Bringing AI from Hype to Reality for Routine Clinical Practice: Addressing the Gaps and Opportunities for NCI

Eliot Siegel, M.D., FSIIM, FACR

Professor and Vice Chairman University of Maryland School of Medicine Department of Diagnostic Radiology and Nuclear Medicine

Chief Imaging VA Maryland Healthcare System

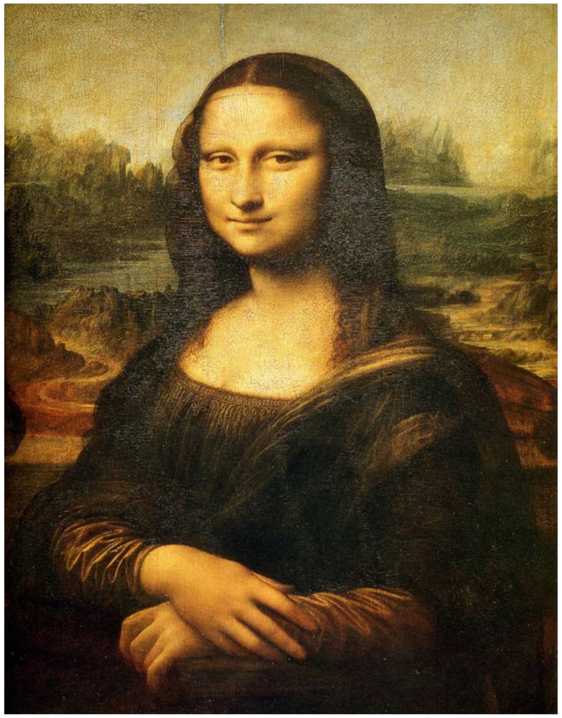
2004 Kick-off for National EMR Initiative (Meaningful Use)
George W. Bush Visit to Baltimore VA Medical Center

Introduction

- This coming June marks the 30th anniversary of operation of the world's first filmless radiology department at the Baltimore VA
- In addition to the goal of any image any time and anywhere when it was needed for clinical care
- We also wanted to move to digital to take advantage of advances in computer aided detection and diagnosis and quantification which was a rapidly emerging field in the 1990s
- When asked about the predictions about the future that turned out to be the farthest off, I would have been quite surprised to know almost 30 years later that these AI algorithms were still not in routine use and had almost no impact on patient cancer care

Introduction

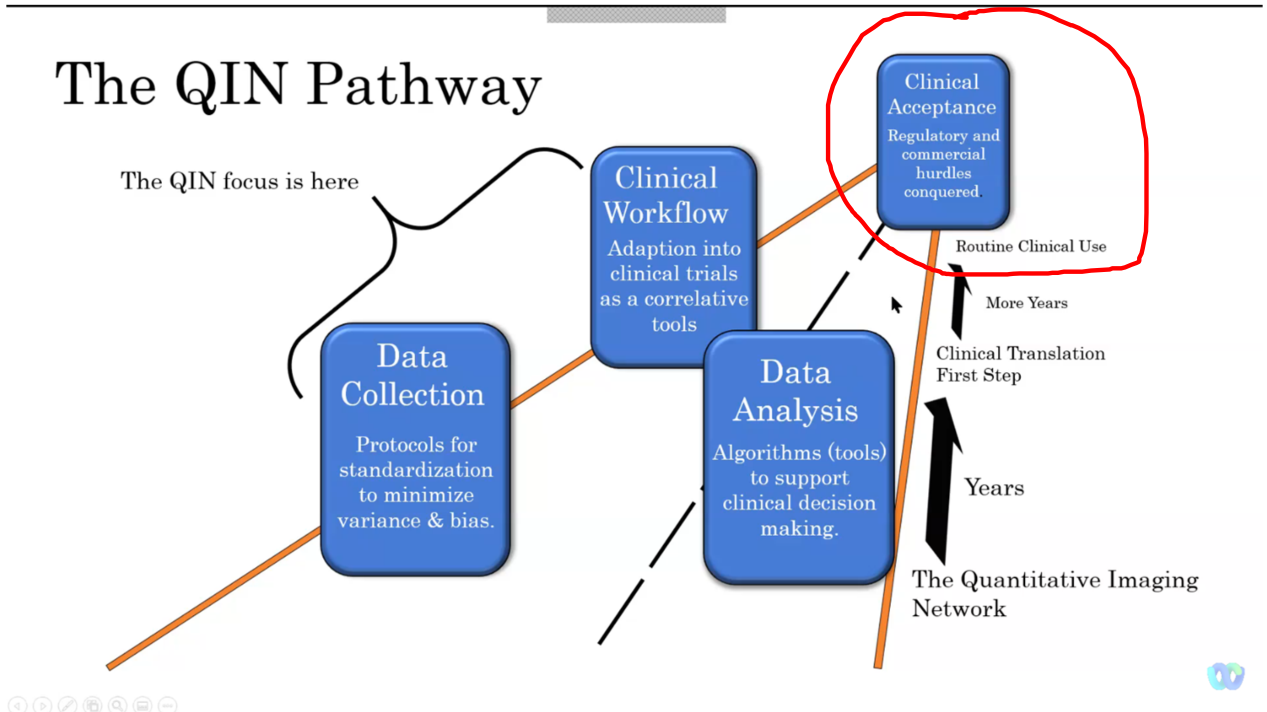
- When I was first asked by Dr. Daniel Sullivan of the Cancer Imaging Program to serve as lead for the caBIG imaging workspace, the question from the caBIG leadership was why would imaging even be a part of the program since (unlike genomics) imaging was just qualitative "artwork" without a real quantitative component?
- Many on this virtual presentation were involved in the many creative and innovative projects including TCIA/NCIA, ePAD, AIM annotation and image mark-up schema and many others that have endured today and influenced imaging at NCI
  - A major part of my goal for the original cancer imaging archive was that it could serve as a training set for computer aided detection and what subsequently became referred to as "AI"



Call to Action

- RSNA and ACR have done great work identifying use cases and promoting AI algorithm development with competitions such as offered by Kaggle, but as was the case for TCIA/NCIA, and the many advances in imaging informatics in the era of caBIG, NCI may need (and be in the best position) to play a major role in moving things forward again
- This presentation is intended to serve as a "call to action" to NCI to consider providing even greater leadership in AI translation and implementation in routine cancer care

Dr. Robert Nordstrom's QIN Slide



Tools Ready for Clinical Validation and Utility From QIN

- 3D Slicer
- ePAD
- PyRadiomics
- Automated PET Phantom Analysis & Reporting Tool (APPART)
- PET Tumor Segmentation
- Quantitative DWI QC
- Aegis SER
- AutoPERCIST
- Functional Analysis Platform of imFIAT
- IB Clinic
- MiViewer
- Solid Tumor Segmentation
- Spectroscopic MRI Clinical Interface

Quantitative Imaging and Clinical Trials

- Quantitative Imaging (in clinical trials): the extraction of measurable information from medical images to assess the status or change in a status of normal and disease
  - Sits at the crossroads of imaging, analytics, and informatics to provide quantitative tools for clinical decision support
  - May offer valuable anatomic, physiologic, metabolic and molecular information, provide important insights into disease location and extent, and reduce the need for multiple biopsies

I want that but don't have it for my everyday routine clinical practice in oncology!

Vast Majority of Cancer Patients are Treated outside of Clinical Trials

1. Percentage of patients enrolled in clinical trials:
   a. 18.9% of patients at NCI-designated centers participated in treatment trials, compared with 3.9% of patients from community cancer programs, 4.7% from integrated network cancer programs, and 5.0% from academic comprehensive cancer programs.
      i. Unger JM, Fleury M. Nationally representative estimates of the participation of cancer patients in clinical research studies according to the commission on cancer. *J Clin Oncol*. 2021 (suppl 28; abstr 74). doi:10.1200/JCO.2020.39.28_suppl.74
2. So the vast majority of cancer patients are not enrolled in a clinical trial and these algorithms need to be delivered to those 95% of cancer patients

The Great Divide Challenge – Validation of QIN tools in Clinical Trials

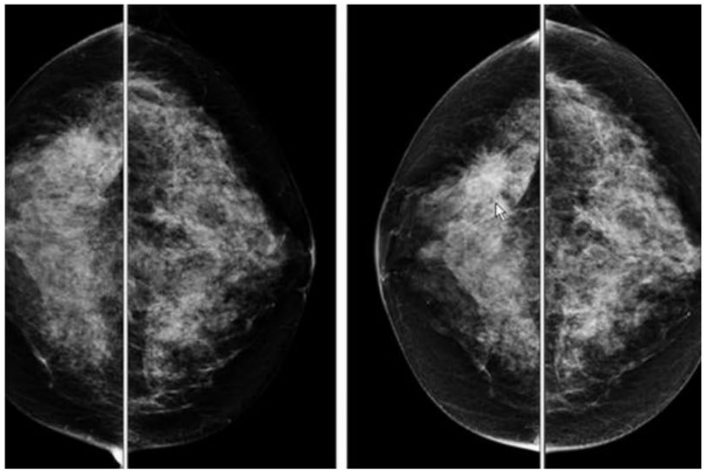
J Eary 2018

In the News Last Week:
Why We Need Greater NCI Involvement in AI for Cancer Care

Innovate Healthcare
HealthImaging
INSIGHTS IN IMAGING & INFORMATICS

More than 6,000 mammograms reviewed after radiology group misses dozens of cancers

Hannah Murphy | *October 26, 2022* | *Breast Imaging*



A radiology group in Arizona is under fire after allegedly missing dozens of breast malignancies, some of which were "screaming cancer," according to a new NBC News investigation.

"I literally couldn't believe what I was seeing," breast surgeon, Dr. Beth Dupree, who consulted with numerous patients impacted by the missed diagnoses, told the network. "These were misses that were not subtle."

Combined AI and Radiologist Assessment Mammography

- Article published in JAMA Network Open:
  - An ensemble of AI algorithms combined with radiologist assessment in a single-reader screening environment improved overall accuracy

Original Investigation | Imaging

Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms

Thomas Schaffter, PhD; Diana S. M. Buist, PhD, MPH; Christoph I. Lee, MD, MS; Yaroslav Nikulin, MS; Dezső Ribli, MSc; Yuanfang Guan, PhD; William Lotter, PhD; Zequn Jie, PhD; Hao Du, BEng; Sijia Wang, MSc; Jiashi Feng, PhD; Mengling Feng, PhD; Hyo-Eun Kim, PhD; Francisco Albiol, PhD; Alberto Albiol, PhD; Stephen Morrell, B Bus Sc, MiF, M Res; Zbigniew Wojna, MSI; Mehmet Eren Ahsen, PhD; Umar Asif, PhD; Antonio Jimeno Yepes, PhD; Shivanthan Yohanandan, PhD; Simona Rabinovici-Cohen, MSc; Darvin Yi, MSc; Bruce Hoff, PhD; Thomas Yu, BS; Elias Chaibub Neto, PhD; Daniel L. Rubin, MD, MS; Peter Lindholm, MD, PhD; Laurie R. Margolies, MD; Russell Bailey McBride, PhD, MPH; Joseph H. Rothstein, MSc; Weiva Sieh, MD, PhD; Rami Ben-Ari, PhD; Stefan Harrer, PhD; Andrew Trister, MD, PhD; Stephen Friend, MD, PhD; Thea Norman, PhD; Berkman Sahiner, PhD; Fredrik Strand, MD, PhD; Justin Guinney, PhD; Gustavo Stolovitzky, PhD; and the DM DREAM Consortium

Abstract

IMPORTANCE Mammography screening currently relies on subjective human interpretation. Artificial intelligence (AI) advances could be used to increase mammography screening accuracy by reducing missed cancers and false positives.

OBJECTIVE To evaluate whether AI can overcome human mammography interpretation limitations with a rigorous, unbiased evaluation of machine learning algorithms.

DESIGN, SETTING, AND PARTICIPANTS In this diagnostic accuracy study conducted between September 2016 and November 2017, an international, crowdsourced challenge was hosted to foster AI algorithm development focused on interpreting screening mammography. More than 1100 participants comprising 126 teams from 44 countries participated. Analysis began November 18, 2016.

MAIN OUTCOMES AND MEASUREMENTS Algorithms used images alone (challenge 1) or combined images, previous examinations (if available), and clinical and demographic risk factor data (challenge 2) and output a score that translated to cancer yes/no within 12 months. Algorithm accuracy for breast cancer detection was evaluated using area under the curve and algorithm specificity compared with radiologists' specificity with radiologists' sensitivity set at 85.9% (United States) and 83.9% (Sweden). An ensemble method aggregating top-performing AI algorithms and radiologists' recall assessment was developed and evaluated.

RESULTS Overall, 144 231 screening mammograms from 85 580 US women (952 cancer positive ≤12 months from screening) were used for algorithm training and validation. A second independent validation cohort included 166 578 examinations from 68 008 Swedish women (780 cancer positive). The top-performing algorithm achieved an area under the curve of 0.858 (United States) and 0.903 (Sweden) and 66.2% (United States) and 81.2% (Sweden) specificity at the radiologists' sensitivity, lower than community-practice radiologists' specificity of 90.5% (United States) and 98.5% (Sweden). Combining top-performing algorithms and US radiologist assessments resulted in a higher area under the curve of 0.942 and achieved a significantly improved specificity (92.0%) at the same sensitivity.

CONCLUSIONS AND RELEVANCE While no single AI algorithm outperformed radiologists, an ensemble of AI algorithms combined with radiologist assessment in a single-reader screening

Key Points

Question How do deep learning algorithms perform compared with radiologists in screening mammography interpretation?

Findings In this diagnostic accuracy study using 144 231 screening mammograms from 85 580 women from the United States and 166 578 screening mammograms from 68 008 women from Sweden, no single artificial intelligence algorithm outperformed US community radiologist benchmarks; including clinical data and prior mammograms did not improve artificial intelligence performance. However, combining best-performing artificial intelligence algorithms with single-radiologist assessment demonstrated increased specificity.

Meaning Integrating artificial intelligence to mammography interpretation in single-radiologist settings could yield significant performance improvements, with the potential to reduce health care system expenditures and address resource scarcity experienced in population-based screening programs.

+ Invited Commentary

+ Supplemental content

Author affiliations and article information are

What are the Current Challenges to Bring AI to Cancer Patients?

- Difficult to know which of the many commercial AI algorithms perform well in generalized practice, and also how do those developed using NIH/NCI funds perform
  - Unfortunately, FDA clearance does not imply clinical efficacy
- Clinical acquisition protocols are much more variable than those used for clinical trials and AI software may be brittle and vulnerable to these differences
  - Image quality control is a related issue
- How closely does a given AI algorithm conform to the patient population it is being applied to?
- Shouldn't the AI algorithm to learn/improve over time?
- Incorporation of personalized patient data and history into algorithm performance
- Lack of integration of AI into the workflow
  - AI as standalone application only works when there is a single algorithm used and even then, workflow integration in PACS or dictation software or advanced visualization software is key

Current Challenges to Bring AI to Cancer Patients?

- Medicolegal issues including storing results and high dimensional multi-parametric data
- Delivery of AI algorithms to oncologists and other non-radiologist providers
- Lack of a standard platform or a standard for platforms
- Security and privacy issues
- Reimbursement challenges
- Discovering algorithms and datasets indexed by NCI or NLM similar to PubMed for publications
- Imprecise methods for screening patients such as simply applying the NLST criteria and then determining how rapidly and intensively to follow up

Challenge: What Help is There for Radiologists in Clinical Practice to Choose the Right AI Algorithm?

- FDA clearance is more focused on process of developing AI algorithms and ensuring that claims made are reasonable based on development process
  - Does not presume to judge clinical effectiveness or accuracy of the AI algorithm
- NCI AI algorithms are much more rigorously tested in multiple facilities and have greater validation than the majority of FDA approved tools
  - How to get FDA clearance of NCI tools?
  - Is there the possibility of making NCI tools/algorithms available for clinical use whether or not they have been FDA cleared? There is no requirement that radiologists/oncologists must be limited to only FDA cleared tools except for reimbursement purposes but there is very limited reimbursement now for AI

Challenge: Uniform Acquisition Protocols
QIBA: Quantitative Imaging Biomarkers Alliance

- QIBA is an initiative to advance quantitative imaging and the use of imaging biomarkers in clinical trials and clinical practice by engaging researchers, healthcare professionals and industry
- This involves:
  - Collaborating to identify needs, barriers, and solutions to develop and test consistent, reliable, valid, and achievable quantitative imaging results across imaging platforms, clinical sites, and time.
  - Accelerating the development and adoption of hardware and software standards needed to achieve accurate and reproducible quantitative results from imaging methods.

Challenge to Acquire PET/CT Images in Standardized Fashion

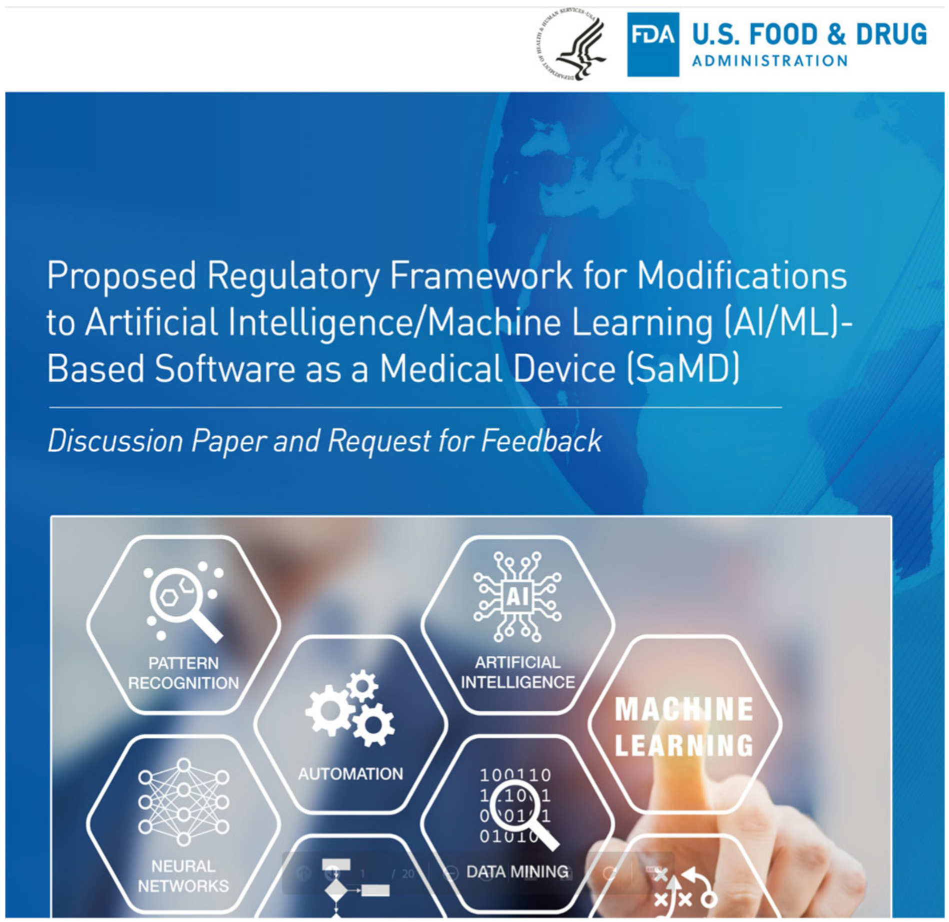
- Despite efforts such as QIBA which unlike UPICT is very clinically focused although there has been very limited adoption
- Efforts working RSNA to motivate clinical interest in PET/CT SUV standardization through an accreditation process has met without high enthusiasm by outpatient facilities
  - Overall, motivation for clinical trial conformity is that it is mandated by trial and ability to participate in trial and get paid
  - There is no such similar mandate for clinical practice ironically which has led to major discrepancies in measurements, SUV, distance, MRI perfusion etc. in actual patient care

Defining Image Quality in Clinical Practice?

- Is there a quantitative metric of image quality?
  - UCSF project looked at over 800 CT scans with over 120 CT readers with Duke and UMD looking at physics aspects and my task was to create a machine learning algorithm to predict radiologists rating of studies where that rating varied considerably with project originally designed to look at image quality trade-offs with radiation dose
- How can we assess it in clinical practice when there is no central reading core lab etc.?
- There are commercial algorithms for mammography, for example that assess factors such as patient positioning and other quality metrics
- Acquisition protocols such as Iterative Reconstruction in CT or a high spatial frequency reconstruction kernel can be destructive to texture information that might be used by an AI algorithm

Challenge of Generalizing an AI Algorithm to Specific Population and Ability of the Algorithm to Learn Over time

- Clinical practice may be fine tuned for a particular location and particular patient population unlike the typical case with a clinical trial which lasts limited period of time with “experts” selected to determine “truth”
- FDA has a white paper looking at the intriguing possibility of software that changes with feedback from users and learns as a resident or fellow would learn over time
  - I am not aware of any vendors that have taken advantage of this opportunity at this point
- Project at U of Maryland where we applied corrective lens for “astigmatism” of differences from NIH algorithm to local BMD scores at U of MD
  - Substantial improvement in performance of BMD prediction algorithm developed at NIH



FDA Role Different for Clinical Practice Algorithms Than Clinical Trials for Research

- https://aicentral.acrdsi.org/ ACR Data Science Institute website lists 190 AI algorithms cleared by FDA but what about post market surveillance?
- FDA has strong interest in post-market surveillance but there is not currently a mechanism for agreement or disagreement by end users with the algorithm to be communicated back to vendors or FDA so FDA does not track clinical efficacy of cleared software but would like to as they do with pharmaceuticals
- Clinical practice heterogeneous and often different than where software was developed and tested, this may be to a lesser degree with clinical trials
- Academic level expert interpretation in clinical trials vs. less subspecialty high level expertise in clinical practice
- Current project is working with FDA to create a standardized method and data elements to report feedback on algorithm performance

https://aicentral.acrdsi.org

Apps | Beach House | Lung Cancer Risk C... | Longevity | Recipies | Blood Test for Lung... | Bookmarks | Bloomberg.com | Citibank® - Sign On | Service Chief Appro... | Daniel L. Rubin Lab:... | History of UML | IBM Valuing diversit... | New Tab | » | Other bo

AI CENTRAL

DATA SCIENCE INSTITUTE™
AMERICAN COLLEGE OF RADIOLOGY

ACR.or

Home | Data Science Institute | Editorial Board | Email DSI

Welcome to ACR Data Science Institute AI Central. This site is intended to provide easy-to-access, detailed information regarding FDA cleared AI medical products that are related to radiology and other imaging domains. Our editorial board and staff are continuously reviewing data from FDA public facing documents, vendor information and physician user feedback to provide you with up-to-date information that will help you to make appropriate purchasing decisions.

Check back regularly to see which new algorithms are available and have been added to the list. Send information on AI algorithms that are not listed and report missing information to DSI@acr.org.

Best,
Keith J. Dreyer, DO, PhD, FACR, FSIIM
Chairman of Editorial Board, AI Central Editorial

Board members: Christoph Wald, MD, PhD, MBA, FACR - Bibb Allen Jr., MD, FACR - Sheela Agarwal, MD, MBA - Bernardo C. Bizzo, MD, PhD - Judy W. Gichoya, MD, MS - Jay Patti, MD

Free Text Search	Company	Type	Finding	Body Area	Subspeciality	Modality	Age	Date Cleared
	(All)	(All)	(All)	(All)	(All)	(All)	(All)	Last 20 years

Date Cleared	Product	Company	Type	Finding	BodyArea	Subspeciality	Modality	Age
2022-04-26	Aidoc BriefCase for iPE Tr...	Aidoc Medical, Ltd.	CADt	Pulmonary embolism	Pulmonary...	Chest Imaging	CT	Adult only
2022-04-22	HealthOST	NanoxAI Ltd.	Image Processing/...	Spine display ,Vascular mea...	Spine	Neuroradiology	CT	Adult only
2022-04-15	AI-Rad Companion Brain...	Siemens Healthineers	Image Processing/...	Brain anatomy	Brain	Neuroradiology	MR	Unspecifiec
2022-04-06	ClariPulmo	Clari Pi	Image Processing/...	Pneumonia,COVID-19	Chest	Chest Imaging	CT	Adult only
2022-03-21	BriefCase- CTA	Aidoc Medical, Ltd.	CADt	Aneurysm	Brain	Neuroradiology	CTA	Unspecifiec
2022-03-14	BriefCase- Ptx	Aidoc Medical, Ltd.	CADt	Pneumothorax	Chest	Chest Imaging	XRAY	Unspecifiec
2022-03-14	Arterys MICA	Arterys Inc.	Image Processing/...	Cardiac measurements	Heart	Cardiac Imaging	MR	Unspecifiec
2022-03-10	ClearRead Xray Pneumot...	Riverain Technologies	CADt	Pneumothorax	Chest	Chest Imaging	XRAY	Adult only
2022-03-07	Syngo.CT Applications	Siemens Healthineers	Image Processing/...	Cardiac measurements	All	All	MR	Unspecifiec
2022-03-01	BoneView	Gleamer	CADe/x	Fracture	Upper Extr...	Musculoskeletal Im...	XRAY	Adult only
2022-02-24	Annalise Enterprise CXR ...	Annalise-AI Pty Ltd.	CADt	Pneumothorax	Chest	Chest Imaging	XRAY	Adult only
2022-02-18	Viz Aneurysm, Viz ANX	Viz. ai, inc.	CADt	Aneurysm	Brain	Neuroradiology	CT	Unspecifiec
2022-02-18	Deep Learning Image Re...	GE Medical Systems	Image Processing/...	Image Quality improvement	All	All	CT	Adult and p
2022-02-03	Contour ProtégéAI	Mim Software	Image Processing/...	Radiation treatment planning	All	All	CT,MR	Adult only

Filters

Challenge: Incorporating clinical history to determine a priori probability of disease

Clinical History is utilized in clinical practice and should be routinely used to augment AI performance

Bayesian a priori probability of disease is critical and determining a patient's chances of developing a given disease such as breast or lung cancer can have a major impact on AI algorithm utility

PLCO Example of Using Extensive Patient Data To Predict a priori Probability of Disease

- Published in 2009, the PLCO Screening Trial enrolled ~155,000 participants to determine whether certain screening exams reduced mortality from prostate, lung, colorectal and ovarian cancer
- The Prostate, Lung, Colorectal and Ovarian Cancer (PLCO) Screening Trial dataset provides an unparalleled resource for matching patients with the outcomes of demographically or diagnostically comparable patients
- These matched data can be used to inform a more sophisticated, personalized diagnostic decision-making process by tailoring imaging and testing follow-up intervals or even guiding intervention and prognosis
- They can also be incorporated into CAD algorithms to improve diagnostic efficacy by provided a priori likelihood of disease information.

PLCO Dataset

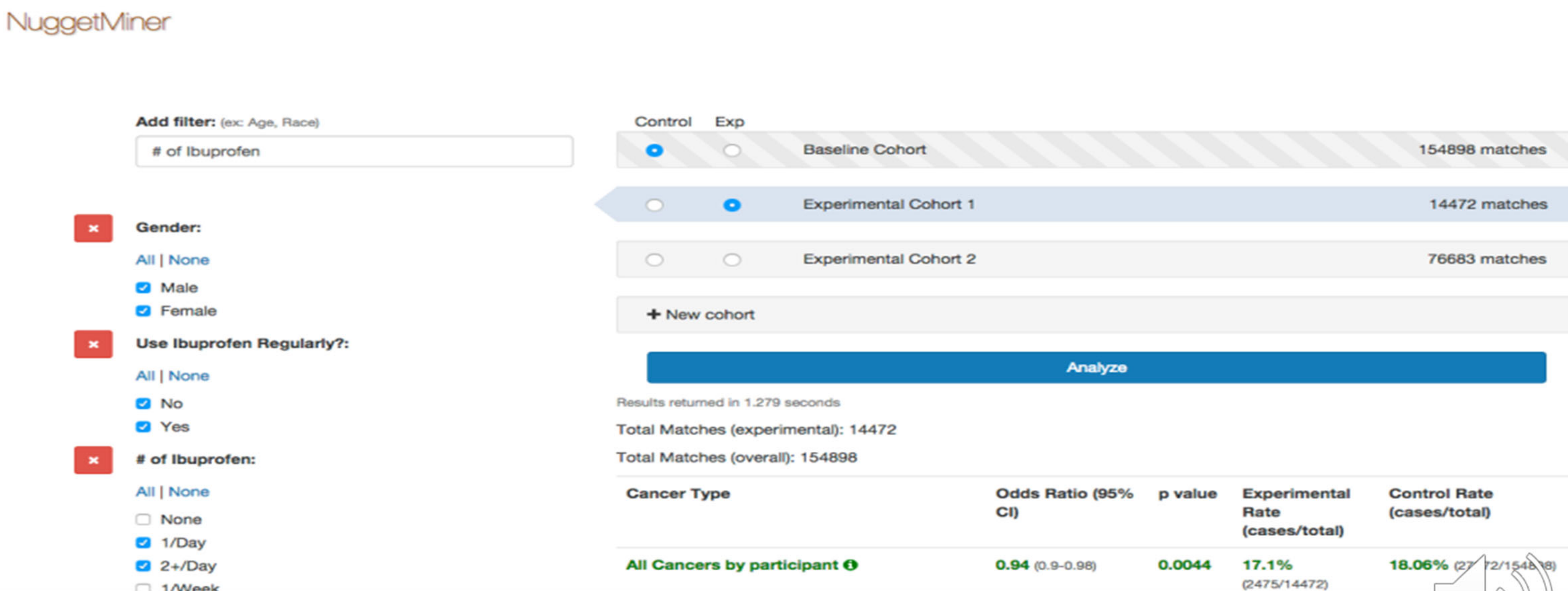
Table 2. Modified Logistic-Regression Prediction Model ($PLCO_{M2012}$) of Cancer Risk for 36,286 Control Participants Who Had Ever Smoked.*

Variable	Odds Ratio (95% CI)	P Value	Beta Coefficient
Age, per 1–yr increase†	1.081 (1.057–1.105)	<0.001	0.0778868
Race or ethnic group‡			
White	1.000		Reference group
Black	1.484 (1.083–2.033)	0.01	0.3944778
Hispanic	0.475 (0.195–1.160)	0.10	−0.7434744
Asian	0.627 (0.332–1.185)	0.15	−0.466585
American Indian or Alaskan Native	1		0
Native Hawaiian or Pacific Islander	2.793 (0.992–7.862)	0.05	1.027152
Education, per increase of 1 level†§	0.922 (0.874–0.972)	0.003	−0.0812744
Body-mass index, per 1-unit increase†	0.973 (0.955–0.991)	0.003	−0.0274194
Chronic obstructive pulmonary disease (yes vs. no)	1.427 (1.162–1.751)	0.001	0.3553063
Personal history of cancer (yes vs. no)	1.582 (1.172–2.128)	0.003	0.4589971
Family history of lung cancer (yes vs. no)	1.799 (1.471–2.200)	<0.001	0.587185
Smoking status (current vs. former)	1.297 (1.047–1.605)	0.02	0.2597431
Smoking intensity¶			−1.822606
Duration of smoking, per 1-yr increase†	1.032 (1.014–1.051)	0.001	0.0317321
Smoking quit time, per 1-yr increase†	0.970 (0.950–0.990)	0.003	−0.0308572
Model constant			−4.532506

* To calculate the 6-year probability of lung cancer in an individual person with the use of categorical variables, multiply the variable or the level beta coefficient of the variable by 1 if the factor is present and by 0 if it is absent. For continuous

NuggetMiner

“Instant Research”
Personalized Clinical Care



PLCO Participants Who Qualify for NLST

PLCO Participants Who Qualify for the National Lung Screening Trial:

All | None

- [ ] No
- [x] Yes

Age at BQ: 49 - 78

Height (inches): 48 - 84

Weight (lbs) at Baseline: 70 - 399

Gender:

All | None

- [x] Male
- [x] Female

Cigarette Smoking Status:

All | None

- [x] Never Smoked Cigarettes
- [x] Current Cigarette Smoker
- [x] Former Cigarette Smoker

Ever Smoked Cigars?:

All | None

Analyze

Results returned in 1.033 seconds

Print results

Permanent link

Total Matches (experimental): 29719

Total Matches (overall): 114697

Cancer Type	Relative Risk (95% CI)	p value	Experimental Rate (cases/total)	Control Rate (cases/total)
Mortality	2.09 (2.04-2.14)	<0.0001	24.19% (7188/29719)	11.59% (13289/114697)
All Cancers by participant	1.42 (1.39-1.45)	<0.0001	23.73% (7052/29719)	16.69% (19139/114697)
Lung	7.57 (7.05-8.12)	<0.0001	7.42% (2204/29719)	0.98% (1124/114697)
Prostate	0.8 (0.76-0.84)	<0.0001	9.37% (1690/18031)	11.78% (6266/53201)
Breast	0.73 (0.67-0.79)	<0.0001	2.32% (688/29719)	3.16% (3622/114697)
Bladder	2.54 (2.28-2.83)	<0.0001	1.82% (540/29719)	0.72% (821/114697)
Colorectum	1.28 (1.16-1.41)	<0.0001	1.8% (536/29719)	1.41% (1616/114697)
NonHodgkin's Lymphoma	0.83 (0.72-0.96)	0.0111	0.71% (210/29719)	0.85% ([illegible]0/114697)
Melanoma	0.71 (0.61-0.82)	<0.0001	0.71% (210/29719)	0.99% ([illegible]34/114697)

This project made possible by the PLCO and CDAS

Using Machine Learning to Determine Which Patients Should be Screened Rather than NLST Criteria for Example Using PLCO

Then Determine Time Period and Type of Follow Up Beyond Criteria Such as Lung-RADS Using NLST Database to Create malignancy Similarity Index

ORIGINAL ARTICLE

Clinical Impact and Generalizability of a Computer-Assisted Diagnostic Tool to Risk-Stratify Lung Nodules With CT

Scott J. Adams, MD, PhD[a], David K. Madtes, MD[b], Brent Burbridge, MD[c], Josiah Johnston, PhD[d], Ilya G. Goldberg, PhD[d], Eliot L. Siegel, MD[e], Paul Babyn, MDCM[f], Viswam S. Nair, MD, MS[b,g], Michael E. Calhoun, PhD[h]

Abstract

Objective: To evaluate whether an imaging classifier for radiology practice can improve lung nodule classification and follow-up.

Methods: A machine learning classifier was developed and trained using imaging data from the National Lung Screening Trial (NSLT) to produce a malignancy risk score (malignancy Similarity Index [mSI]) for individual lung nodules. In addition to NLST cohorts, external cohorts were developed from a tertiary referral lung cancer screening program data set and an external nonscreening data set of all nodules detected on CT. Performance of the mSI combined with Lung-RADS was compared with Lung-RADS alone and the Mayo and Brock risk calculators.

Results: We analyzed 963 subjects and 1,331 nodules across these cohorts. The mSI was comparable in accuracy (area under the curve = 0.89) to existing clinical risk models (area under the curve = 0.86-0.88) and independently predictive in the NLST cohort of 704 nodules. When compared with Lung-RADS, the mSI significantly increased sensitivity across all cohorts (25%-117%), with significant increases in specificity in the screening cohorts (17%-33%). When used in conjunction with Lung-RADS, use of mSI would result in earlier diagnoses and reduced follow-up across cohorts, including the potential for early diagnosis in 42% of malignant NLST nodules from prior-year CT scans.

Conclusion: A computer-assisted diagnosis software improved risk classification from chest CTs of screening and incidentally detected lung nodules compared with Lung-RADS. mSI added predictive value independent of existing radiological and clinical variables. These results suggest the generalizability and potential clinical impact of a tool that is straightforward to implement in practice.

Key Words: Artificial intelligence, CT, lung cancer, pulmonary nodule, radiomics

J Am Coll Radiol 2022;■:■-■.

INTRODUCTION

Lung cancer is the leading cause of cancer death, with 160,000 deaths per year in the United States [1]. We now understand that earlier lung cancer detection reduces mortality based largely on two large, prospective randomized clinical trials of lung cancer screening using low-dose CT, which demonstrated a 20% to 24% reduction in lung cancer mortality [2,3]. Lung cancer screening

[a]Department of Medical Imaging, University of Saskatchewan, Saskatoon, Canada; Scientific Director of the National Medical Imaging Clinic in Saskatoon.
[b]Clinical Research Division, Fred Hutchinson Cancer Research Center,
[g]Division of Pulmonary, Critical Care and Sleep Medicine, University of Washington School of Medicine, Seattle, Washington.
[h]Chief Technology Officer, RevealDx, Seattle, Washington.

AI Must Support Comparison Of Current And Prior Studies Which Is Not Currently Used For Mammography, Lung Cancer Etc.

But Is Routine For Radiologists
This Requires Databases To Have Longitudinal Data For AI Training

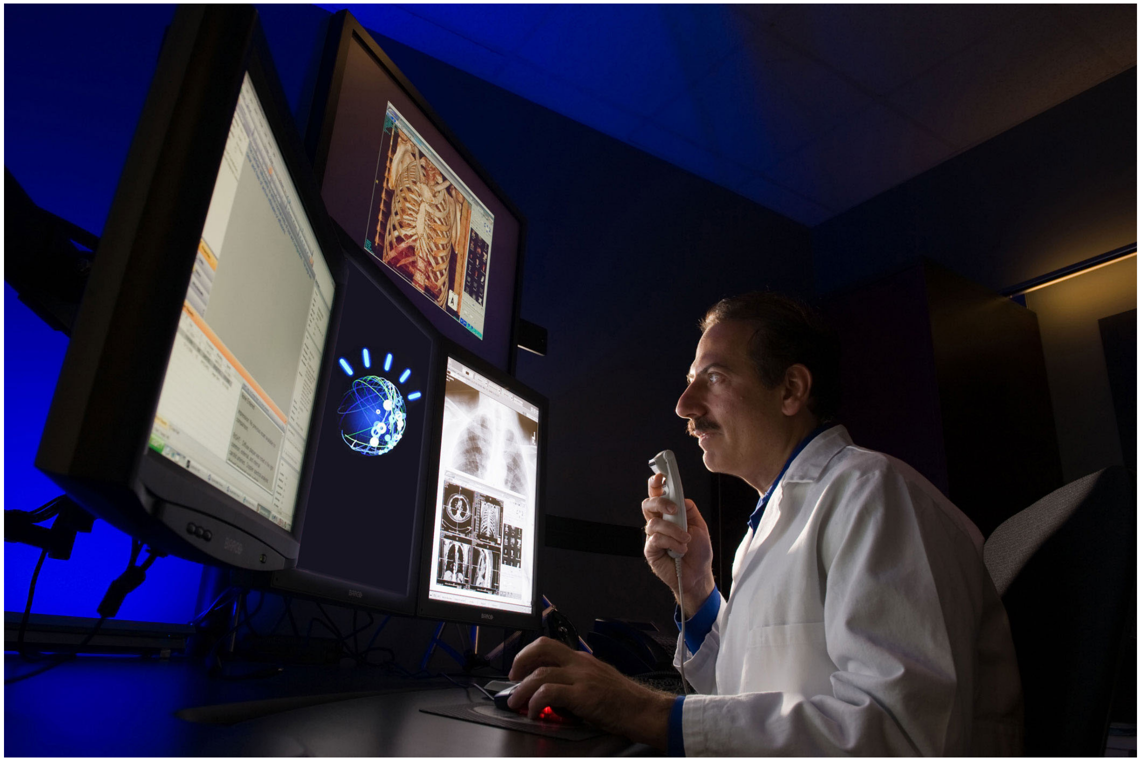
Challenge: In Order to Maintain Clinical Productivity, We Need to Integrate AI

- Fundamentally different workflow in clinical practice than reviewing for clinical trials
- How to consume quantitative AI in clinical practice?
  - Application consumed in separate window via cloud or local server or combination of the two
  - PACS
  - Platforms
    - Third party: Blackford/Bayer
    - Speech Recognition: Nuance
    - Advanced visualization
    - PACS
- Very little conformance with limited standards for AI incorporation into clinical workflow and today's platforms are proprietary

Medicolegal Mindset in Clinical Practice: Saving AI/CAD Annotations and Measurements and Diagnoses

- Across the US in mammography over 80% of mammo practices indicated they used CAD but almost none saved the markings for medicolegal reasons
- Medico legal issues are major issue for clinical use of quantitative algorithms but not for clinical trials for the most part
- Do we save markings clinically moving forward including quantitative data measured by algorithm?
    - Currently not done for mammography and many are not doing this for AI applications in stroke, intracranial hemorrhage, pulmonary embolism detection, etc.

How Can We Make Imaging Quantitative Data Available To the Next Generation of Clinical Analytic Systems



Capturing Quantitative Data in Clinical Practice

- Augmented Report (Radiology reporting equivalent of "appendix" at end of a research paper with supplemental materials)
  - Can capture quantitative data in clinical trials database but where does quantitative data go in radiology report with paradigm for text report to get generated by radiologist and then go into the EMR?
  - What about when a single study such as CT/PA generates 20 algorithms evaluating in a quantitative way BMD, coronary artery calcification, interstitial lung disease, lung volumes, flow dynamics, burden of emboli, pulmonary hypertension estimate, chamber sizes, etc. Where do we store that information if not in the radiologist report?
  - Do these data get stored in DICOM SR? If so, is it stored with the PACS? How do we make it available to other algorithms across the outpatient or hospital network or networks?

- But no concerted effort that I’m aware of tackles the challenge to make these quantitative data available in machine readable format to the EMR for clinical/analytics- decision support
- This will be absolutely mandatory for radiology and nuclear medicine to stay relevant in oncologic and other clinical practice as data continue to become more complex and guidelines are created for quality clinical practice
- Making data available to other algorithms in EMR
  - Radiology is not an independent island and as other specialties develop their own AI algorithms those might need to query radiology reports or supplemental data
    - E.g. oncology decision support algorithms may want to directly take data in standardized format from radiology supplemental data outside of a radiology report

AI In Clinical Practice
Should AI Be Directed at Oncologists for Direct Consumption?

- Could quantitative AI be consumed directly by oncologists or do radiologists need to review and comment on all results?
- Are the QIN Tools, for example, designed to be used by radiologists, oncologists, primary care specialists, or even patients as is becoming commonplace with dermatology apps?
- Should those tools be available on hospital Enterprise Imaging Systems which are typically web based or just available on radiologist PACS workstations or advanced visualization workstations?
- In order for imaging and AI tools to be useful they need to be in a format that can be easily understood and reviewed by oncologists

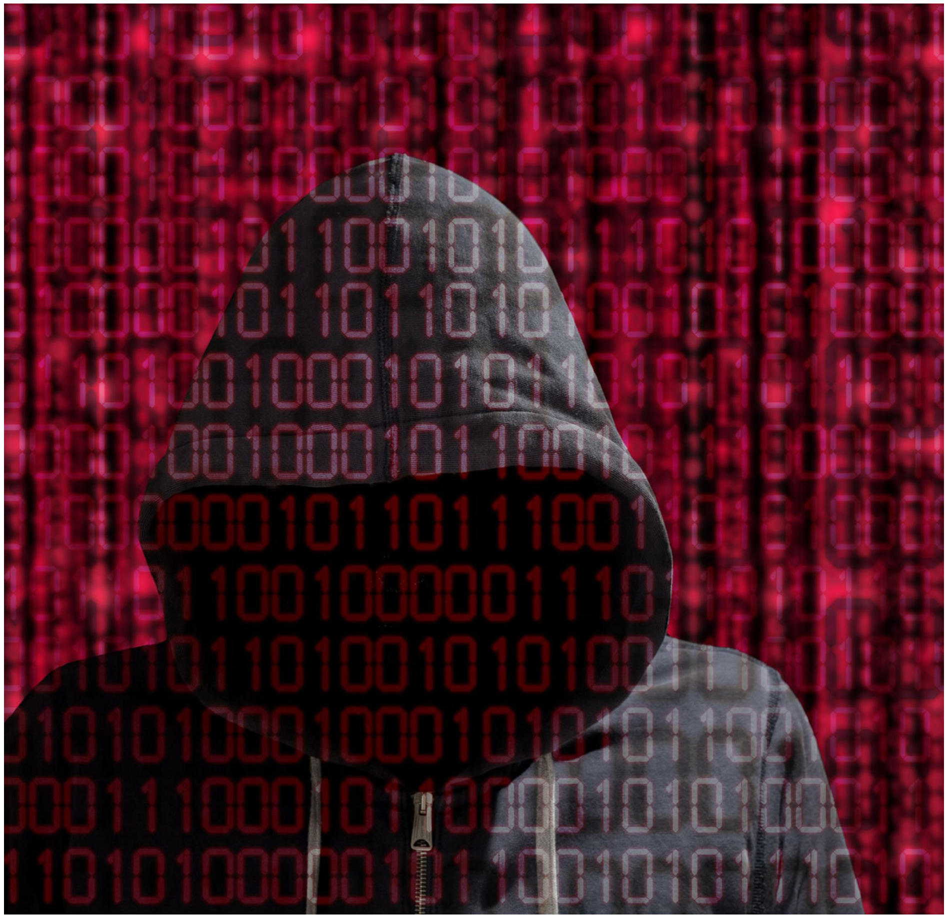
Clinical Evaluation of AI Algorithms
Could NCI funded Algorithms Be Run on Commercial AI Platforms that Could Also Support Local AI Algorithm Development?

- Some commercial platforms currently let users run their own "home grown" algorithms or downloaded research algorithms outside of FDA clearance
  - There is no regulatory prohibition against using these in routine clinical care
- Could NCI engage with these platforms and provide non-FDA approved software directly for research, clinician trials, and clinical care through the platforms?
- Are there other ways that I can access NCI and other AI software when I do routine clinical interpretation?

- There is currently no consensus about a platform for AI that would allow assessment of agreement by the radiologist or oncologist and allow consumption of AI algorithms not only purchased with FDA clearance but those from NIH/NCI and others
- Would need to have a viewer to be able to directly interact with algorithms such as tumor segmentation

IT Security

- With a high percentage of approved algorithms only available in the cloud, IT privacy and security concerns are substantial especially given the current environment of cyberthreats
- It has been difficult and time consuming to vet a single AI commercial algorithm and this challenge increases exponentially for numerous AI algorithms whether consumed in the cloud or locally
- AI platforms may mitigate this somewhat by having a single platform vendor but understanding how information flows out of that platform will still present security challenges
  - The VA and DoD and Indian Health federal sector has particular challenges with consumption of software in the cloud but this is slowly improving



Reimbursement for AI

Reimbursement for AI Still In Very Early Stages

- In a groundbreaking development, CMS has recently approved for a three-year period, a $1,040 technical fee reimbursement for Viz.ai, currently with a cap of $25,000 per institution using the NTAP (New Technology Add-on Payment) for hospital inpatients
- This has paved the potential for additional possibilities for AI for rapid diagnosis in order to enable more rapid treatment or to more rapidly determine that treatment is not needed
- There has also been recent approval by AMA of a test CPT code for reimbursement for radiology artificial intelligence (AI) for detection of vertebral body fractures and estimation of bone mineral density on CT scans of the abdomen and pelvis is a milestone

Reimbursement for AI

- However, there is currently no payor reimbursement for the dozens of other FDA cleared AI software and imaging departments and hospitals have had to absorb costs for AI
  - AI software may be bundled into cost of study as was the case with mammography CAD which initially reimbursed $12 per study and then that was withdrawn
- In the future, AI software may be considered to be just part what is expected with a PACS or advanced visualization or enterprise imaging solution

Discovering and Consuming Databases and AI Algorithms from NIH and Others
No NLM Resource

- At best, freely sharable databases are accessed using their own idiosyncratic web portal
- Currently no index of databases or their content
- No standards exist to describe how databases can “advertise” their content and availability (free or business model) and their data provenance and sources and peer review, etc.
- Would be wonderful project to investigate the creation of an XML standard for describing the content of databases
- Nuclear medicine is a very much overlooked source of databases which has put the NM community behind in AI applications

SOLICITATION

AIMI Datasets

PERSON OF CONTACT	PROPOSAL DUE DATE	NAICS CODES
Connor Cigrang (connor.cigrang@nih.gov)	Tuesday, June 14, 2022 (Due in 3 weeks 4 days)	5417, 541714, 541990

Description

To further advance AI in Medical Imaging (AIMI) large datasets, acquired through routine standard of care, are needed to train and evaluate the performance of the ML/AI algorithms. The datasets need to be correctly de-identified to maintain patient privacy while at the same time preserving as much scientifically relevant information as possible. Large datasets from the existing standard of care radiology practice, along with companion clinical data, are needed for the training and development of ML/AI algorithms by the research community.

RESEARCH CATEGORY
Research & Development Services & Subcontracts

SOLICITATION NUMBER
S22-068

NCI OA CONTRACT NUMBER
75N91019D00024 – NCI-FFRDC

Proposal Instructions

Offerors must contact Connor Cigrang, Subcontract Administrator, for the official RFP Document and Attachments (please provide your organization's SAM.gov Unique Entity Identifier in your inquiry).

Interested vendors are advised to submit a request for the RFP package to the Subcontract Administrator by May 27, 2022. All questions and requests for clarification are due by 5PM, June 3, 2022 for which answers will be provided to all interested vendors by 5PM, June 7, 2022.

Extending DICOM Standard to Optimize It for AI/ML Development and Deployment

- Hospital PACS are not designed as research repositories and function poorly in that capacity
- AI/ML research could be facilitated by storing downsampled versions of images or compressed versions of images routinely using only a tiny fraction of storage currently required for full datasets
- Work to create more robust and secure repository and strategies to make access to research databases much easier and more efficient
- Current DICOM databases used for research need enhancements in encryption and access control for use in the cloud

What Could NCI Do to Facilitate AI Implementation in Routine Clinical Practice

- Create a well-defined mechanism for not only researchers but also clinicians and vendors to find and utilize NIH/NCI funded datasets and also algorithms
  - E.g. if I want to use Ron Summers' AI algorithms, how do I do that as a radiologist? As a vendor?
  - Can I use a dataset such as PLCO to develop a commercial application as a vendor? Can I use it in clinical practice as a radiologist or oncologist?
- Could NCI organize/create a forum or even structured process for user rating of algorithms and databases from a research but also a clinical perspective?
- Would it be possible to create a repository of image acquisition protocols developed for clinical trials to be candidates for routine clinical practice? Perhaps these could be automatically uploaded by CT, MRI, and PET/CT vendors
- Could NCI create an index of algorithms or databases similar to NLM PubMed that would also list the type of subjects used to develop the algorithms or populate the databases? Do we already have this for AI algorithms developed for NCI?
- Could NCI sponsor a grant that would require an AI algorithm for example that finds lymph nodes on an abdominal/pelvic scan to learn/improve over time?
  - Or sponsor one that would incorporate personalized patient data and history into algorithm performance

What Could NCI Do to Facilitate AI Implementation in Routine Clinical Practice

- Create a platform that could be NCI standard required for AI algorithms that would also include a visualization engine and allow user feedback in a standardized way
- Revise current guidelines such as screening based on NLST trial to be personalized based on databases such as PLCO that personalize risk for a single patient who might have a combination of factors based on ethnicity, family history, etc. that could be him/her at risk without history of smoking for example
- Revise guidelines for follow-up for lung nodules and other findings based on evaluation of pixel plus patient information or change in lesion characteristics over time
- Create software to facilitate routine parallel research repositories that are already downsampled/compressed for machine learning with improved encryption
- Create standard mechanism for representing and storing quantitative data in a routine scan such as coronary calcification, lung volume, splenic volume, bone mineral density, etc.

Bringing AI from Hype to Reality for Routine Clinical Practice: Addressing the Gaps and Opportunities for NCI

Eliot Siegel, M.D., FSIIM, FACR

Professor and Vice Chairman University of Maryland School of Medicine Department of Diagnostic Radiology and Nuclear Medicine

Chief Imaging VA Maryland Healthcare System