leveraging Registry Data to Conduct Virtual Clinical Trials
8	Prediction of COVID Pneumonia Outcome using Radiomic Feature Analysis
9	Radiomics & Machine Intelligence of COVID-19 for detection and diagnosis on chest radiographs and thoracic CTs
10	Visualization & Explainability of Machine Intelligence of COVID-19 for prognosis and monitoring therapy
11	Investigation of image-based biomarkers for radiogenomics of COVID-19
12	Determining COVID-19 image data quality, provenance, and harmonization

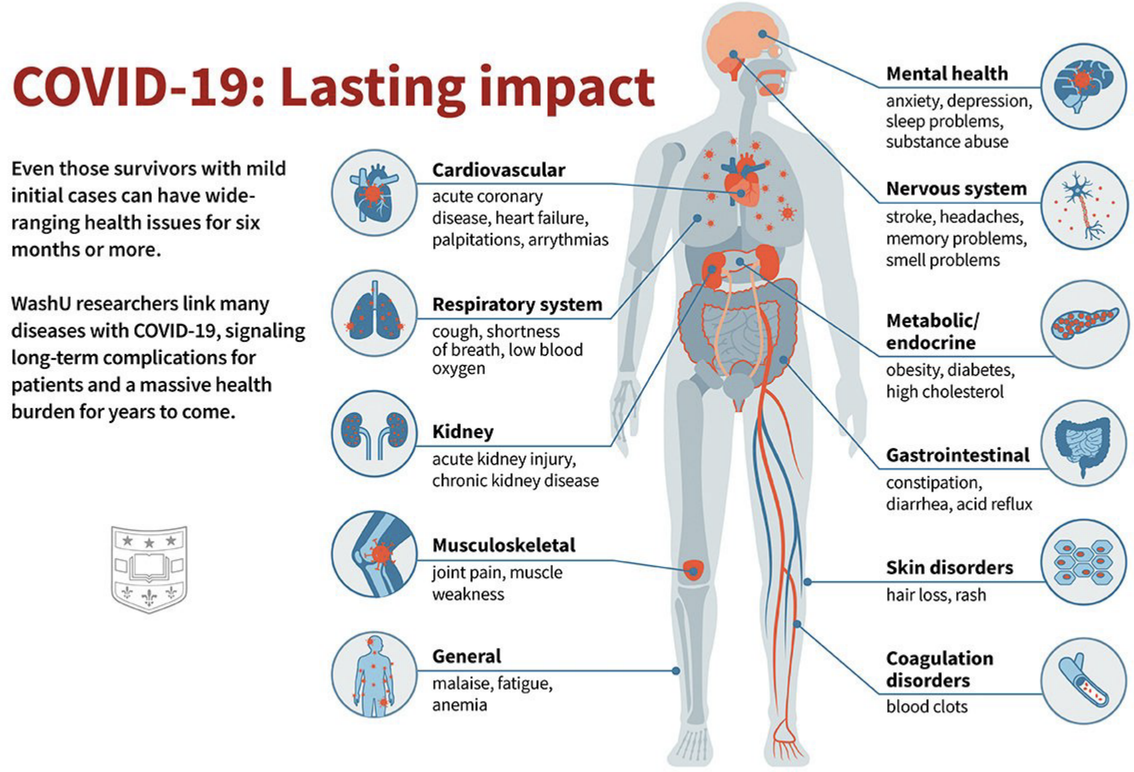
Trans-MIDRC scientific workgroups

- Grand Challenges Work Group
  - Created to coordinate effort on all aspects of challenges
  - Potential to merge top performing algorithms to benefit the common good
- Bias and Diversity Work Group
  - Goal of assessing and mitigating bias in data and ML
  - Diversity in MIDRC

What is next for MIDRC?

MIDRC
MEDICAL IMAGING AND DATA RESOURCE CENTER.

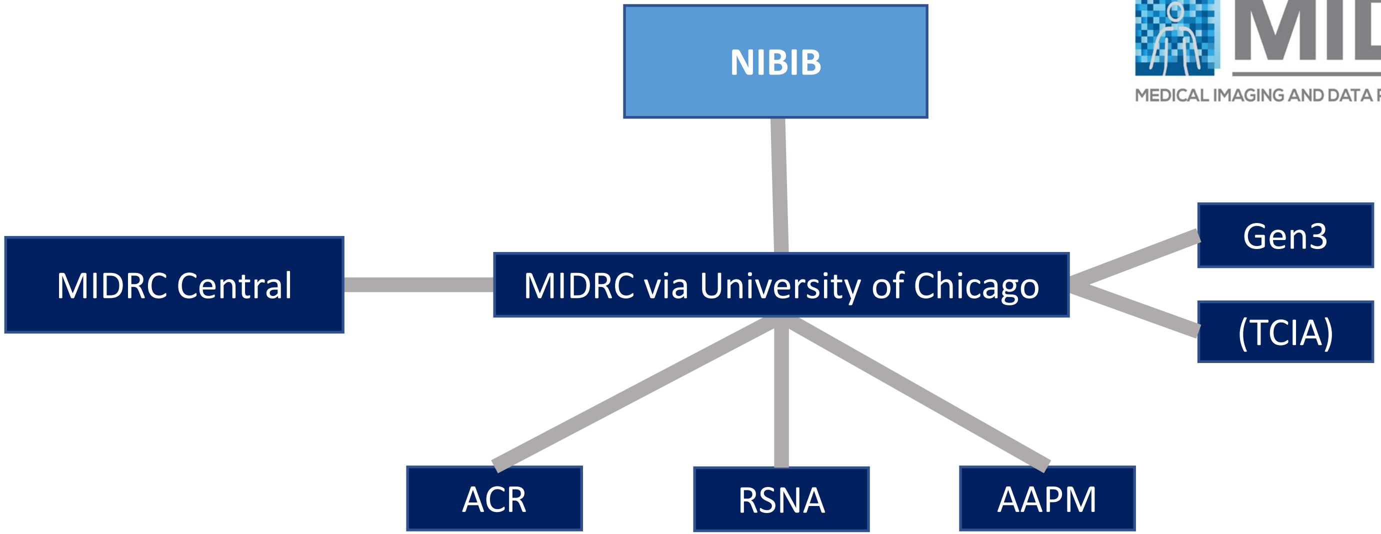
- **Beyond chest radiographs and thoracic CTs for COVID-19**
  - Include images of the (
    (**brain**)
  - Include images beyond
    images to **monitor pos**
  - Collaborate with the n
    across clinical data, im
- **Beyond COVID-19**
  - MIDRC, with its develo
    tools, will be ready for
- Thus, MIDRC will be an **in**
  - Currently funded for tv
  - Require additional funds to continue with other diseases.



Field in 10 years (exciting directions & challenges)

4. Comprehensive data registries (future of MIDRC) to support de-biased AI development, testing, & translation for public good
   - Resource center – repository with additional aspects including, e.g., analysis, harmonization, metrology, sequestered data, evaluation, computational enclaves, centralized & federated, and user interfaces
   - "DICOM-image-type" organization for all other –omics (pathomics, EHR, genomics)
   - Collaboration and interconnectivity with other data commons
   - Collaboration with regulatory/FDA to expedite translation & post-market evaluation
5. Open shared data resources and a willingness to give data for the public good
   - A patient has already benefited through medical care.
   - A hospital/medical center has already benefitted through reimbursement.
   - Now let the public benefit with the MIDRC second usage of the images.
   - Need to change the culture of medical imaging.

MIDRC
MEDICAL IMAGING AND DATA RESOURCE CENTER.



Two Major Scientific Components

<u>Creation of Open Discovery Data Repository</u>: 5 Technology Development Projects
along with three data science subcommittees and advisory committees

<u>Machine Intelligence Computational Capabilities</u>: 12 Collaborative Research Projects
along with multiple trans-MIDRC scientific workgroups

MIDRC Data Dashboard

MIDRC
MEDICAL IMAGING AND DATA RESOURCE CENTER.

Total ingested into MIDRC

# of Imaging Studies

41,071

Undergoing MIDRC Data Quality and Harmonization

# of Imaging Studies

38,927

Released by MIDRC

# of Imaging Studies

2,144

Quality checked
Diversity assessed
Clinical Task AI

Goal of 60,000 curated imaging studies to be released by MIDRC by Sept 2021

Database vs. Repository vs. Resource Center

"Rough definitions":

- Database – collection of data
- Repository – data stored and managed
- Resource center – repository with additional aspects including browse/search, analysis, metrology, evaluation, and user interfaces

Key Bottleneck to Successful and Meaningful Machine Learning Algorithms: Lack of Diverse Data

Area	Current State of the Art
Data needs for machine learning research	Few public image data sets are available, mostly small in size and lacking real-world variation.

Area	Current State
Software use cases for AI	AI algorithms are being created based on use cases developed at single institutions working with single developers, limiting diversity and generalizability to widespread clinical practice.

Langlotz CP, Allen B, Erickson BJ, et al. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. Radiology. 2019;291: 781–791. https://doi.org/10.1148/radiol.2019190613

Allen B Jr, Seltzer SE, Langlotz CP, et al. A Road Map for Translational Research on Artificial Intelligence in Medical Imaging: From the 2018 National Institutes of Health/RSNA/ACR/The Academy Workshop. J Am Coll Radiol. 2019. https://doi.org/10.1016/j.jacr.2019.04.014

MIDRC
MEDICAL IMAGING AND DATA RESOURCE CENTER.

Radiology

ORIGINAL RESEARCH • SPECIAL REPORT

Ethics of Using and Sharing Clinical Imaging Data for Artificial Intelligence: A Proposed Framework

David B. Larson, MD, MBA • David C. Magnus, PhD • Matthew P. Lungren, MD, MPH • Nigam H. Shah, MBBS, PhD • Curtis P. Langlotz, MD, PhD

> "After clinical data are used to provide care, the primary purpose for acquiring the data is fulfilled. At that point, clinical data should be treated as a form of public good. All who interact with or control the data have an obligation to ensure that the data are used for the benefit of future patients and of society."

https://pubs.rsna.org/doi/pdf/10.1148/radiol.2020192536

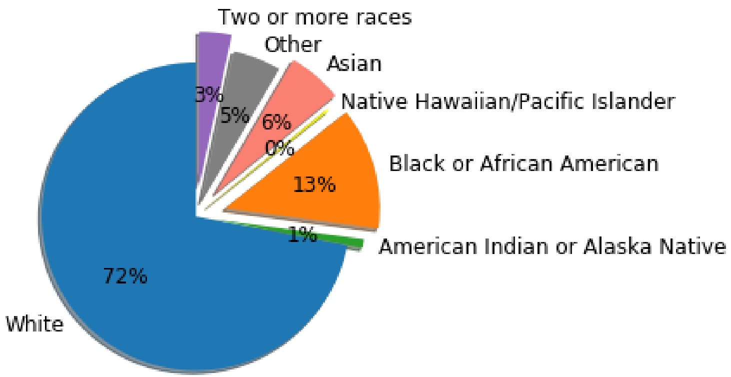
Diversity of the data is an essential component in the developing and testing of unbiased AI

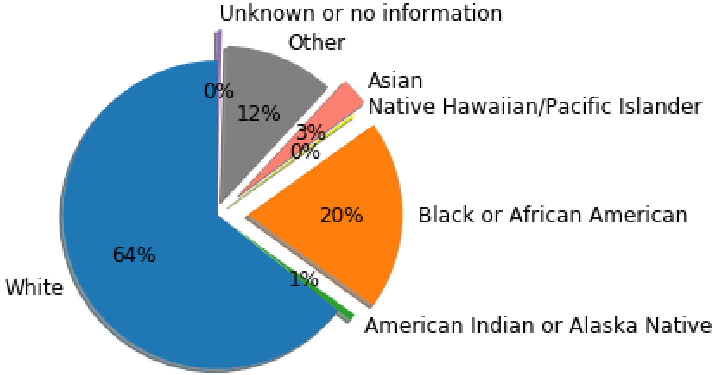
MEDICAL IMAGING AND DATA RESOURCE CENTER

US Census

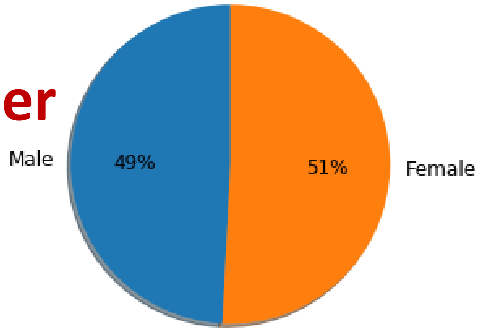
MIDRC

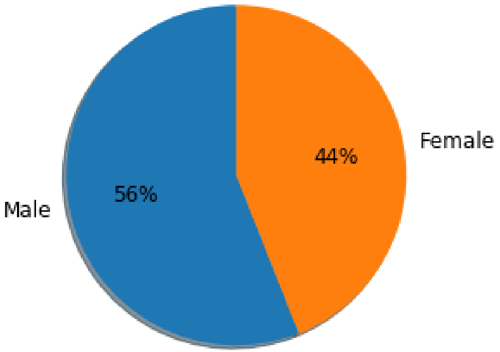
Race





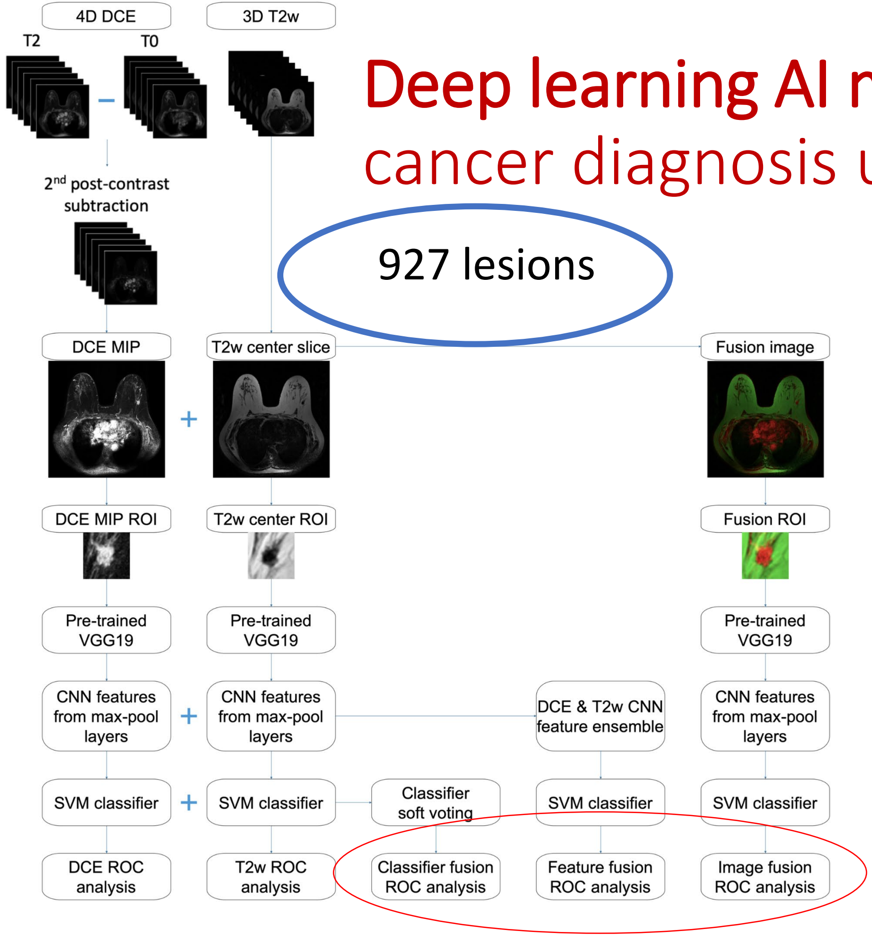
Gender



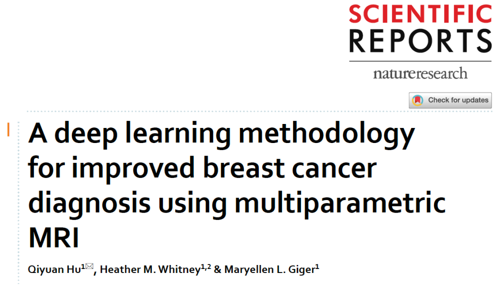


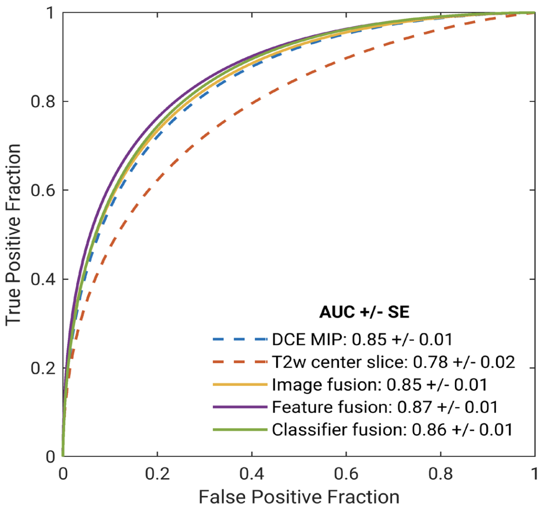
As of incoming data
May 31, 2021

Deep learning AI methodology for improved breast cancer diagnosis using multiparametric MRI



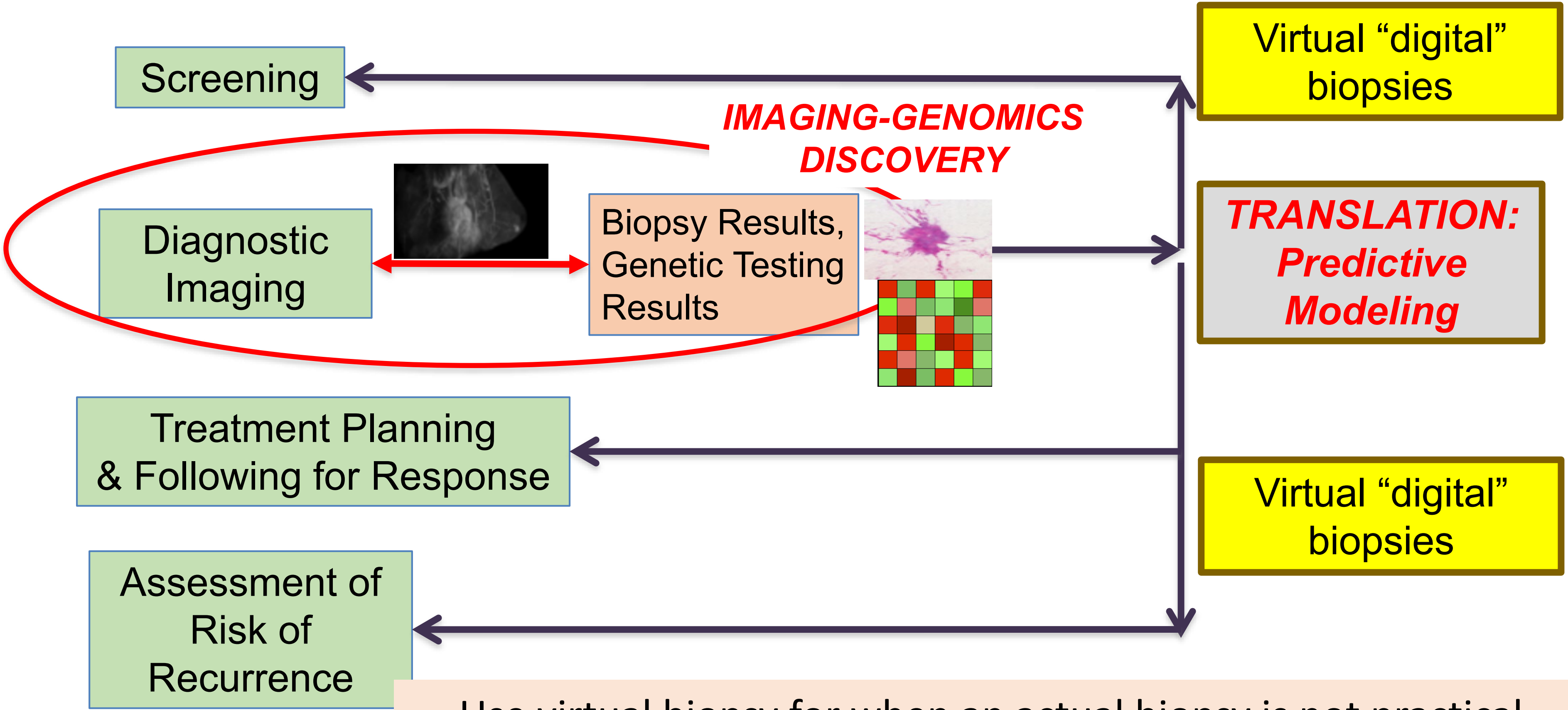
927 lesions





Hu Q, Whitney HM, Giger ML. A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI. Sci Rep. 2020 Jun 29;10(1):10536. doi: 10.1038/s41598-020-67441-4.

Discovery and Predictive Modeling for Personalized Patient Care



Use virtual biopsy for when an actual biopsy is not practical

Giger Lab; U Chicago