AMIA

INFORMATICS PROFESSIONALS. LEADING THE WAY.

Artificial Intelligence and Data Science Year in Review

Yuan Luo, PhD, FACMI, FAMIA, FIAHSI

Northwestern University Feinberg School of Medicine

@yuanhypnosluo

Genevieve Melton-Meaux, MD, PhD, FACMI

University of Minnesota School of Medicine

@MeltonMeaux

Disclosure

Industry: Walmart (Luo)

Funding: NIH (Luo, Melton), AHRQ (Melton), PCORI (Melton), FDA (Melton), Minnesota State (Melton)

The content is solely our responsibility and does not represent the official views of our institutions or the NIH, AHRQ, PCORI, FDA or the State of Minnesota.

Copyrights are reserved by the authors who provided the slides that we adapted from.

Learning Objectives

After participating in this session, the learner should be better able to:

- Have a sense of the scope of AI and Data Science for health care
- Summarize the state-of-the-art research in AI and Data Science for health care
- Identify research and development opportunities in this space and advance the field

Process and caveats

The content is developed by the informatics community, for the informatics community, we are not boiling the ocean

Teamwork makes dreamwork, we thank talented volunteers including Diego Trujillo, Malvika Pillai, Suhana Bedi, Tina Yi Jin Hsieh, Yikuan Li, Sicheng Zhou, Hanyin Wang, Yidi Huang, Young Sang Choi, Yu Huang, Wenxin Chen, and generous help from Jim Cimino

Each volunteer picked a topic, for which they were asked to nominate papers that are significant (new model, new insights, new task)

Focused on original research of AI and data science for health care, exclude reviews, viewpoints …

Audiovisual materials requested from authors for "interesting" papers, tried our best to acknowledge names, all content from authors or original papers

Genevieve and I finalized the selection and organized the slides to divide-and-conquer the presentation

Apologies for anything missed or misjudged, mistakes are all our responsibility

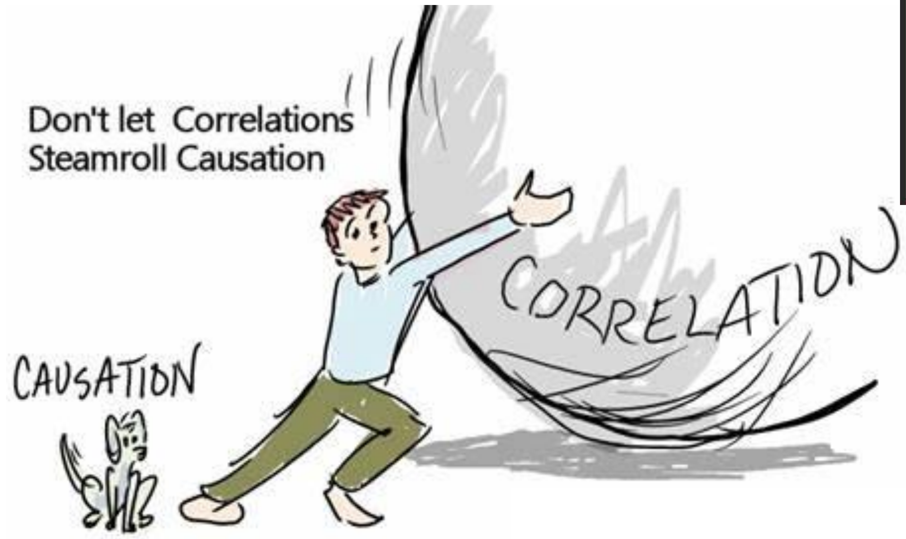
Which subjects best describe your interest (can select >1)?

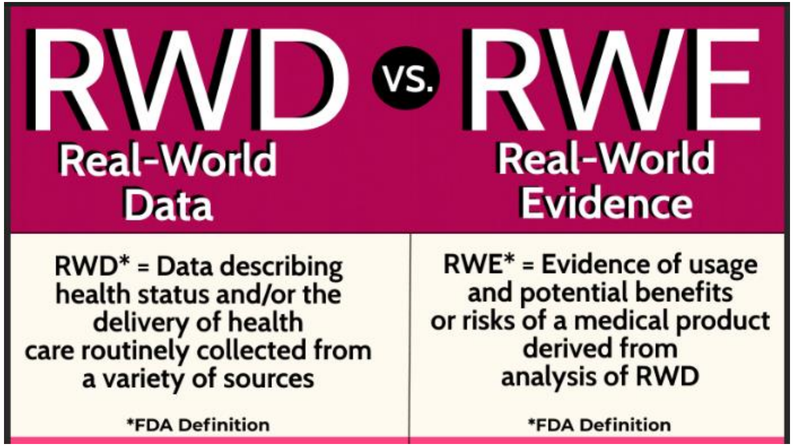
- Real-World Evidence and Causal Inference
- Multi-modal AI and Generative AI
- Clinical Decision Support
- Imaging Informatics
- Evaluation and Implementation
- Large Language Models and Natural Language Processing
- Ethical, Legal and Social Issues
- Public and Global Health Informatics

Topics

- Real-World Evidence and Causal Inference
- Multi-modal AI and Generative AI
- Clinical Decision Support
- Imaging Informatics
- Evaluation and Implementation
- Large Language Models and Natural Language Processing
- Ethical, Legal and Social Issues
- Public and Global Health Informatics

Real-World Evidence and Causal Inference





Real-World Evidence and Causal Inference

nature communications

Published: 11 December 2023

High-throughput target trial emulation for Alzheimer's disease drug repurposing with real-world data

Chengxi Zang, Hao Zhang, Jie Xu, Hansi Zhang, Sajjad Fouladvand, Shreyas Havaldar, Feixiong Cheng, Kun Chen, Yong Chen, Benjamin S. Glicksberg, Jin Chen, Jiang Bian & Fei Wang

Nature Communications **14**, Article number: 8180 (2023)

Target Trial Emulation

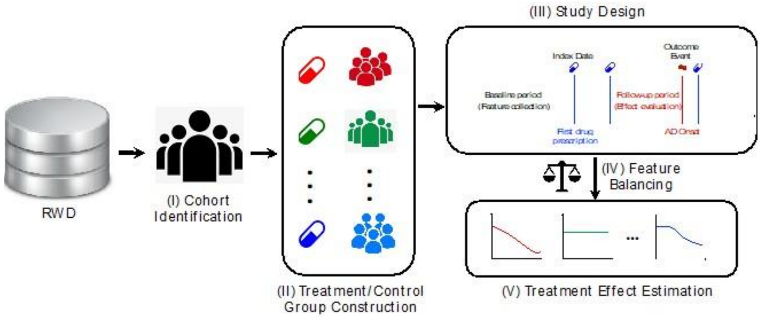
Randomized Controlled Trials (RCTs)

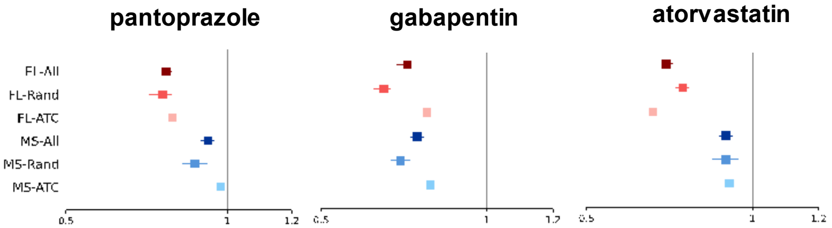
- Stringent inclusion/exclusion criteria
- Expensive in terms of both time and money
- Not representative of real-world patients who will receive the treatment

Target Trial Emulation

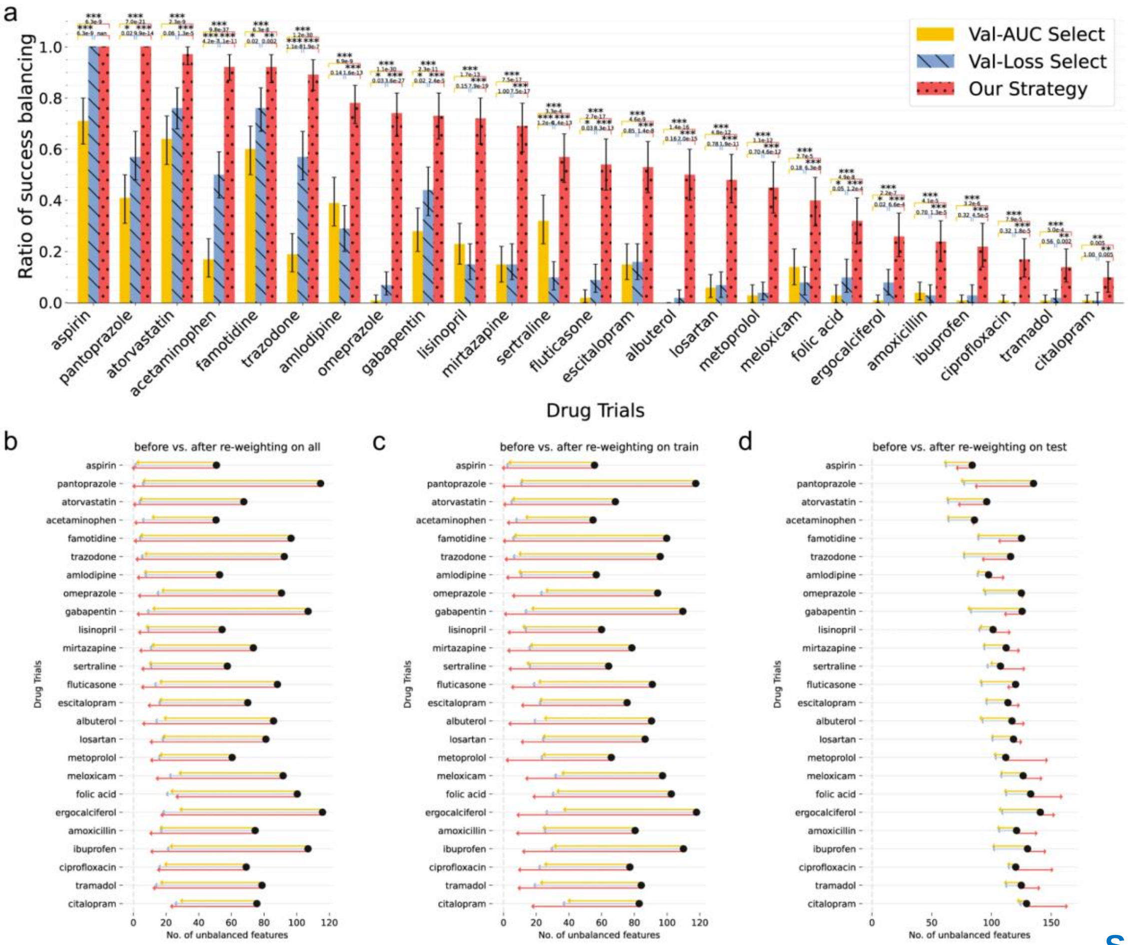
- Mimicking RCTs using real world data (e.g., electronic health records)
- Large sample size
- Real world evidence

TTE pipeline for identification of drugs that could be beneficial to Alzheimer's disease





Slides courtesy of Fei Wang



AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

Slides courtesy of Fei Wang

Real-World Evidence and Causal Inference

INFORMATICS PROFESSIONALS. LEADING THE WAY.

Multiply Robust Federated Estimation of Targeted Average Treatment Effects

Larry Han
Department of Health Sciences
Northeastern University
Boston, MA 02115
lar.han@northeastern.edu

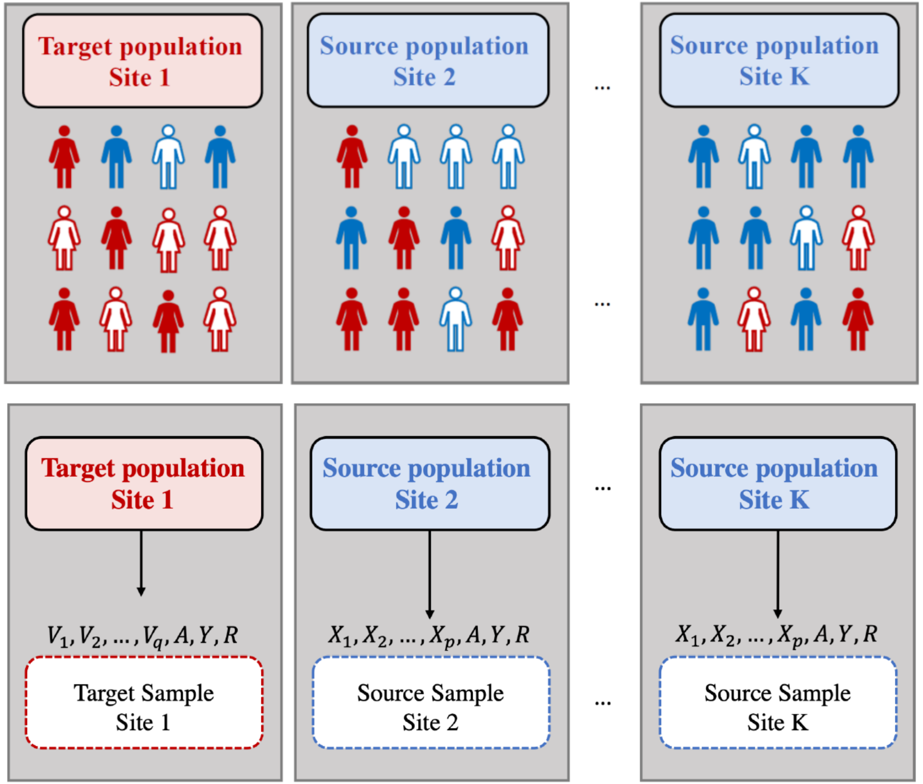
Zhu Shen
Department of Biostatistics
Harvard University
Boston, MA 02115
zhushen@g.harvard.edu

Jose Zubizarreta
Departments of Health Care Policy, Biostatistics, and Statistics
Harvard University
Boston, MA 02115
zubizarreta@hcp.med.harvard.edu

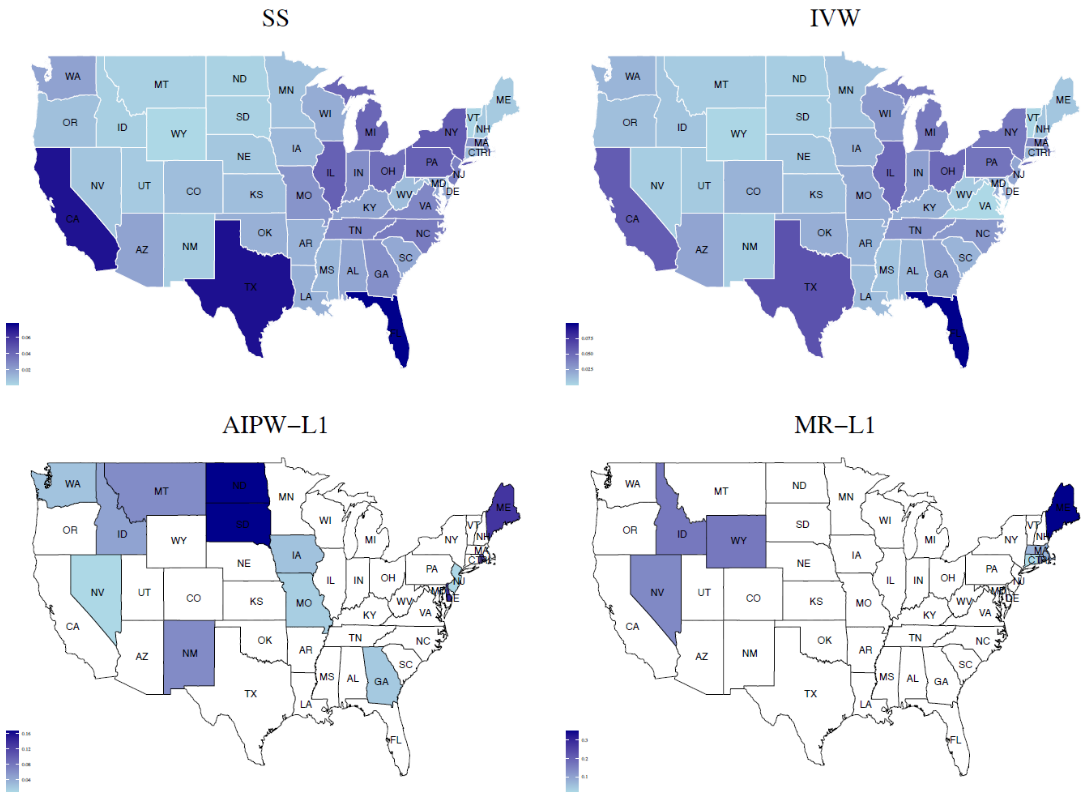
Advances in Neural Information Processing Systems 36 (NeurIPS 2023)

Federated estimators across data sources for causal inferences

In multi-source settings, how can we make optimal use of available data to make causal inferences for a target population of interest?



Federation weights across states for the PCI treatment effect in Maine with four federated estimators



Real World Evidence and Causal Inference

ELSEVIER

The Journal of Thoracic and Cardiovascular Surgery

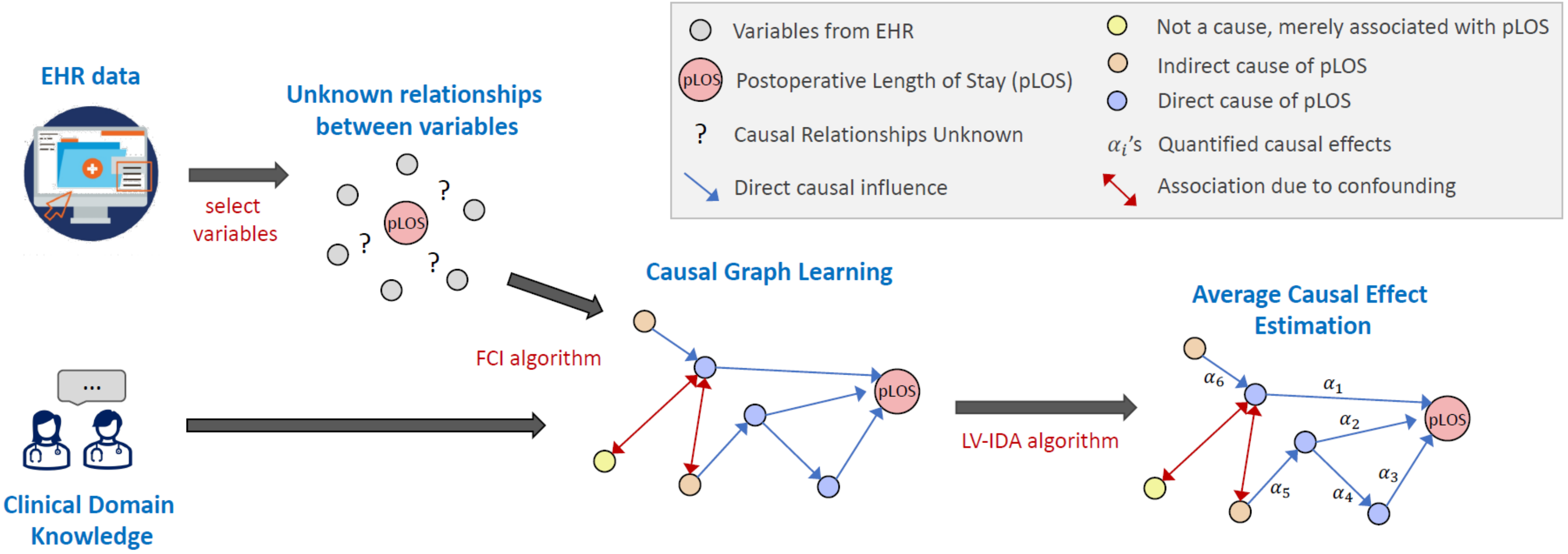
Volume 166, Issue 5, November 2023, Pages e446-e462

Adult: Perioperative Management: Evolving Technology

Causal determinants of postoperative length of stay in cardiac surgery using causal graphical learning

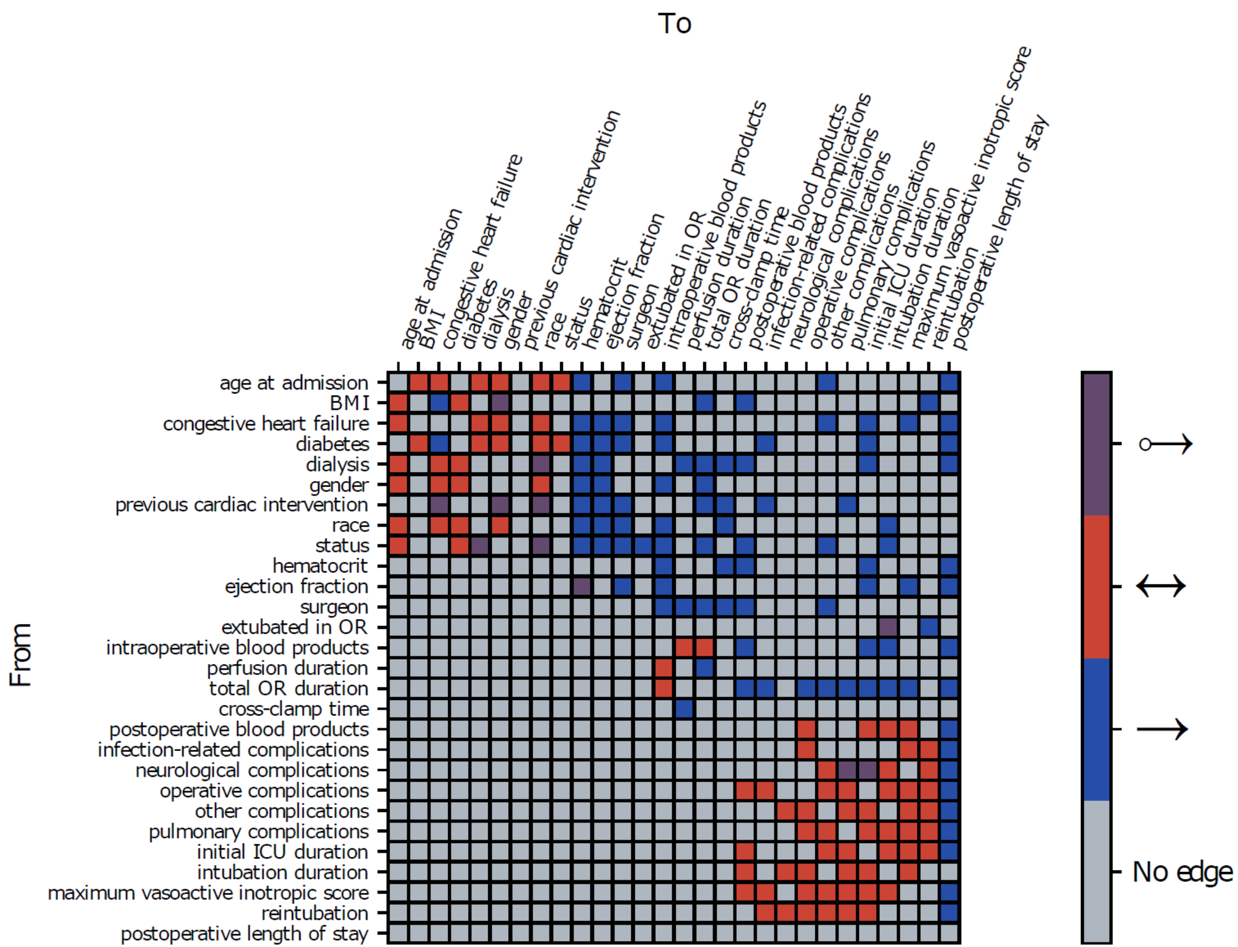
Jaron J.R. Lee PhB [a b], Ranjani Srinivasan MS [b c], Chin Siang Ong MBBS, PhD [d], Diane Alejo BA [e], Stefano Schena MD, PhD [e], Ilya Shpitser PhD [a b], Marc Sussman MD [e], Glenn J.R. Whitman MD [e], Daniel Malinsky PhD [f]

Causal Determinants of Postoperative Length of Stay in Cardiac Surgery



Causal Graphical Learning uncovers true causal relationships and their strengths, leading to actionable insights that can reduce post-operative length of stay.

pLOS: postoperative Length of Stay
EHR: Electronic Health Records
FCI: Fast Causal Inference
LVIDA: Latent Variable Intervention-calculus when the DAG is Absent


To
age at admission
BMI
congestive heart failure
diabetes
dialysis
gender
previous cardiac intervention
race
status
hematocrit
ejection fraction
surgeon
extubated in OR
intraoperative blood products
perfusion duration
total OR duration
cross-clamp time
postoperative blood products
infection-related complications
neurological complications
operative complications
other complications
pulmonary complications
initial ICU duration
intubation duration
maximum vasoactive inotropic score
reintubation
postoperative length of stay
From
No edge

Real World Evidence and Causal Inference

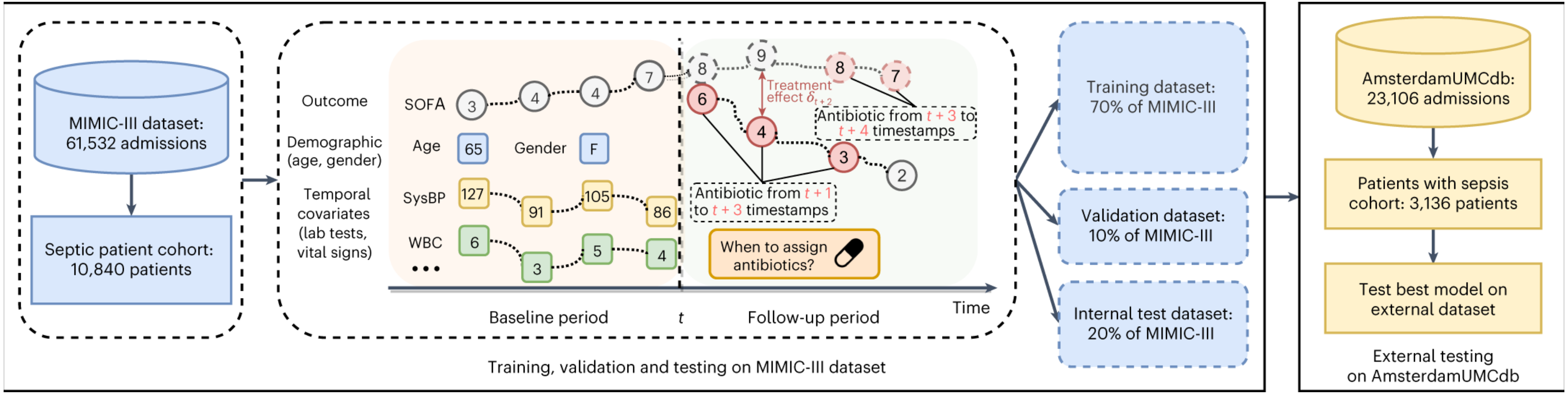
nature machine intelligence

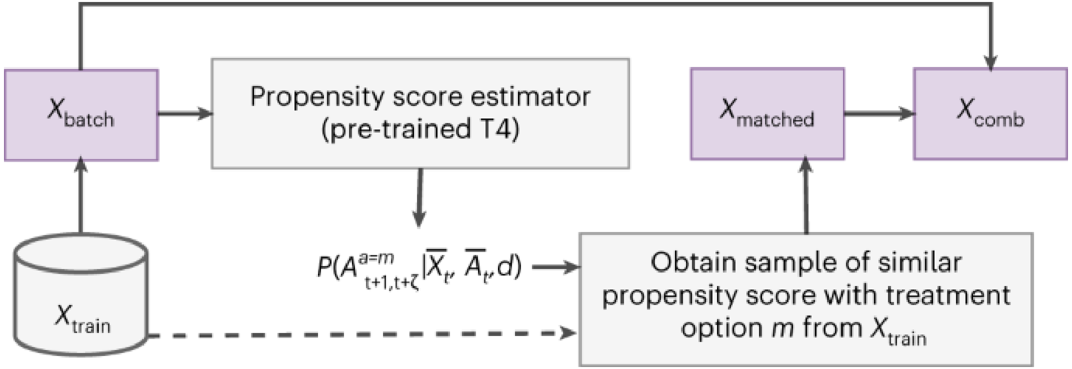
Published: 06 April 2023

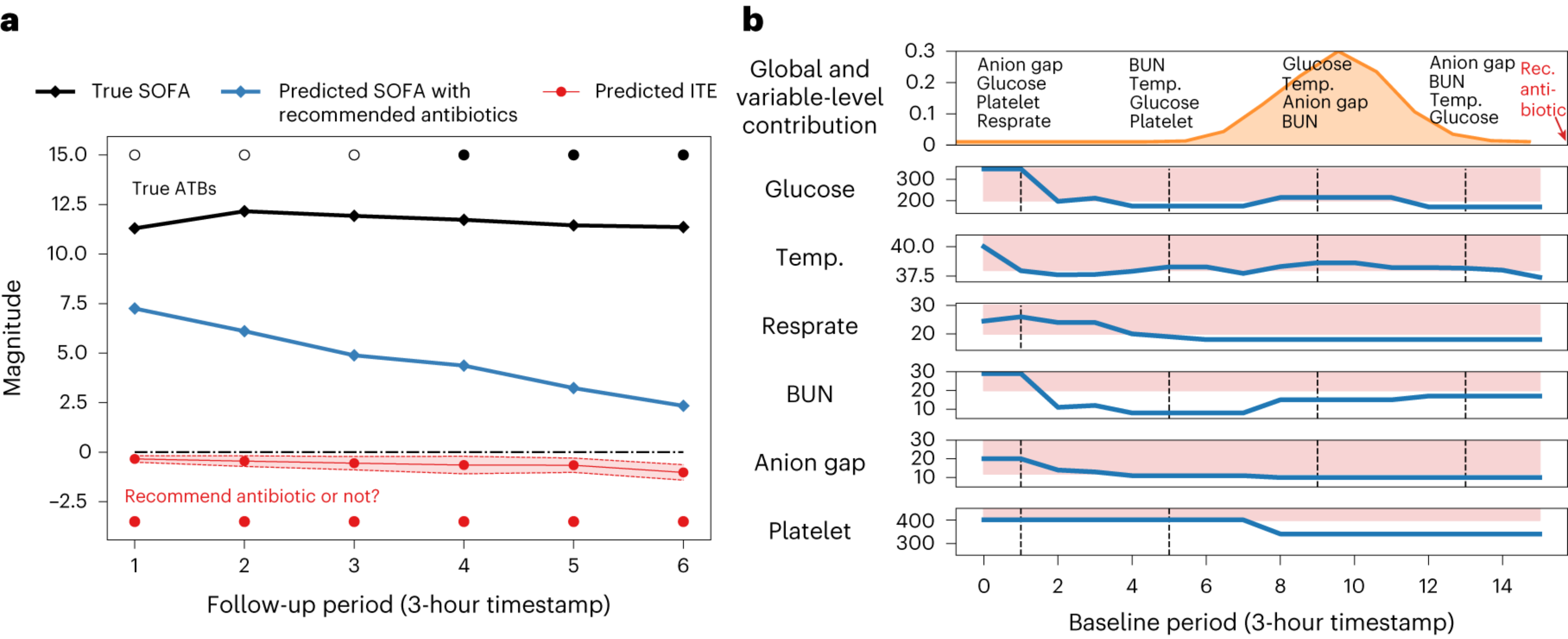
Estimating treatment effects for time-to-treatment antibiotic stewardship in sepsis

Ruoqi Liu, Katherine M. Hunold, Jeffrey M. Caterino & Ping Zhang

Nature Machine Intelligence **5**, 421–431 (2023)


MIMIC-III dataset: 61,532 admissions
Septic patient cohort: 10,840 patients
Outcome
SOFA
Demographic (age, gender)
Age
Gender
Temporal covariates (lab tests, vital signs)
SysBP
WBC
Treatment effect
Antibiotic from t + 3 to t + 4 timestamps
Antibiotic from t + 1 to t + 3 timestamps
When to assign antibiotics?
Baseline period
t
Follow-up period
Time
Training, validation and testing on MIMIC-III dataset
Training dataset: 70% of MIMIC-III
Validation dataset: 10% of MIMIC-III
Internal test dataset: 20% of MIMIC-III
AmsterdamUMCdb: 23,106 admissions
Patients with sepsis cohort: 3,136 patients
Test best model on external dataset
External testing on AmsterdamUMCdb


X_{batch}
Propensity score estimator (pre-trained T4)
$X_{matched}$
X_{comb}
$P(A^{a=m}_{t+1,t+\zeta} | \overline{X}_t, \overline{A}_t, d)$
Obtain sample of similar propensity score with treatment option m from X_{train}
X_{train}



a
True SOFA
Predicted SOFA with recommended antibiotics
Predicted ITE
True ATBs
Recommend antibiotic or not?
Magnitude
15.0
12.5
10.0
7.5
5.0
2.5
0
−2.5
1
2
3
4
5
6
Follow-up period (3-hour timestamp)
b
Global and variable-level contribution
0.3
0.2
0.1
0
Anion gap
Glucose
Platelet
Resprate
BUN
Temp.
Glucose
Platelet
Glucose
Temp.
Anion gap
BUN
Anion gap
BUN
Temp.
Glucose
Rec. anti-biotic
Glucose
300
200
Temp.
40.0
37.5
Resprate
30
20
BUN
30
20
10
Anion gap
30
20
10
Platelet
400
300
0
2
4
6
8
10
12
14
Baseline period (3-hour timestamp)

Real World Evidence and Causal Inference

Journal of Biomedical Informatics

Volume 139, March 2023, 104298

Original Research

Causal knowledge graph construction and evaluation for clinical decision support of diabetic nephropathy

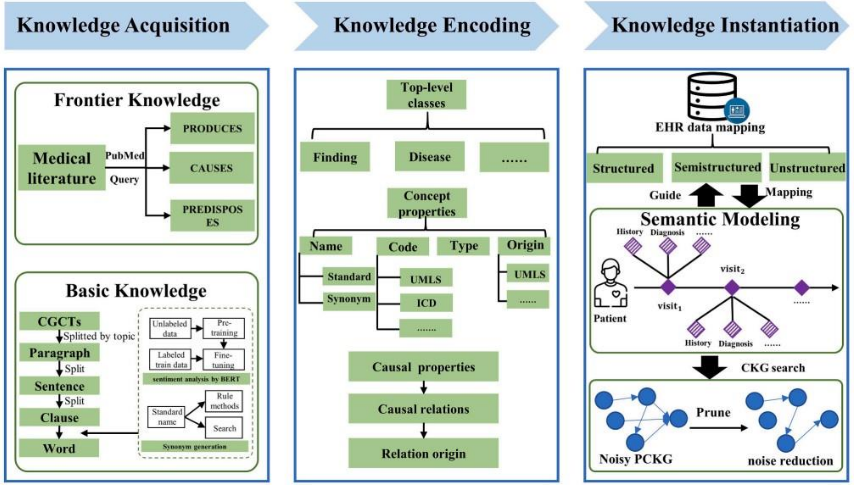
Kewei Lyu [a, 1], Yu Tian [a, 1], Yong Shang [b], Tianshu Zhou [b], Ziyue Yang [a], Qianghua Liu [a], Xi Yao [c], Ping Zhang [c], Jianghua Chen [c], Jingsong Li [a, b]

Objective

- ✓ Causal knowledge (CK) plays a critical role in clinical decision-making of diagnosing complex medical conditions such as diabetic nephropathy.
- ✓ A comprehensive evaluation based on real-world data and methods for handling potential knowledge noise of causal knowledge graphs (CKGs) are still lacking.

Methods

- ➢ Propose a **semi-automated construction frame work of a large-scale and high-quality CKG**
- ➢ Design **pruning strategies** to reduce the potential knowledge noise introduced by medical literature



Causal knowledge graph

- ➢ 153,289 concepts
- ➢ 1,719,968 causal triplets

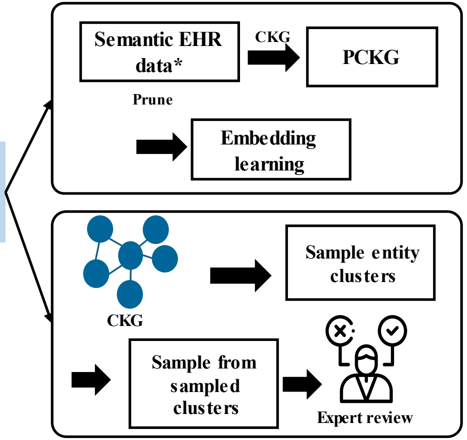
Slides courtesy of Jingsong Li

Results

1 Knowledge completeness evaluation

The CKG **covered 93.9% of the causal relationships** between diseases and diagnostic evidence in clinical texts.

2 Knowledge accuracy evaluation



CKG	Overall average accuracy(±SD)	The average accuracy of triples from SemMedDB (±SD)
Unpruned	67.5% ± 5.0%	65.4% ± 5.2%
Pruned based cooccurrence ≥ 1e-4	73.0% ± 4.4%	72.9% ± 3.2%
Pruned based causality ratio ≥ 0.3	72.8% ± 6.5%	70.1% ± 9.6%

The knowledge accuracy of the CKG was **significantly improved after pruning**.

Conclusion

This paper contributes to advancing clinical decision support systems by incorporating causal knowledge to support more informed and accurate diagnostic decisions in the context of diabetic nephropathy.

Proposes a framework for constructing and evaluating the CKG, and designs pruning strategies for noise reduction.

Contributes to a better diagnosis and management of complex medical conditions such as diabetic nephropathy.

*This study retrospectively analyzed the **EHR data** from the First Affiliated Hospital of Zhejiang University (FAHZJU) from October 2010 to October 2020, a total of 1427 inpatients with 3045 hospitalization records were ultimately enrolled. The clinical research ethics committee approved this study. (No. 2022-683)

Slides courtesy of Jingsong Li

Real World Evidence and Causal Inference

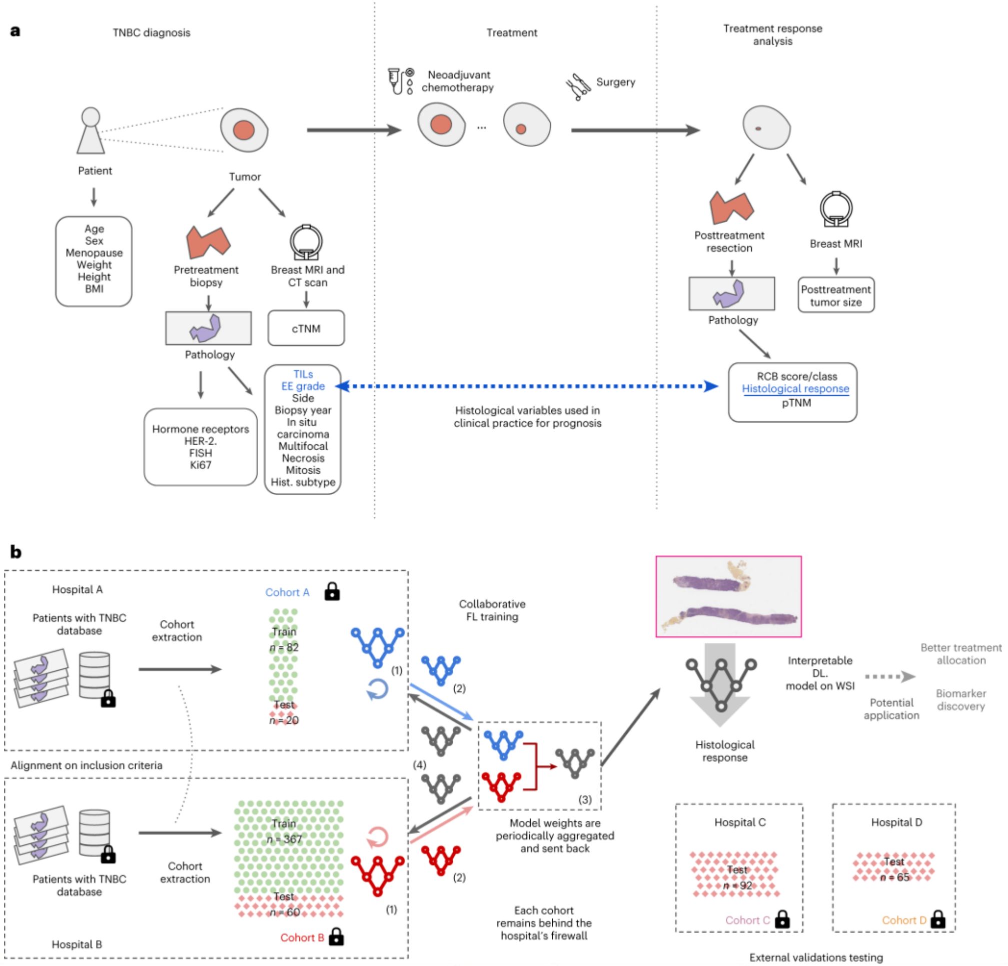
nature medicine

Article | Published: 19 January 2023

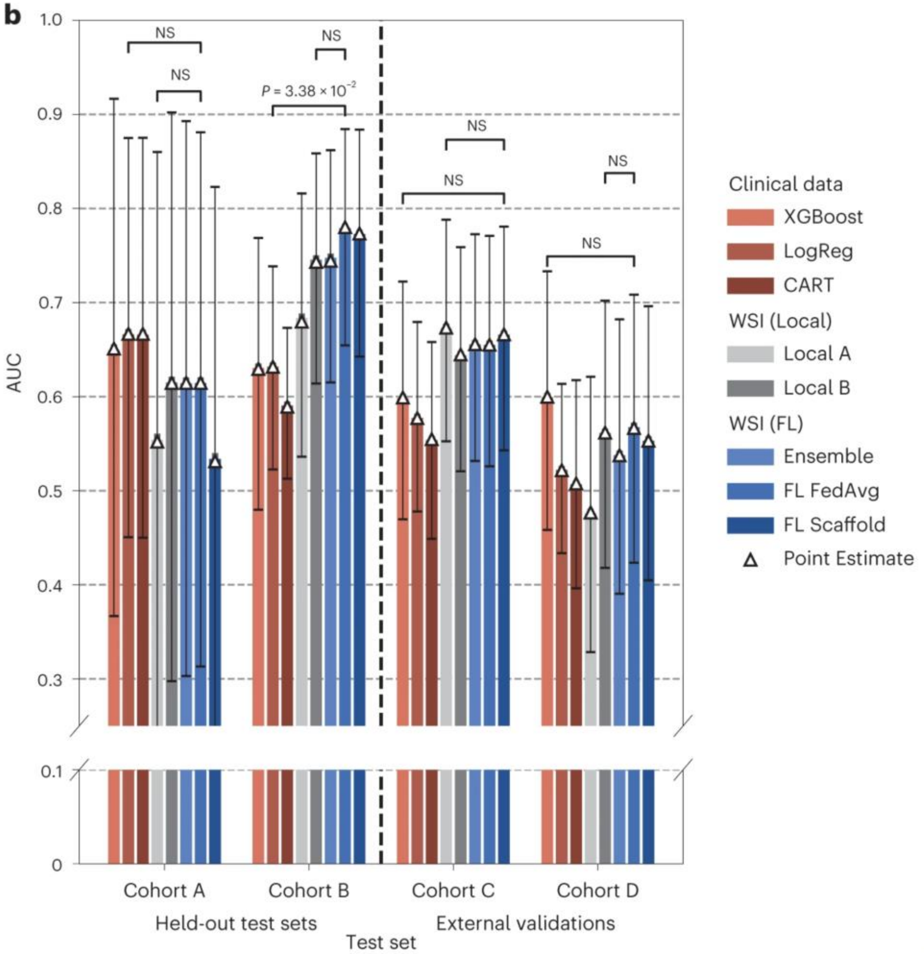
Federated learning for predicting histological response to neoadjuvant chemotherapy in triple-negative breast cancer

Jean Ogier du Terrail, Armand Leopold, Clément Joly, Constance Béguier, Mathieu Andreux, Charles Maussion, Benoît Schmauch, Eric W. Tramel, Etienne Bendjebbar, Mikhail Zaslavskiy, Gilles Wainrib, Maud Milder, Julie Gervasoni, Julien Guerin, Thierry Durand, Alain Livartowski, Kelvin Moutet, Clément Gautier, Inal Djafar, Anne-Laure Moisson, Camille Marini, Mathieu Galtier, Félix Balazard, Rémy Dubois, Jeverson Moreira, Antoine Simon, Damien Drubay, Magali Lacroix-Triki, Camille Franchet, Guillaume Bataillon & Pierre-Etienne Heudel

Nature Medicine **29**, 135–146 (2023)



Slides courtesy of Jingsong Li



Slides courtesy of Jingsong Li

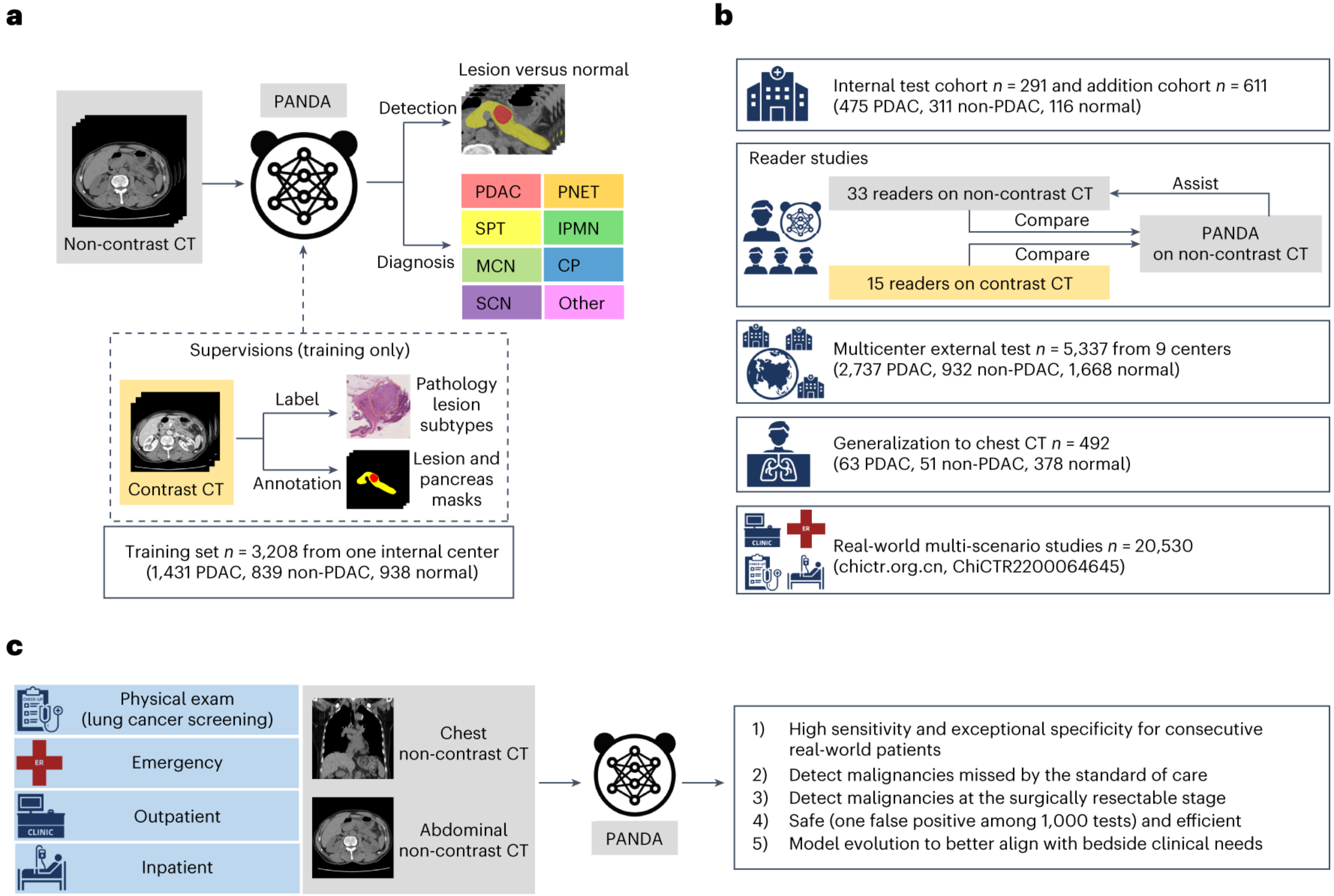
Real World Evidence and Causal Inference

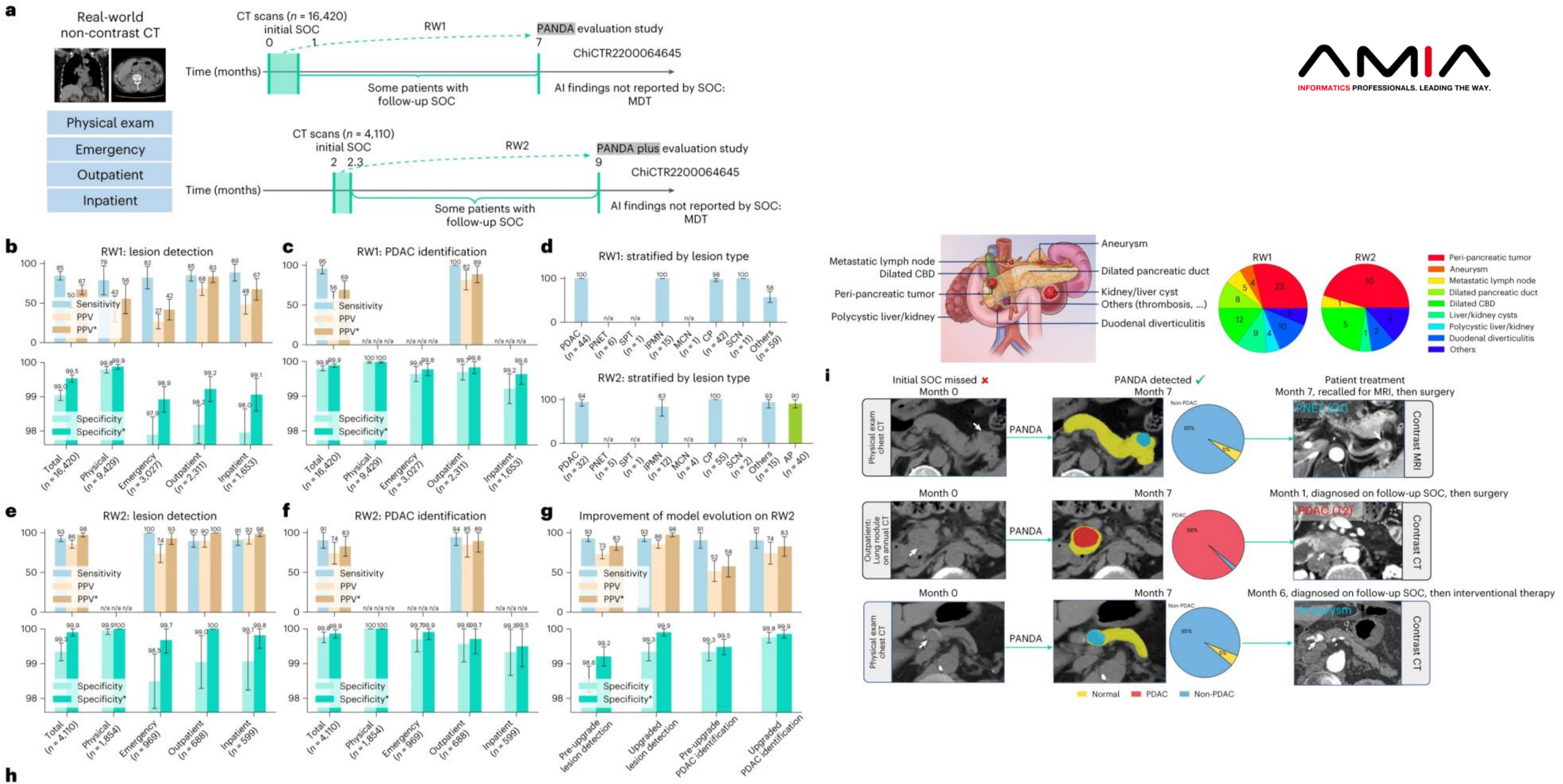
Article | Open access | Published: 20 November 2023

Large-scale pancreatic cancer detection via non-contrast CT and deep learning

Kai Cao, Yingda Xia, Jiawen Yao, Xu Han, Lukas Lambert, Tingting Zhang, Wei Tang, Gang Jin, Hui Jiang, Xu Fang, Isabella Nogues, Xuezhou Li, Wenchao Guo, Yu Wang, Wei Fang, Mingyan Qiu, Yang Hou, Tomas Kovarnik, Michal Vocka, Yimei Lu, Yingli Chen, Xin Chen, Zaiyi Liu, Jian Zhou, Chuanmiao Xie, Rong Zhang, Hong Lu, Gregory D. Hager, Alan L. Yuille, Le Lu, Chengwei Shao, Yu Shi, Qi Zhang, Tingbo Liang, Ling Zhang & Jianping Lu

Nature Medicine **29**, 3033–3043 (2023)


a
PANDA
Lesion versus normal
Detection
Non-contrast CT
PDAC
PNET
SPT
IPMN
MCN
CP
SCN
Other
Diagnosis
Supervisions (training only)
Label
Pathology lesion subtypes
Contrast CT
Annotation
Lesion and pancreas masks
Training set n = 3,208 from one internal center
(1,431 PDAC, 839 non-PDAC, 938 normal)
b
Internal test cohort n = 291 and addition cohort n = 611
(475 PDAC, 311 non-PDAC, 116 normal)
Reader studies
33 readers on non-contrast CT
Assist
Compare
PANDA
on non-contrast CT
Compare
15 readers on contrast CT
Multicenter external test n = 5,337 from 9 centers
(2,737 PDAC, 932 non-PDAC, 1,668 normal)
Generalization to chest CT n = 492
(63 PDAC, 51 non-PDAC, 378 normal)
CLINIC
ER
Real-world multi-scenario studies n = 20,530
(chictr.org.cn, ChiCTR2200064645)
c
Physical exam
(lung cancer screening)
ER
Emergency
CLINIC
Outpatient
Inpatient
Chest
non-contrast CT
Abdominal
non-contrast CT
PANDA
1) High sensitivity and exceptional specificity for consecutive real-world patients
2) Detect malignancies missed by the standard of care
3) Detect malignancies at the surgically resectable stage
4) Safe (one false positive among 1,000 tests) and efficient
5) Model evolution to better align with bedside clinical needs


a
Real-world non-contrast CT
Physical exam
Emergency
Outpatient
Inpatient
CT scans (n = 16,420) initial SOC
RW1
PANDA evaluation study
ChiCTR2200064645
Time (months)
Some patients with follow-up SOC
AI findings not reported by SOC: MDT
CT scans (n = 4,110) initial SOC
RW2
PANDA plus evaluation study
ChiCTR2200064645
Time (months)
Some patients with follow-up SOC
AI findings not reported by SOC: MDT
b
RW1: lesion detection
Sensitivity
PPV
PPV*
Specificity
Specificity*
Total (n = 16,420)
Physical (n = 9,429)
Emergency (n = 3,027)
Outpatient (n = 2,311)
Inpatient (n = 1,653)
c
RW1: PDAC identification
d
RW1: stratified by lesion type
RW2: stratified by lesion type
e
RW2: lesion detection
Total (n = 4,110)
Physical (n = 1,854)
Emergency (n = 969)
Outpatient (n = 688)
Inpatient (n = 599)
f
RW2: PDAC identification
g
Improvement of model evolution on RW2
Pre-upgrade lesion detection
Upgraded lesion detection
Pre-upgrade PDAC identification
Upgraded PDAC identification
h
Aneurysm
Metastatic lymph node
Dilated CBD
Dilated pancreatic duct
Peri-pancreatic tumor
Kidney/liver cyst
Others (thrombosis, ...)
Polycystic liver/kidney
Duodenal diverticulitis
RW1
RW2
Peri-pancreatic tumor
Aneurysm
Metastatic lymph node
Dilated pancreatic duct
Dilated CBD
Liver/kidney cysts
Polycystic liver/kidney
Duodenal diverticulitis
Others
i
Initial SOC missed
Month 0
PANDA detected
Month 7
Patient treatment
Month 7, recalled for MRI, then surgery
Physical exam chest CT
PANDA
PNET (G1)
Contrast MRI
Month 0
Month 7
Month 1, diagnosed on follow-up SOC, then surgery
Outpatient: Lung nodule on annual CT
PDAC (T2)
Contrast CT
Month 0
Month 7
Month 6, diagnosed on follow-up SOC, then interventional therapy
Physical exam chest CT
Aneurysm
Contrast CT
Normal
PDAC
Non-PDAC

Real World Evidence and Causal Inference

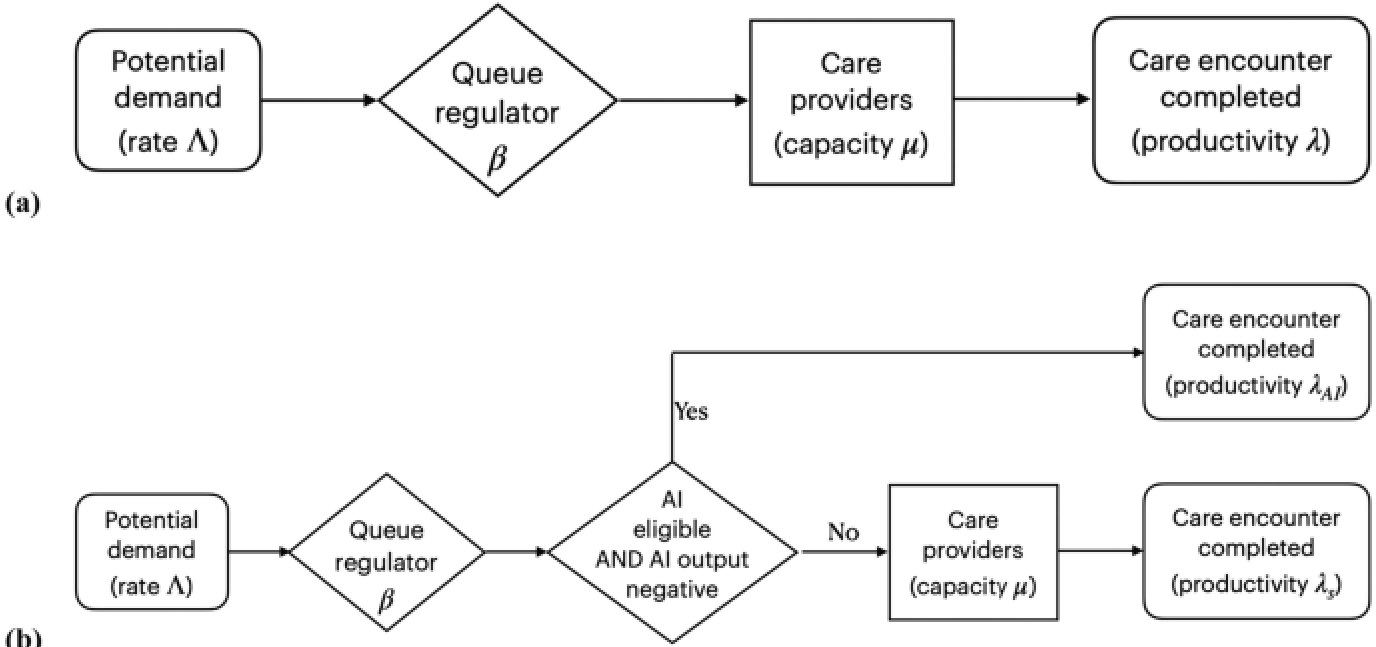
Article | Open access | Published: 04 October 2023

Autonomous artificial intelligence increases real-world specialist clinic productivity in a cluster-randomized trial

Michael D. Abramoff, Noelle Whitestone, Jennifer L. Patnaik, Emily Rich, Munir Ahmed, Lutful Husain, Mohammad Yeadul Hassan, Md. Sajidul Huq Tanjil, Dena Weitzman, Tinglong Dai, Brandie D. Wagner, David H. Cherwek, Nathan Congdon & Khairul Islam

npj Digital Medicine **6**, Article number: 184 (2023)

Healthcare productivity model based on rational queueing theory



Productivity outcomes by study group

	Control (95% CI)	Intervention (95% CI)
Completed care encounters among clinic patients with diabetes		
Care encounter involved specialist	920	858
Care encounter completed by AI-only	0	331
Total	920	1189
Total number of specialist hours in clinic	819	763
Clinic productivity (95% CI) for diabetes patients: number of completed care encounters per hour per specialist physician[a]	$\lambda_{d,c} = 1.14$ (1.02, 1.25)	$\lambda_{d,AI} = 1.59$ (1.3, 1.80)
Clinic productivity (95% CI) for all patients number of completed care encounters per hour per specialist physician[b]	$\lambda_{c} = 3.36$ (3.08, 3.63)	$\lambda_{AI} = 4.05$ (3.66, 4.43)
Specialist productivity adjusted for patient complexity for diabetes patients	$\lambda_{ca,d,c} = 1.19$	$\lambda_{caAI} = 3.15$

Multi-model and Generative AI

slido

What data modalities do you think of when developing or using multi-modal AI?

ⓘ Start presenting to display the poll results on this slide.

Multi-modal AI and Generative AI

nature medicine

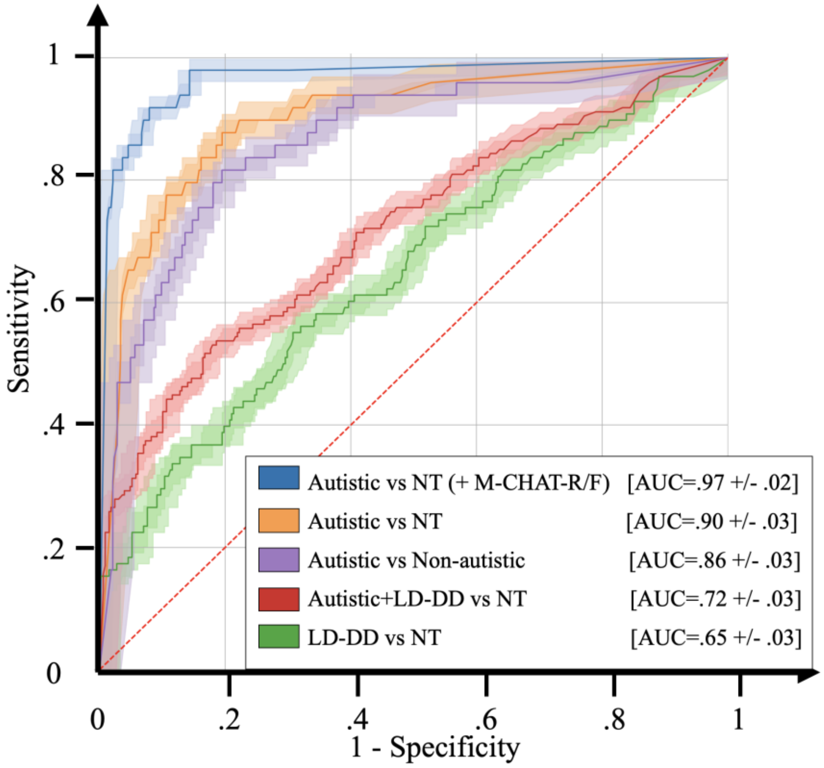
Article | Open access | Published: 02 October 2023

Early detection of autism using digital behavioral phenotyping

Sam Perochon, J. Matias Di Martino, Kimberly L. H. Carpenter, Scott Compton, Naomi Davis, Brian Eichner, Steven Espinosa, Lauren Franz, Pradeep Raj Krishnappa Babu, Guillermo Sapiro & Geraldine Dawson

Nature Medicine **29**, 2489–2497 (2023)

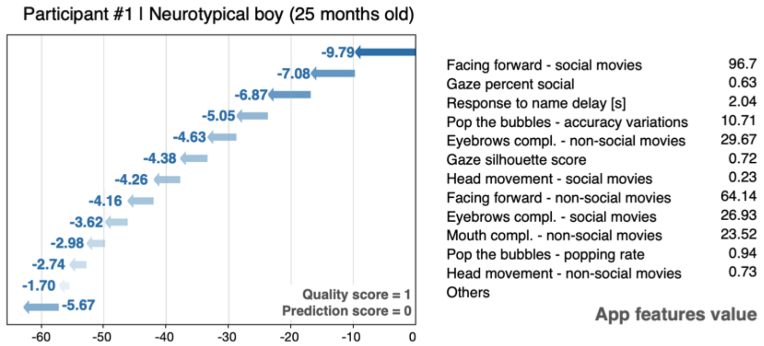
SenseToKnow's AI-based algorithm combines 23 digital phenotypes for detection of autism

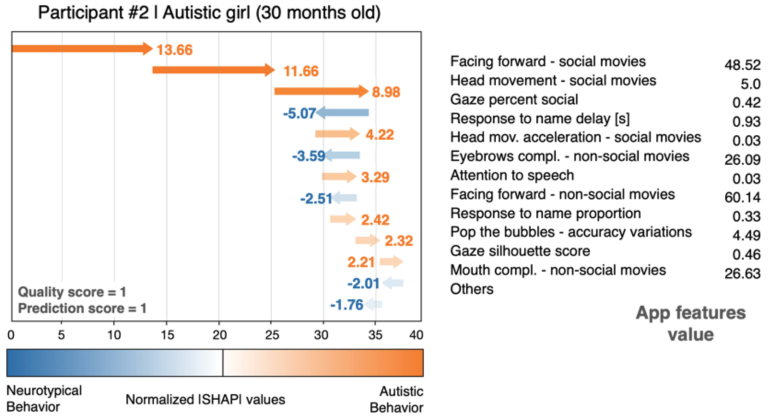


- App was administered during well-child visit in primary care clinics to 475 17–36-month-old toddlers (49 autism; 98 developmental delay)
- Automatically assesses quality of the app administration
- Sensitivity = 87.9%
  Specificity = 80.8%
- No difference in sensitivity and PPV for children of different sex, race, or ethnicity

Slides courtesy of Geraldine Dawson

App generates each child's unique behavioral profile





- Prediction is accompanied by a rating of the confidence level of the prediction
- App generates individualized digital behavioral profile for each child
- Can be used for early intervention planning

Slides courtesy of Geraldine Dawson

Towards Unifying Medical Vision-and-Language Pre-training via Soft Prompts

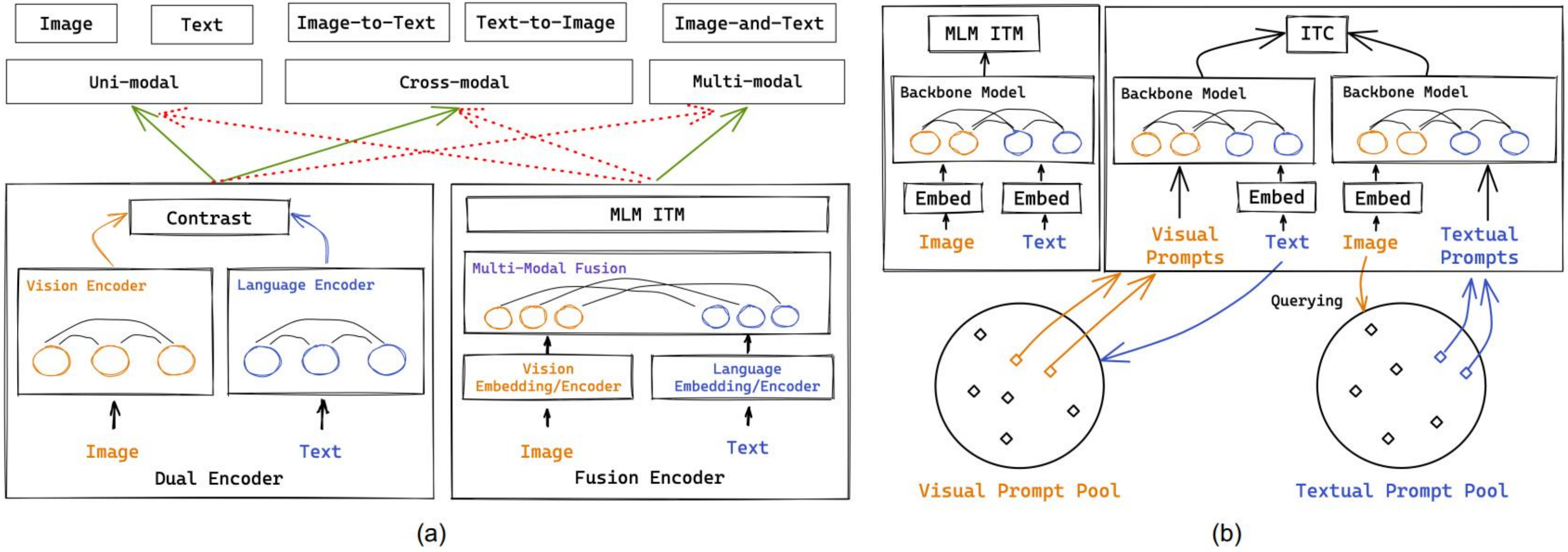
Zhihong Chen[1,2*] Shizhe Diao[3*]
Benyou Wang[1,2†] Guanbin Li[4†] Xiang Wan[2]

[1]The Chinese University of Hong Kong, Shenzhen [2]Shenzhen Research Institute of Big Data
[3]The Hong Kong University of Science and Technology [4]Sun Yat-sen University

zhihongchen@link.cuhk.edu.cn sdiaoaa@connect.ust.hk
wangbenyou@cuhk.edu.cn liguanbin@mail.sysu.edu.cn wanxiang@sribd.com

International Conference on Computer Vision (ICCV 2023)

PTUnifier unifies fusion-encoders & dual-encoders for VLP



(a) (b)

PTUnifier works in a task- and architecture-agnostic manner robust to missing modalities

Methods	Uni-Modal				Cross-Modal			Multi-Modal		
	Image		Text		Image-to-Text		Text-to-Image			
	CheXpert AUROC	PNAS AUROC	RadNLI Acc	MIMIC RL	MIMIC BL4	ROCO R@1	ROCO R@1	VQA-RAD Acc	SLAKE Acc	MedVQA-2019 Acc
$Study_1$	ConVIRT [65]		ClinicalBERT [2]	TransABS [37]	R2Gen [11]	ViLT [26]		CPRD [34]		
	87.3	81.3	72.6	43.8	8.0	11.9	9.8	72.7	82.1	-
$Study_2$	GLoRIA [20]		IFCC [42]	WGSum [19]	M2Trans [42]	METER [17]		MMBERT [25]		
	88.1	88.6	77.8	45.1	10.5	14.5	11.3	72.0	-	77.9
PTUnifier (ours)	90.1	90.6	80.0	46.2	10.7	21.0	20.8	78.3	85.2	79.3

Multi-modal AI and Generative AI

Med-Flamingo: a Multimodal Medical Few-shot Learner

Michael Moor* CORRESPONDENCE TO: MDMOOR@CS.STANFORD.EDU
Qian Huang*
Shirley Wu
Michihiro Yasunaga
Yash Dalmia
Jure Leskovec†
Department of Computer Science, Stanford University, Stanford, USA

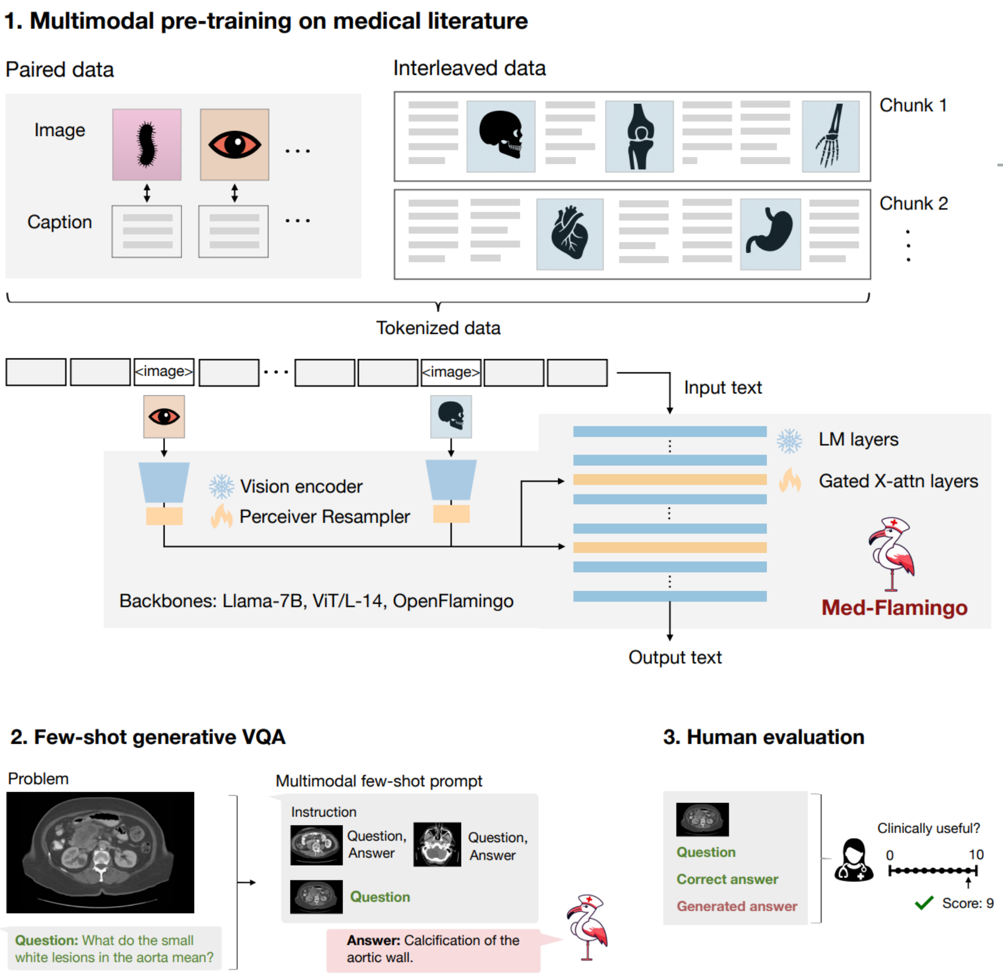
Cyril Zakka
Department of Cardiothoracic Surgery, Stanford Medicine, Stanford, USA

Eduardo Pontes Reis
Hospital Israelita Albert Einstein, São Paulo, Brazil

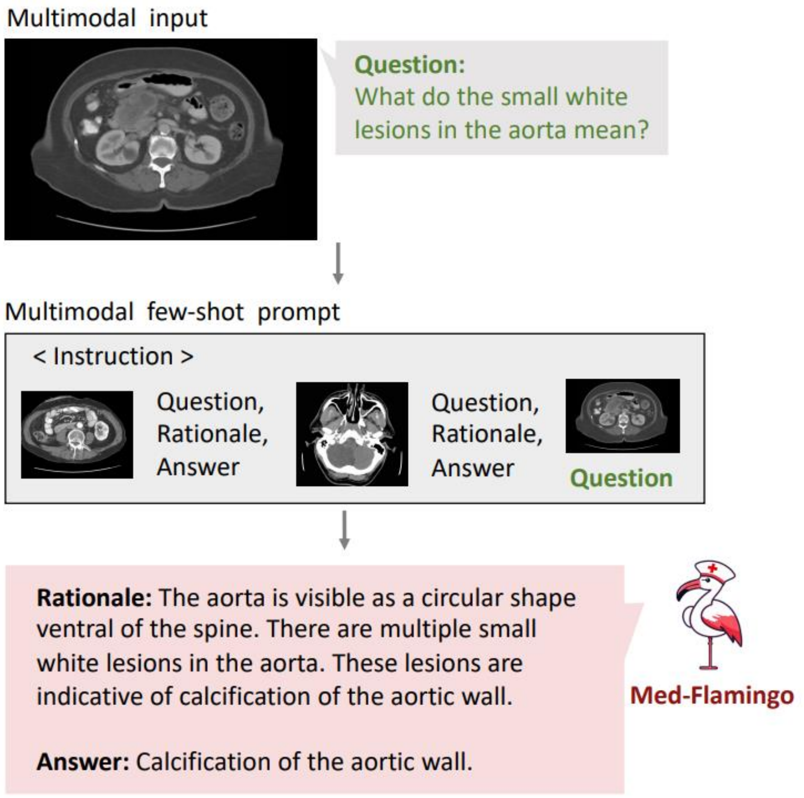
Pranav Rajpurkar
Department of Biomedical Informatics, Harvard Medical School, Boston, USA

Proceedings of Machine Learning Research 225:353–367, 2023

Machine Learning for Health (ML4H) 2023



INFORMATICS PROFESSIONALS. LEADING THE WAY.



Med-Flamingo performance

Dataset	Model	Clinical eval. score	BERT-sim	Exact-match
VQA-RAD	MedVINT zero-shot	4.63	0.628	0.167
	MedVINT fine-tuned (∼$2K$ samples)	2.87	0.611	0.133
	OpenFlamingo zero-shot	4.39	0.490	0.000
	OpenFlamingo few-shot	<u>4.69</u>	<u>0.645</u>	**0.200**
	Med-Flamingo zero-shot	3.82	0.480	0.000
	Med-Flamingo few-shot	**5.61**	**0.650**	**0.200**
Path-VQA	MedVINT zero-shot	0.13	0.608	0.272
	MedVINT fine-tuned (∼$20K$ samples)	1.23	**0.723**	**0.385**
	OpenFlamingo zero-shot	**2.16**	0.474	0.009
	OpenFlamingo few-shot	<u>2.08</u>	0.669	0.288
	Med-Flamingo zero-shot	1.72	0.521	0.120
	Med-Flamingo few-shot	1.81	<u>0.678</u>	<u>0.303</u>
Visual USMLE	MedVINT zero-shot	0.41	0.421	-
	OpenFlamingo zero-shot	<u>4.31</u>	**0.512**	-
	OpenFlamingo few-shot	3.39	0.470	-
	Med-Flamingo zero-shot	4.18	<u>0.473</u>	-
	Med-Flamingo few-shot	**4.33**	0.431	-

Multi-modal AI and Generative AI

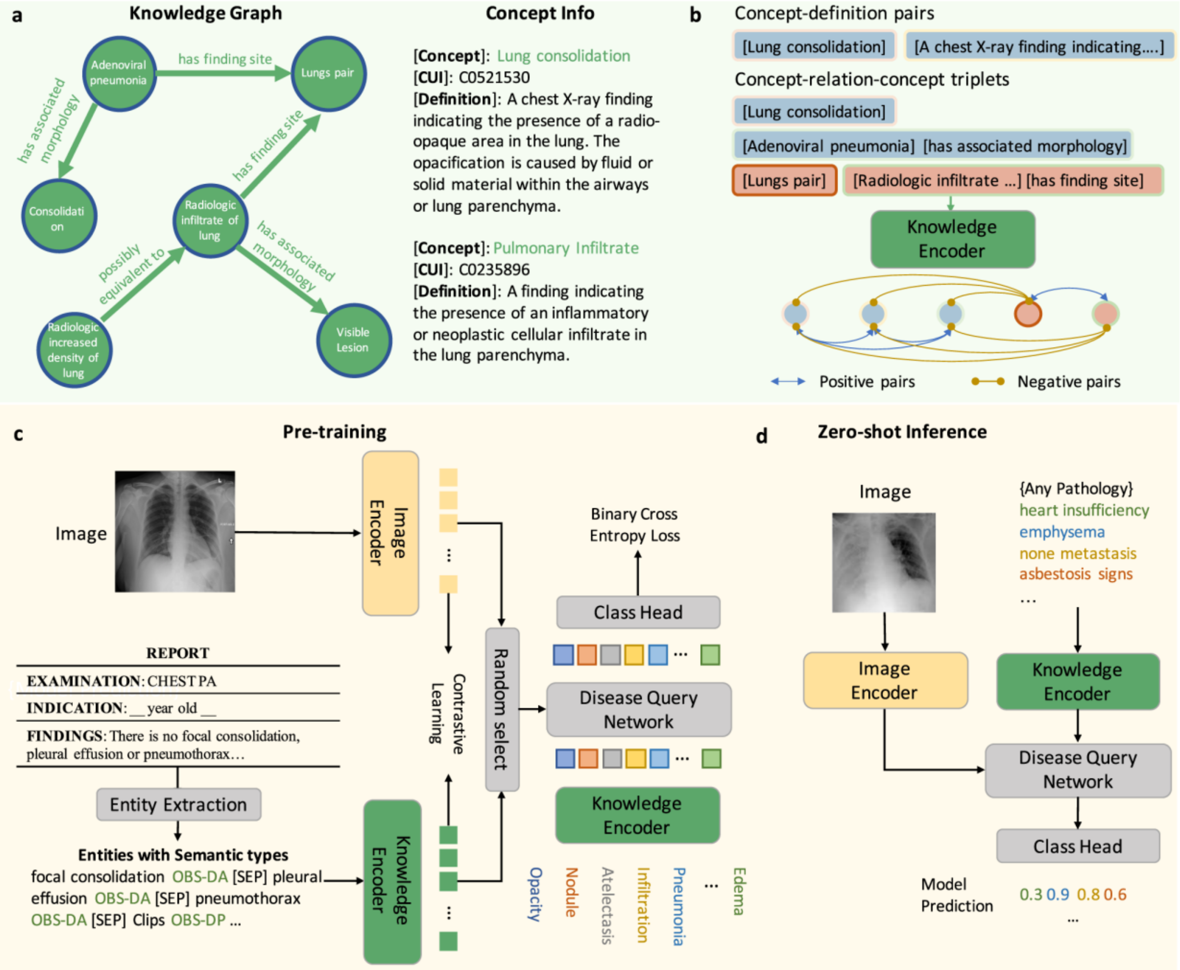
nature communications

Article | Open access | Published: 28 July 2023

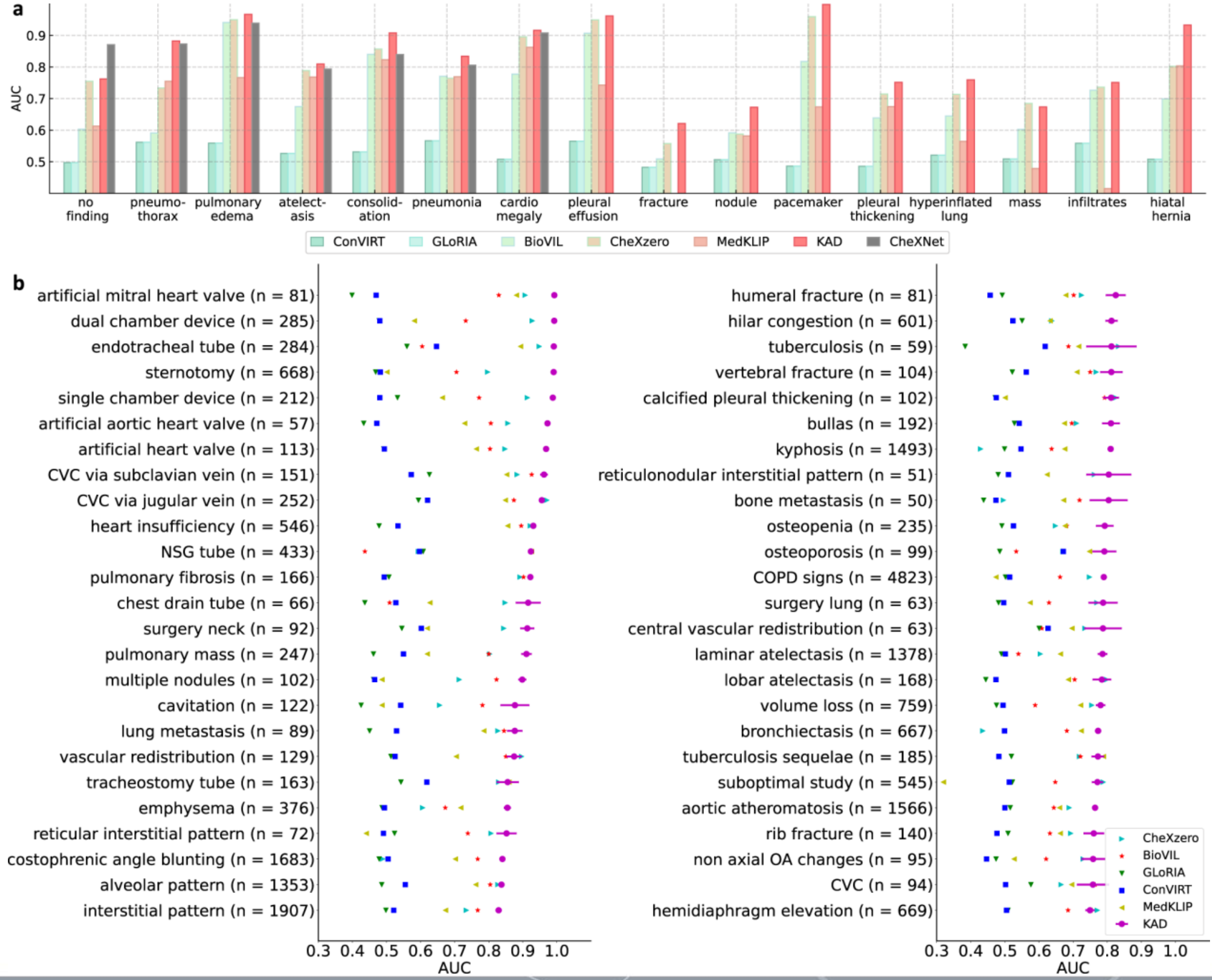
Knowledge-enhanced visual-language pre-training on chest radiology images

Xiaoman Zhang, Chaoyi Wu, Ya Zhang, Weidi Xie & Yanfeng Wang

Nature Communications **14**, Article number: 4542 (2023)



a
Knowledge Graph
Adenoviral pneumonia
has finding site
Lungs pair
has associated morphology
Consolidation
has finding site
Radiologic infiltrate of lung
possibly equivalent to
has associated morphology
Radiologic increased density of lung
Visible Lesion
Concept Info
[Concept]: Lung consolidation
[CUI]: C0521530
[Definition]: A chest X-ray finding indicating the presence of a radio-opaque area in the lung. The opacification is caused by fluid or solid material within the airways or lung parenchyma.
[Concept]: Pulmonary Infiltrate
[CUI]: C0235896
[Definition]: A finding indicating the presence of an inflammatory or neoplastic cellular infiltrate in the lung parenchyma.
b
Concept-definition pairs
[Lung consolidation]
[A chest X-ray finding indicating....]
Concept-relation-concept triplets
[Lung consolidation]
[Adenoviral pneumonia] [has associated morphology]
[Lungs pair]
[Radiologic infiltrate ...] [has finding site]
Knowledge Encoder
Positive pairs
Negative pairs
c
Pre-training
Image
Image Encoder
REPORT
EXAMINATION: CHEST PA
INDICATION: __ year old __
FINDINGS: There is no focal consolidation, pleural effusion or pneumothorax…
Entity Extraction
Entities with Semantic types
focal consolidation OBS-DA [SEP] pleural effusion OBS-DA [SEP] pneumothorax OBS-DA [SEP] Clips OBS-DP ...
Knowledge Encoder
Contrastive Learning
Random select
Binary Cross Entropy Loss
Class Head
Disease Query Network
Knowledge Encoder
Opacity
Nodule
Atelectasis
Infiltration
Pneumonia
...
Edema
d
Zero-shot Inference
Image
{Any Pathology}
heart insufficiency
emphysema
none metastasis
asbestosis signs
...
Image Encoder
Knowledge Encoder
Disease Query Network
Class Head
Model Prediction
0.3 0.9 0.8 0.6
...



a
AUC
no finding
pneumothorax
pulmonary edema
atelectasis
consolidation
pneumonia
cardio megaly
pleural effusion
fracture
nodule
pacemaker
pleural thickening
hyperinflated lung
mass
infiltrates
hiatal hernia
ConVIRT
GLoRIA
BioViL
CheXzero
MedKLIP
KAD
CheXNet
b
artificial mitral heart valve (n = 81)
dual chamber device (n = 285)
endotracheal tube (n = 284)
sternotomy (n = 668)
single chamber device (n = 212)
artificial aortic heart valve (n = 57)
artificial heart valve (n = 113)
CVC via subclavian vein (n = 151)
CVC via jugular vein (n = 252)
heart insufficiency (n = 546)
NSG tube (n = 433)
pulmonary fibrosis (n = 166)
chest drain tube (n = 66)
surgery neck (n = 92)
pulmonary mass (n = 247)
multiple nodules (n = 102)
cavitation (n = 122)
lung metastasis (n = 89)
vascular redistribution (n = 129)
tracheostomy tube (n = 163)
emphysema (n = 376)
reticular interstitial pattern (n = 72)
costophrenic angle blunting (n = 1683)
alveolar pattern (n = 1353)
interstitial pattern (n = 1907)
humeral fracture (n = 81)
hilar congestion (n = 601)
tuberculosis (n = 59)
vertebral fracture (n = 104)
calcified pleural thickening (n = 102)
bullas (n = 192)
kyphosis (n = 1493)
reticulonodular interstitial pattern (n = 51)
bone metastasis (n = 50)
osteopenia (n = 235)
osteoporosis (n = 99)
COPD signs (n = 4823)
surgery lung (n = 63)
central vascular redistribution (n = 63)
laminar atelectasis (n = 1378)
lobar atelectasis (n = 168)
volume loss (n = 759)
bronchiectasis (n = 667)
tuberculosis sequelae (n = 185)
suboptimal study (n = 545)
aortic atheromatosis (n = 1566)
rib fracture (n = 140)
non axial OA changes (n = 95)
CVC (n = 94)
hemidiaphragm elevation (n = 669)
CheXzero
BioViL
GLoRIA
ConVIRT
MedKLIP
KAD
AUC

Multi-modal AI and Generative AI

nature

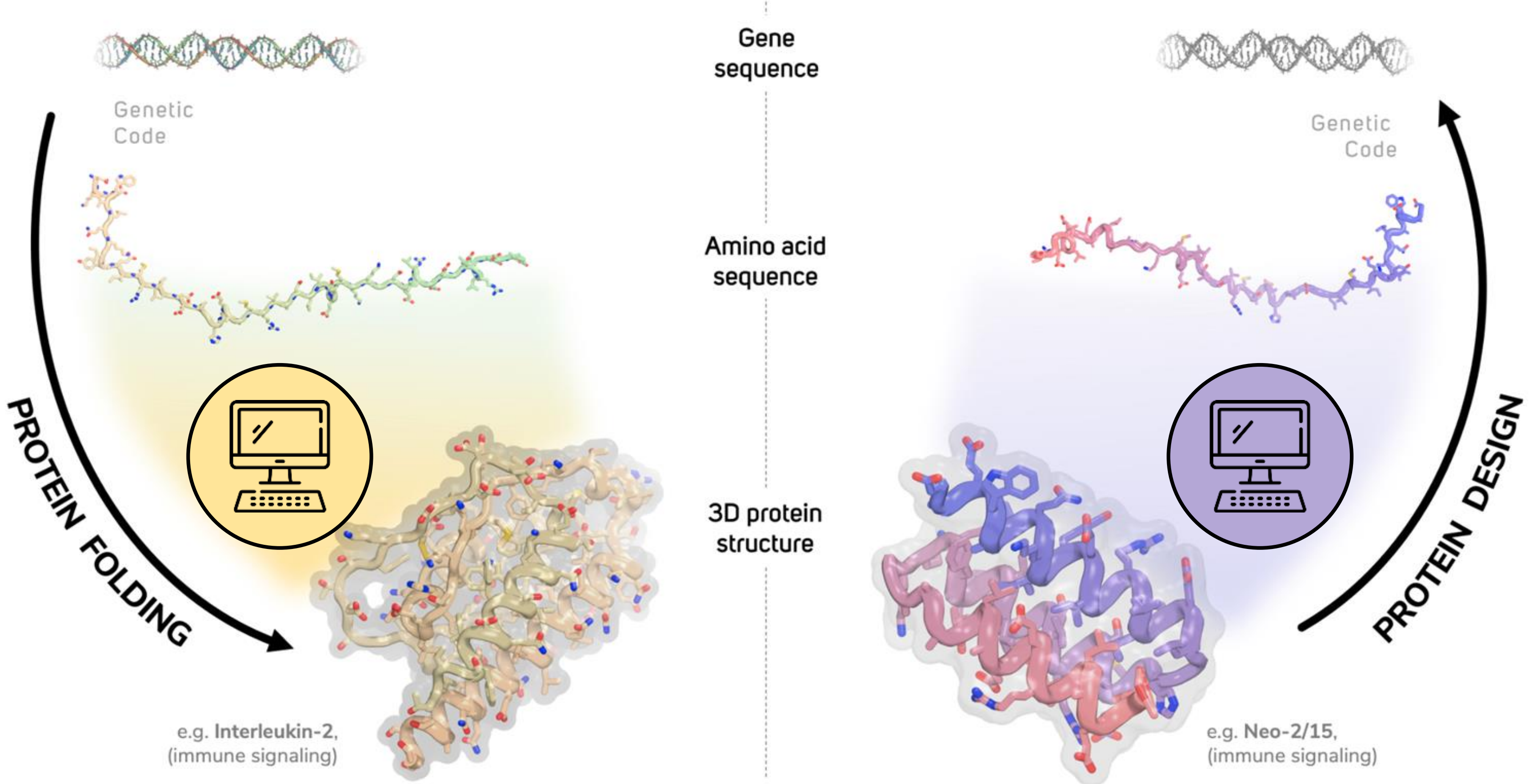
Article | Open access | Published: 11 July 2023

De novo design of protein structure and function with RFdiffusion

Joseph L. Watson, David Juergens, Nathaniel R. Bennett, Brian L. Trippe, Jason Yim, Helen E. Eisenach, Woody Ahern, Andrew J. Borst, Robert J. Ragotte, Lukas F. Milles, Basile I. M. Wicky, Nikita Hanikel, Samuel J. Pellock, Alexis Courbet, William Sheffler, Jue Wang, Preetham Venkatesh, Isaac Sappington, Susana Vázquez Torres, Anna Lauko, Valentin De Bortoli, Emile Mathieu, Sergey Ovchinnikov, Regina Barzilay, ... David Baker

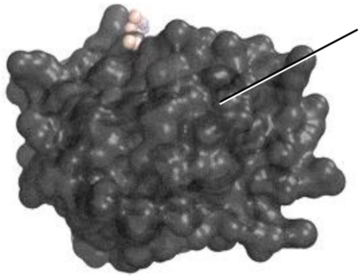
Nature **620**, 1089–1100 (2023)

Slides courtesy of David Baker



RFdiffusion can be used to generate useful proteins in seconds

Slides courtesy of David Baker



The New York Times

A.I. Turns Its Artistry to Creating New Human Proteins

Inspired by digital art generators like DALL-E, biologists are building artificial intelligences that can fight cancer, flu and Covid.

Jan 9, 2023

THE WALL STREET JOURNAL.

Biologists Say Deep Learning Is Revolutionizing Pace of Innovation

A combination of new algorithmic and experimental methods has accelerated research in protein design by a factor of 10.

Mar 22, 2023

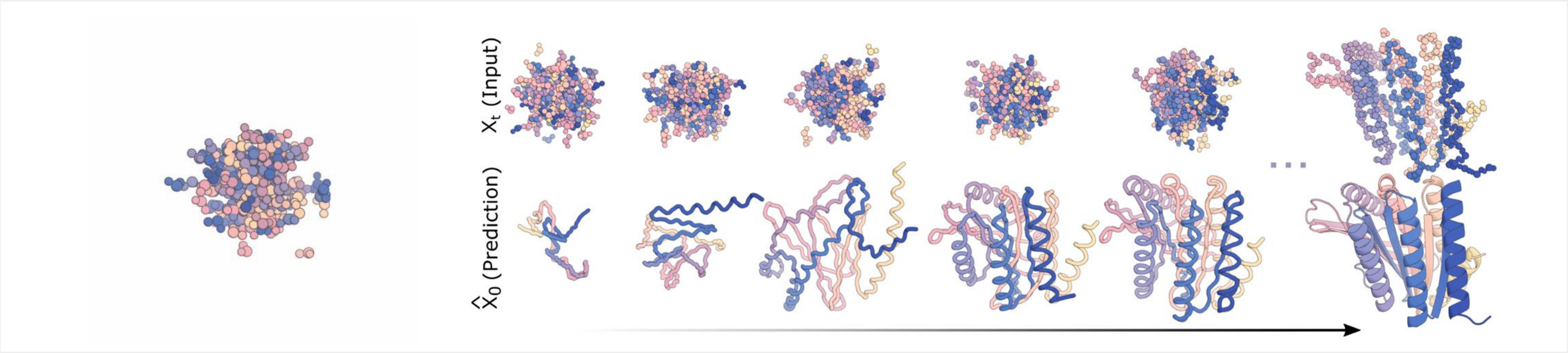
MIT Technology Review

An AI that can design new proteins could help unlock new cures and materials

The machine-learning tool could help researchers discover entirely new proteins not yet known to science.

Sep 18, 2022

RFdiffusion generates new protein structures via progressive denoising

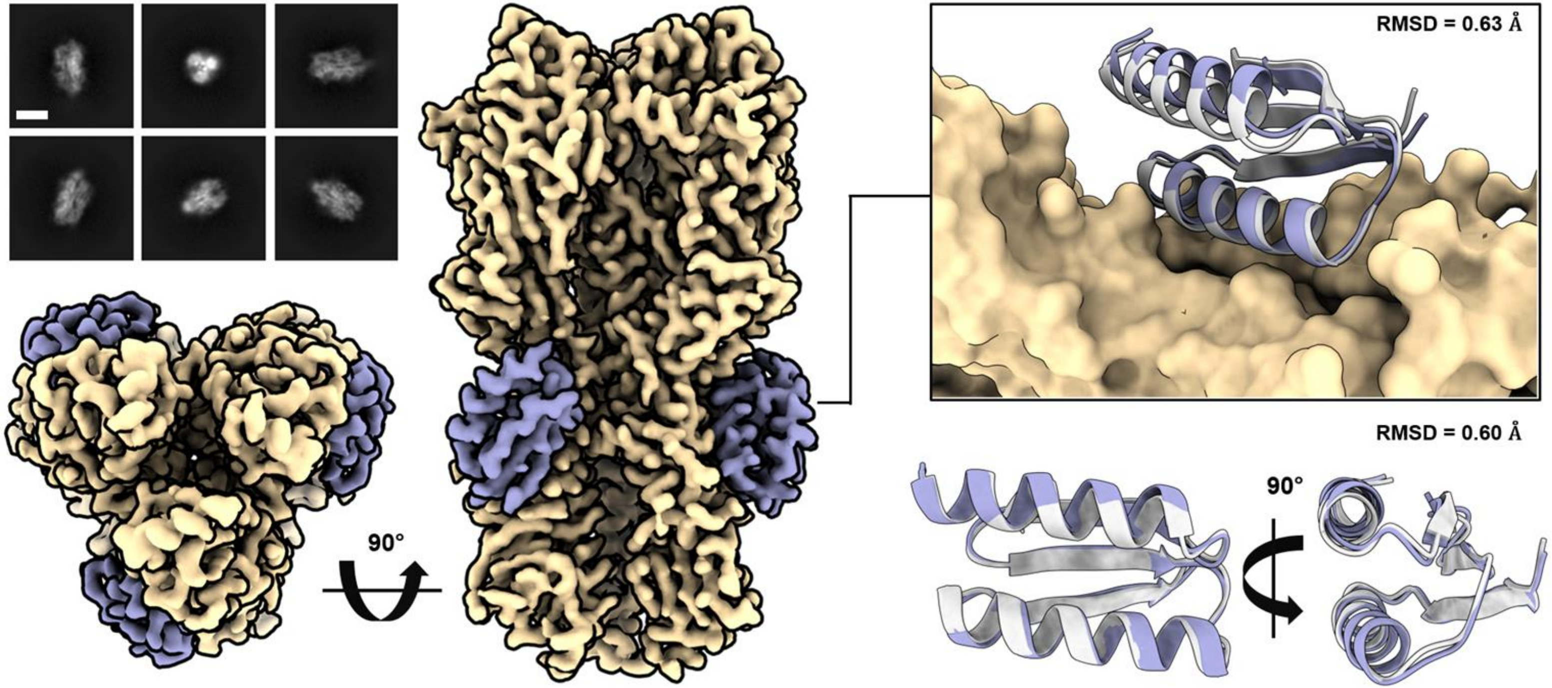


Inspired by deep-learning methods for generating synthetic images.
e.g. DALL-E

Synthetic image trajectory from NVIDIA

Cryo-EM characterization of a diffused influenza-binding protein

2.9 Å resolution



Multi-modal AI and Generative AI

npj | digital medicine

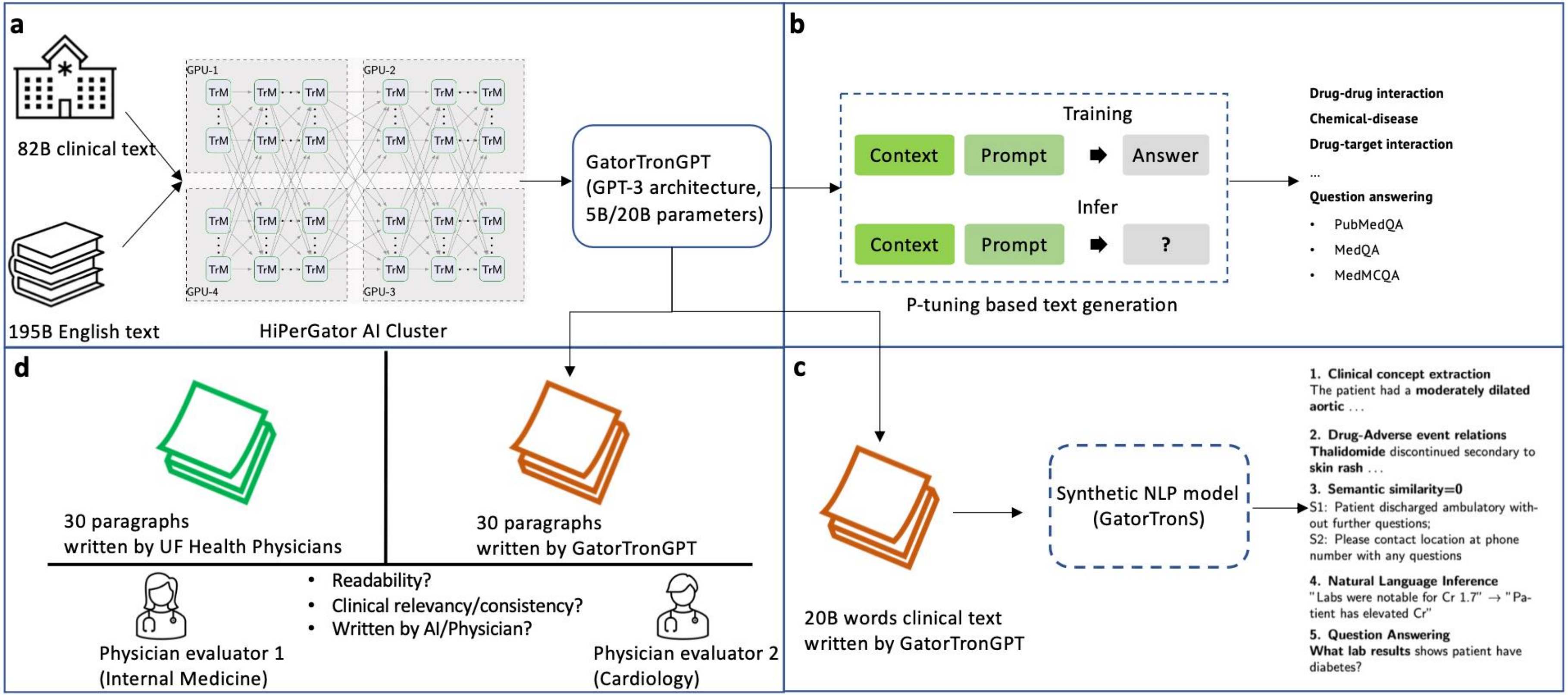
Article | Open access | Published: 16 November 2023

A study of generative large language model for medical research and healthcare

Cheng Peng, Xi Yang, Aokun Chen, Kaleb E. Smith, Nima PourNejatian, Anthony B. Costa, Cheryl Martin, Mona G. Flores, Ying Zhang, Tanja Magoc, Gloria Lipori, Duane A. Mitchell, Naykky S. Ospina, Mustafa M. Ahmed, William R. Hogan, Elizabeth A. Shenkman, Yi Guo, Jiang Bian & Yonghui Wu

npj Digital Medicine **6**, Article number: 210 (2023)

GatorTronGPT - A generative LLM for EHRs



Turing test results

- No significant difference in linguistic readability (p = 0·22; 6·57 of GatorTronGPT compared with 6·93 of human)
- No significant difference in clinical relevance (p = 0·91; 7·0 of GatorTronGPT compared with 6·97 of human)
- Physicians cannot differentiate them (p < 0·001)

a. Percentage of notes correctly identified by human reviewers.

	Ground truth		
	AI (n=30)	**Human** (n=30)	**Total** (n=60)
Physician 1	9 (30·0%)	17 (56·7%)	26 (43·3%)
Physician 2	13 (43·3%)	20 (66·7%)	33 (55%)
Overall	11 (36·7%)	18·5 (61·7%)	29.5 (49·2%)
p-value	< 0·001	= 0·104	< 0·001

b. Ratings of readability and clinical relevance by synthetic and human notes.

	Ground truth		
	AI (n=30)	**Human** (n=30)	**p-value**
Readability	6.57 (1·22)	6.93 (1·09)	0·22
Clinical relevance	7.00 (1·23)	6.97 (1·07)	0·91

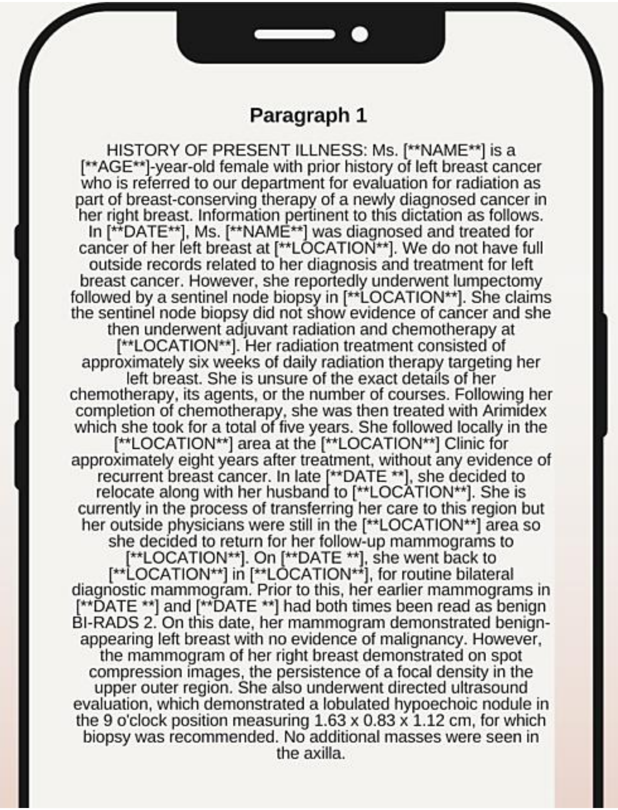
c.

Example 1 (Passed Turing test): HISTORY OF PRESENT ILLNESS: Ms. [**NAME**] is a [**AGE**]-year-old female with prior history of left breast cancer who is referred to the [**LOCATION**] Clinic to discuss surgical management of her left lung nodule which was picked up on CT scanning for surveillance. The patient's chest CT obtained at an outside institution showed an incidental finding of a left lower lobe pleural nodule, which she brought with her to clinic today for further review. Ms. [**NAME**] reports to be doing well, with no change in her baseline level of shortness of breath, but denies any difficulty breathing, cough or hemoptysis. She feels essentially back to baseline and plays golf one day a week. Otherwise, Ms. [**NAME **] has no other pulmonary complaints and has never smoked tobacco or inhaled second hand smoke. She does report several years of intermittent reflux, which worsened after eating rough foods. She underwent an upper GI study for this, which showed a small hiatal hernia as well as a moderate- sized sliding hiatal hernia. She subsequently underwent an abdominal CT for evaluation for mesenteric ischemia and was felt to have a dilated pancreatic duct. Due to suspected head pathology on imaging (although the final report is pending), she underwent an MRI, which confirmed no pancreatic pathology. Given this, Ms. [**NAME **] has been trying to avoid fatty and fried foods, which has improved her reflux. Ms. [**NAME **] does report recent onset of sleep apnea, for which she is using CPAP every night for the last month. She tolerates this well without complaints of dry mouth, gasping for air, or undue sleepiness. Ms. [**NAME **] was seen in our outpatient [**LOCATION**] clinic on [**DATE**] at which time it was determined that in order to obtain a tissue diagnosis of this left lower lobe lesion a VATS approach would be utilized. She presents today for further pre-operative planning. Currently, Ms. [**DATE**] denies any pulmonary symptoms, fevers, chills, cough or weight loss. PAST MEDICAL HISTORY: 1. CARDIAC RISK FACTORS: Hypertension, Hyperlipidemia 2. CARDIAC HISTORY: None 3. OTHER PAST MEDICAL HISTORY: Left lower lobe pulmonary nodule

Example 2 (Failed Turing test): Assessment: ICD-9-CM 1. Pain in joint, pelvic region and thigh, unspecified laterality 719.45 fentaNYL (DURAGESIC) 75 MCG/HR HYDROcodone-acetaminophen (VICODIN HP) 5-300 MG Tablet 2. Migraine NOS/not Intrcbl 346.90 SUMAtriptan Succinate 3. Bone Pain Arthritis 333.90 Dexamethasone Sodium Phosphate 4. Bipolar disorder 296.80 traZODone (DESYREL) 100 MG Tablet prazosin (MINIPRESS) 2 MG capsule carvedilol (COREG) 3.125 MG tablet isosorbide mononitrate (IMDUR) 30 MG CR tablet Refer to Psychiatry clopidogrel (PLAVIX) 75 MG tablet SUMAtriptan Succinate 5. ASTHMA UNSPECIFIED 493.90 albuterol (PROAIR HFA;VENTOLIN HFA) 108 (90 BASE) MCG/ACT inhaler 6. Major depressive disorder, single episode, unspecified 296.20 DULoxetine (CYMBALTA) 60 MG capsule Refer to Psychiatry amitriptyline (ELAVIL) 25 MG tablet traZODone (DESYREL) 100 MG Tablet 7. POST-SURGICAL VARICOSE VEINS of LOWER EXTREMITIES 454.9 fentaNYL (DURAGESIC) 75 MCG/HR 8. Other and unspecified hyperlipidemia 272.4 simvastatin (ZOCOR) 40 MG tablet COMPREHENSIVE METABOLIC PANEL 9. PND (post-nasal drip) 784.91 loratadine (CLARITIN) 10 MG tablet 10. Bipolar I disorder, single manic episode, unspecified 296.00 clonazePAM (KlonoPIN) 1 MG tablet Refer to Psychiatry 11. Allergic rhinitis 477.9 loratadine (CLARITIN) 10 MG tablet 12. Grief reaction 309.0 traZODone (DESYREL) 100 MG Tablet 13. Encounter for long-term (current) use of other medications V58.69 methocarbamol (ROBAXIN) 750 MG tablet COMPREHENSIVE METABOLIC PANEL 14. GERD (gastroesophageal reflux disease) 530.81 lansoprazole (PRE

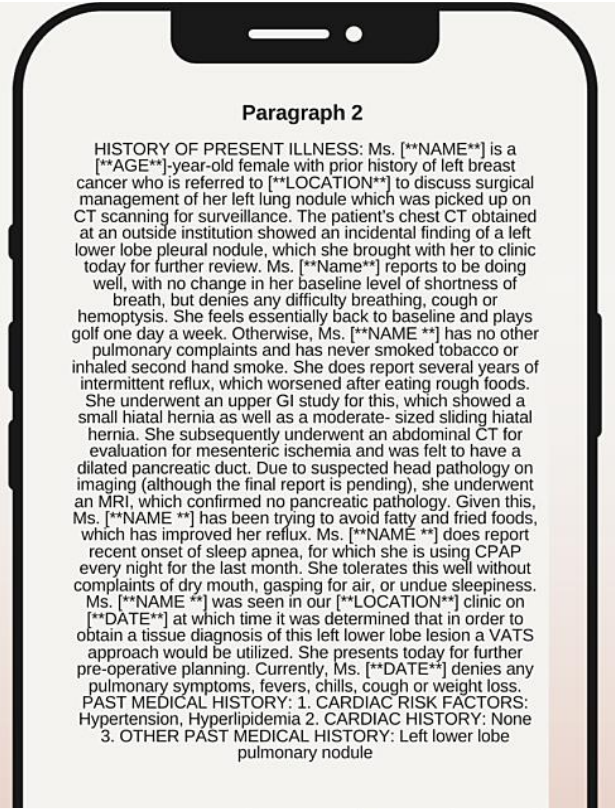
Slides courtesy of Yonghui Wu

Which one is generated by GatorTronGPT?



Paragraph 1

HISTORY OF PRESENT ILLNESS: Ms. [**NAME**] is a [**AGE**]-year-old female with prior history of left breast cancer who is referred to our department for evaluation for radiation as part of breast-conserving therapy of a newly diagnosed cancer in her right breast. Information pertinent to this dictation as follows. In [**DATE**], Ms. [**NAME**] was diagnosed and treated for cancer of her left breast at [**LOCATION**]. We do not have full outside records related to her diagnosis and treatment for left breast cancer. However, she reportedly underwent lumpectomy followed by a sentinel node biopsy in [**LOCATION**]. She claims the sentinel node biopsy did not show evidence of cancer and she then underwent adjuvant radiation and chemotherapy at [**LOCATION**]. Her radiation treatment consisted of approximately six weeks of daily radiation therapy targeting her left breast. She is unsure of the exact details of her chemotherapy, its agents, or the number of courses. Following her completion of chemotherapy, she was then treated with Arimidex which she took for a total of five years. She followed locally in the [**LOCATION**] area at the [**LOCATION**] Clinic for approximately eight years after treatment, without any evidence of recurrent breast cancer. In late [**DATE **], she decided to relocate along with her husband to [**LOCATION**]. She is currently in the process of transferring her care to this region but her outside physicians were still in the [**LOCATION**] area so she decided to return for her follow-up mammograms to [**LOCATION**]. On [**DATE **], she went back to [**LOCATION**] in [**LOCATION**], for routine bilateral diagnostic mammogram. Prior to this, her earlier mammograms in [**DATE **] and [**DATE **] had both times been read as benign BI-RADS 2. On this date, her mammogram demonstrated benign-appearing left breast with no evidence of malignancy. However, the mammogram of her right breast demonstrated on spot compression images, the persistence of a focal density in the upper outer region. She also underwent directed ultrasound evaluation, which demonstrated a lobulated hypoechoic nodule in the 9 o'clock position measuring 1.63 x 0.83 x 1.12 cm, for which biopsy was recommended. No additional masses were seen in the axilla.



Paragraph 2

HISTORY OF PRESENT ILLNESS: Ms. [**NAME**] is a [**AGE**]-year-old female with prior history of left breast cancer who is referred to [**LOCATION**] to discuss surgical management of her left lung nodule which was picked up on CT scanning for surveillance. The patient's chest CT obtained at an outside institution showed an incidental finding of a left lower lobe pleural nodule, which she brought with her to clinic today for further review. Ms. [**Name**] reports to be doing well, with no change in her baseline level of shortness of breath, but denies any difficulty breathing, cough or hemoptysis. She feels essentially back to baseline and plays golf one day a week. Otherwise, Ms. [**NAME **] has no other pulmonary complaints and has never smoked tobacco or inhaled second hand smoke. She does report several years of intermittent reflux, which worsened after eating rough foods. She underwent an upper GI study for this, which showed a small hiatal hernia as well as a moderate- sized sliding hiatal hernia. She subsequently underwent an abdominal CT for evaluation for mesenteric ischemia and was felt to have a dilated pancreatic duct. Due to suspected head pathology on imaging (although the final report is pending), she underwent an MRI, which confirmed no pancreatic pathology. Given this, Ms. [**NAME **] has been trying to avoid fatty and fried foods, which has improved her reflux. Ms. [**NAME **] does report recent onset of sleep apnea, for which she is using CPAP every night for the last month. She tolerates this well without complaints of dry mouth, gasping for air, or undue sleepiness. Ms. [**NAME **] was seen in our [**LOCATION**] clinic on [**DATE**] at which time it was determined that in order to obtain a tissue diagnosis of this left lower lobe lesion a VATS approach would be utilized. She presents today for further pre-operative planning. Currently, Ms. [**DATE**] denies any pulmonary symptoms, fevers, chills, cough or weight loss. PAST MEDICAL HISTORY: 1. CARDIAC RISK FACTORS: Hypertension, Hyperlipidemia 2. CARDIAC HISTORY: None 3. OTHER PAST MEDICAL HISTORY: Left lower lobe pulmonary nodule

Slides courtesy of Yonghui Wu

Multi-modal AI and Generative AI

npj | digital medicine

Article | Open access | Published: 27 May 2023

Generating synthetic mixed-type longitudinal electronic health records for artificial intelligent applications

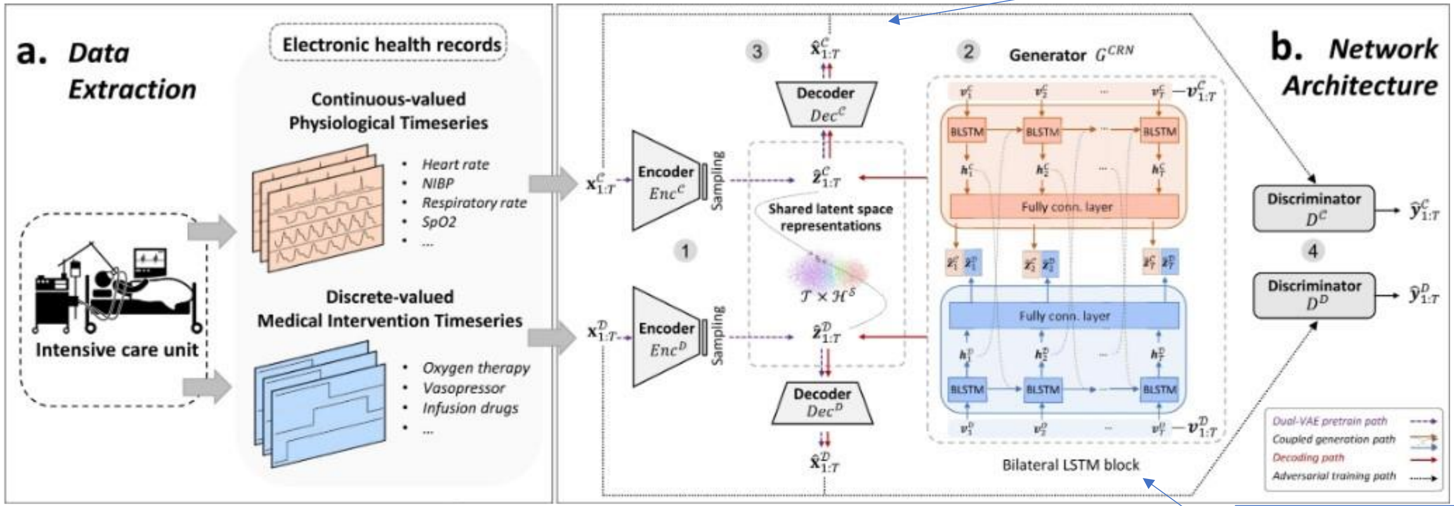
Jin Li, Benjamin J. Cairns, Jingsong Li & Tingting Zhu

npj Digital Medicine **6**, Article number: 98 (2023)

EHR-M-GAN: Synthesizing Mixed-type EHR Data for AI Applications

- Leveraging EHRs for AI research is promising, but patient privacy concerns limit data sharing
- Synthetic data from generative AI models (such as GANs) offers useful alternative
- Current EHR synthesizer limited in data dimensionality and data types

Dual-VAE pretraining for shared latent space representations



Sequentially coupled generator for synthesizing mixed-type timeseries

We propose EHR-M-GAN, a generative adversarial network that synthesizes mixed-type timeseries EHR data mimicking clinical decision-making.

EHR-M-GAN: Synthesizing Mixed-type EHR Data for AI Applications

Slides courtesy of Jingsong Li

- Experimental evaluations demonstrate that EHR-M-GAN is effective in generating **high-fidelity**, **privacy-preserving** clinical timeseries, **improving predictive models** for patient outcomes when used for **augmenting training data**.

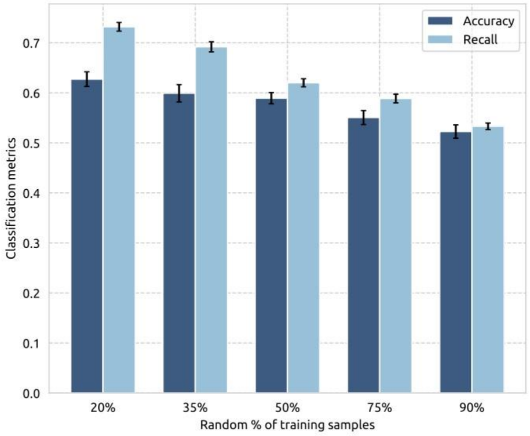


Table 5. Downstream task evaluation with data augmentation ratio α.

Dataset	Treatments	EHR-M-GAN			EHR-M-GAN$_{cond}$		
		$\alpha = 10\%$	$\alpha = 25\%$	$\alpha = 50\%$	$\alpha = 10\%$	$\alpha = 25\%$	$\alpha = 50\%$
MIMIC-III	Vent.	0.828 ± 0.013	0.877 ± 0.014	0.912 ± 0.015 (↑)	0.845 ± 0.022	0.896 ± 0.013 (↑)	**0.923 ± 0.018** (↑)
	Vaso.	0.816 ± 0.015	0.834 ± 0.023	0.859 ± 0.013 (↑)	0.848 ± 0.012 (↑)	0.876 ± 0.017 (↑)	**0.896 ± 0.015** (↑)
eICU	Vent.	0.858 ± 0.008	0.862 ± 0.012	0.873 ± 0.014 (↑)	0.865 ± 0.009	0.879 ± 0.014 (↑)	**0.883 ± 0.016** (↑)
	Vaso.	0.798 ± 0.015	0.805 ± 0.020	0.821 ± 0.028 (↑)	0.813 ± 0.016 (↑)	0.834 ± 0.019 (↑)	**0.839 ± 0.014** (↑)
HiRID	Vent.	0.871 ± 0.025 (↑)	0.882 ± 0.021 (↑)	0.913 ± 0.019 (↑)	0.894 ± 0.015 (↑)	0.906 ± 0.018 (↑)	**0.923 ± 0.021** (↑)
	Vaso.	0.850 ± 0.016	0.874 ± 0.022	0.894 ± 0.018 (↑)	0.883 ± 0.017 (↑)	0.908 ± 0.024 (↑)	**0.913 ± 0.019** (↑)

Downstream tasks are evaluated under the training scenarios of *Train on Synthetic and Real, Test on Real (TSRTR)*. All data from sub-train data A'_{Tr} concated with α of the synthetic data B (augmentation ratio $\alpha = 10\%$, 25% or 50%) is used as the training set. The upper arrow (↑) indicates that the AUROC value under *TSRTR* is higher than *TRTR* in Table 4 for the corresponding task, while the bold arrow (↑) indicates that the value is significantly improved using *t*-test (p≤0.05). Bolded values denote best scores.

Results indicates EHR -M-GAN is sufficiently robust against the **membership inference attack**

Synthetic data from EHR -M-GAN can be used to **boost downstream classifier performance** for predicting clinical interventions

Multi-modal AI and Generative AI

ELSEVIER

Journal of Biomedical Informatics

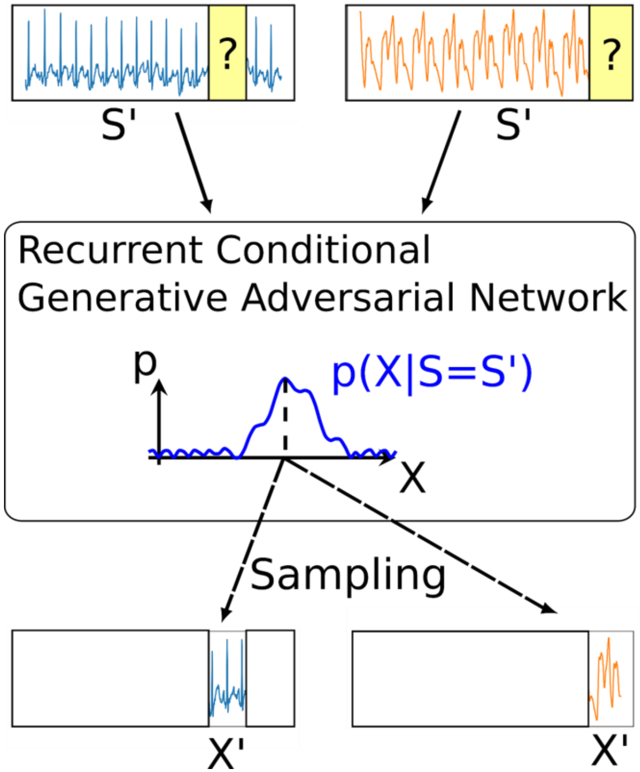
Volume 139, March 2023, 104320

Original Research

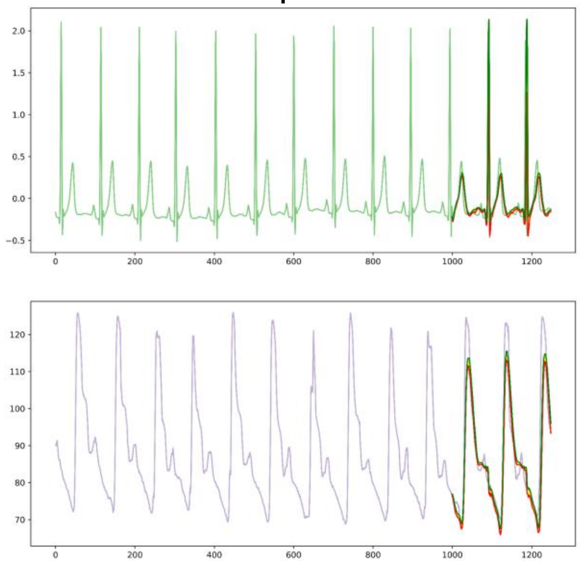
Medical multivariate time series imputation and forecasting based on a recurrent conditional Wasserstein GAN and attention

Sven Festag, Cord Spreckelsen

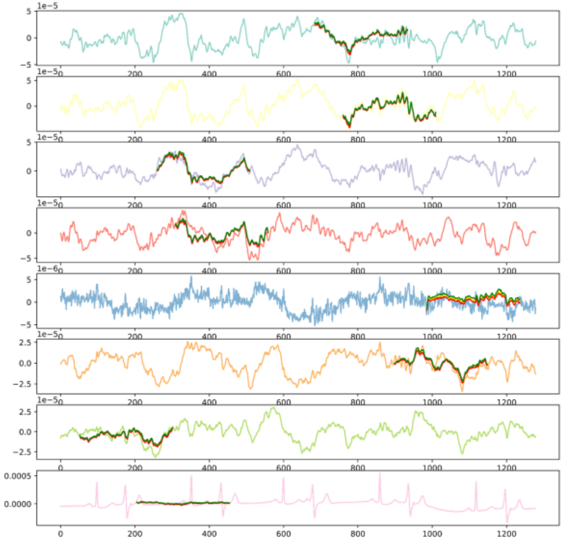
Generative probabilistic system for imputing and forecasting of medical time series



Conditional cross-channel
Forecasting:
Electrocardiogram (ECG) +
Arterial blood pressure

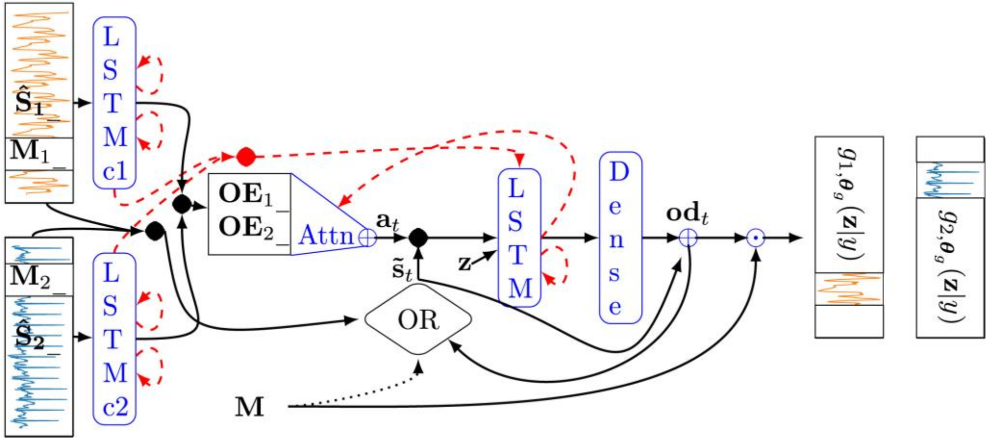


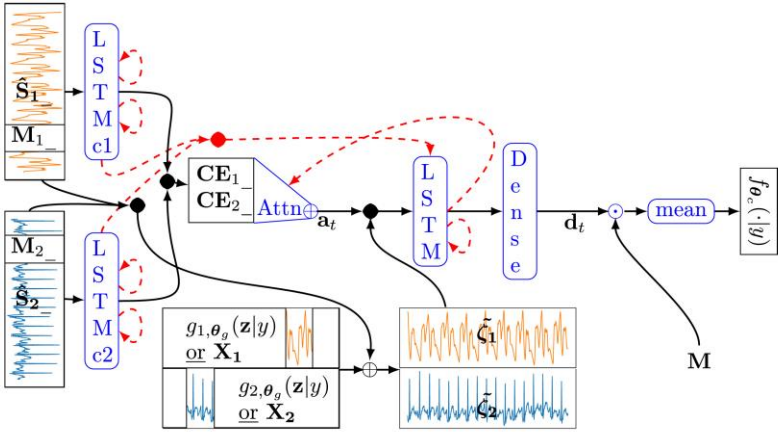
Conditional cross-channel
Imputation:
8-channel Polysomnogram



Slides courtesy of Sven Festag

Generator and critic networks



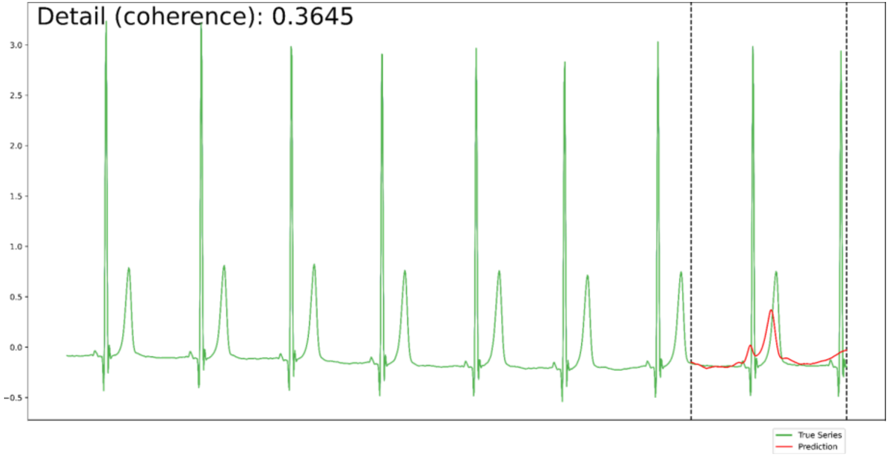


Slides courtesy of Sven Festag

ECG Forecasting

Adversarial training leads to higher detail agreement (Signal Coherence) than direct residual minimisation

No adversarial learning:



Adversarial learning:

Slides courtesy of Sven Festag

Multi-modal AI and Generative AI

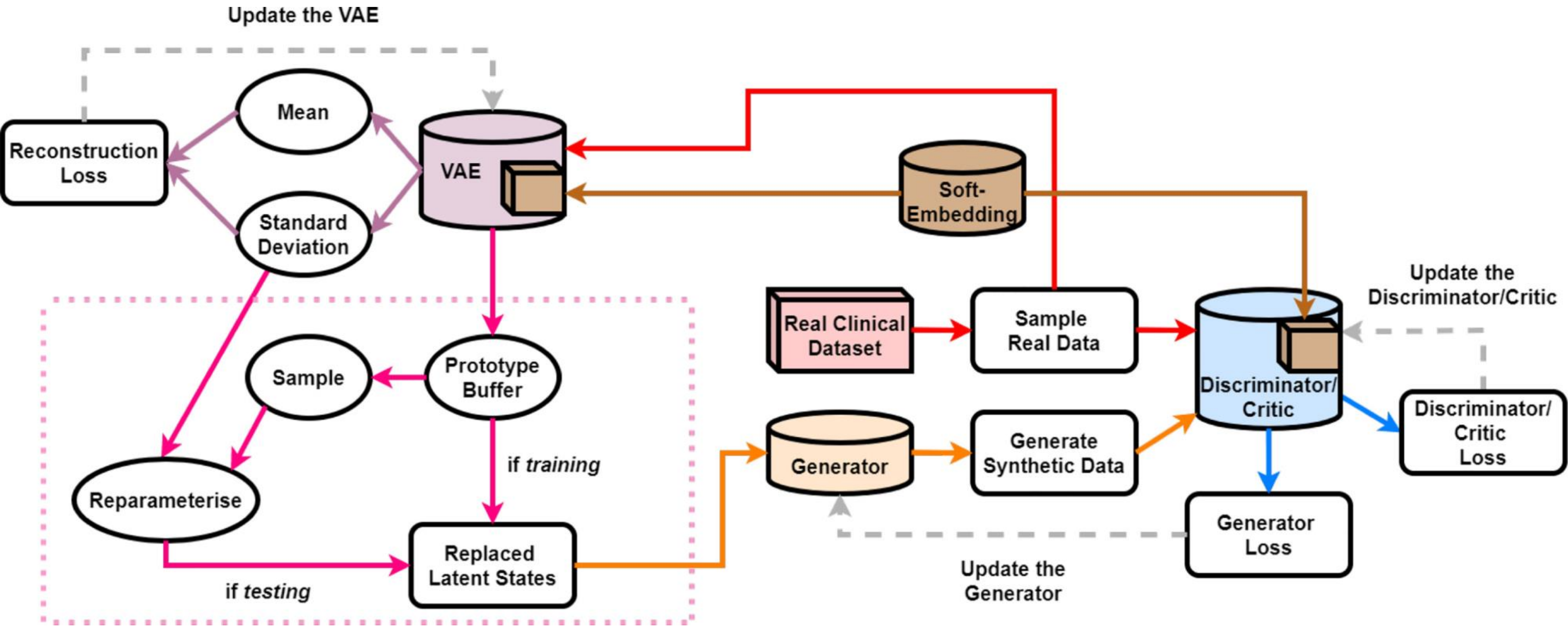
Journal of Biomedical Informatics
Volume 144, August 2023, 104436

Original Research

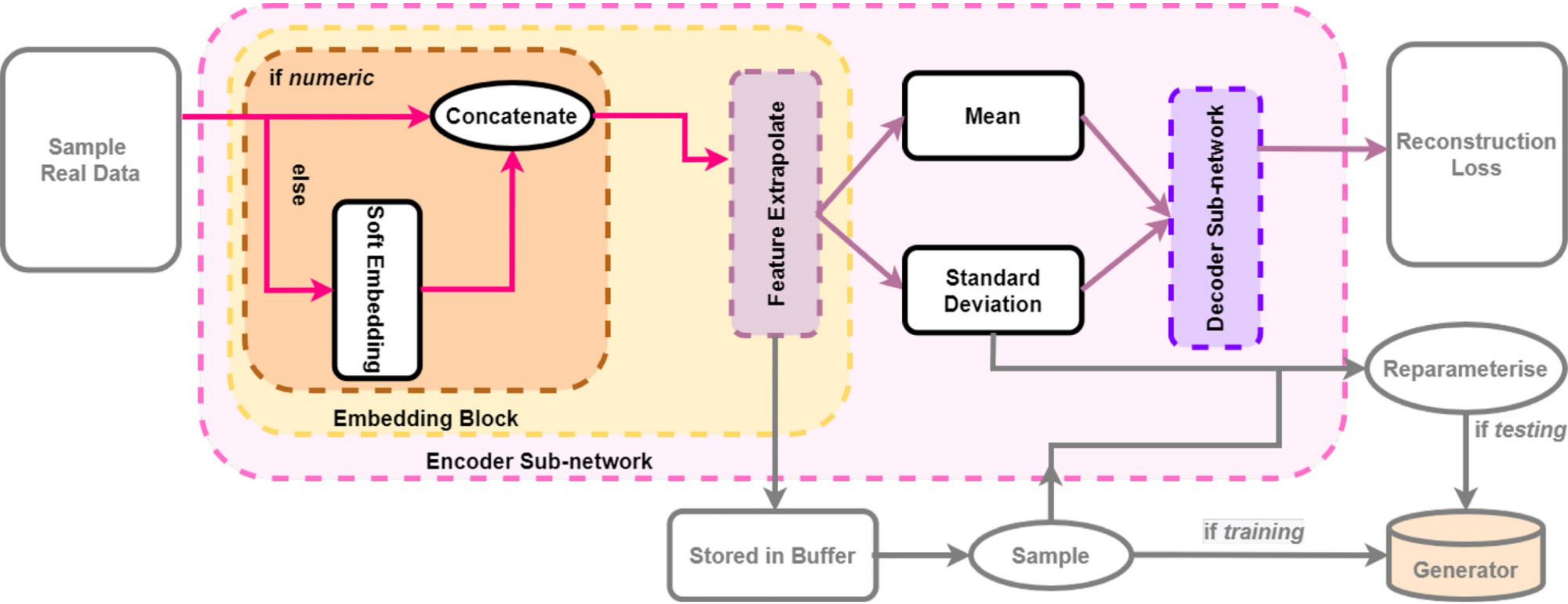
Generating synthetic clinical data that capture class imbalanced distributions with generative adversarial networks: Example using antiretroviral therapy for HIV

Nicholas I-Hsien Kuo [a], Federico Garcia [b c d], Anders Sönnerborg [e], Michael Böhm [f], Rolf Kaiser [f], Maurizio Zazzi [g], EuResist Network study group, Mark Polizzotto [h], Louisa Jorm [a], Sebastiano Barbieri [a]

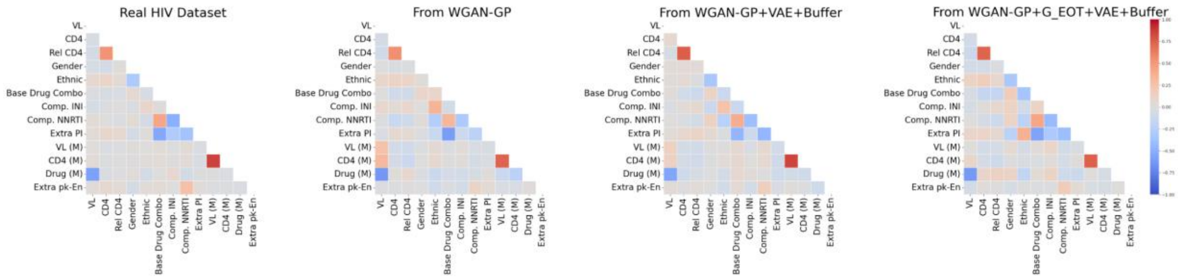
Extending GAN with a buffer which replays observed real data features



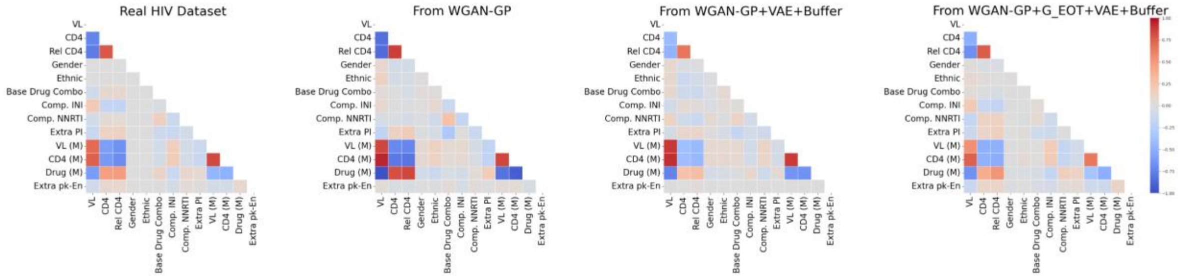
The VAE is built into the critic to collect and extract real data features



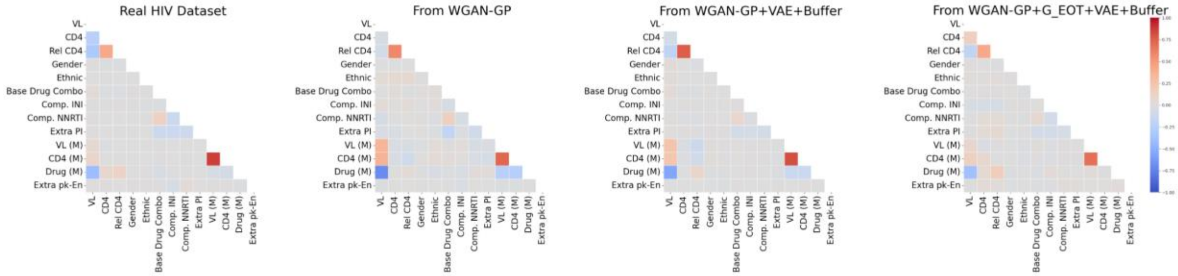
INFORMATICS PROFESSIONALS. LEADING THE WAY.



(a) The classic static correlation.



(b) The dynamic correlation in trends.



(c) The dynamic correlation in cycles.

Multi-modal AI and Generative AI

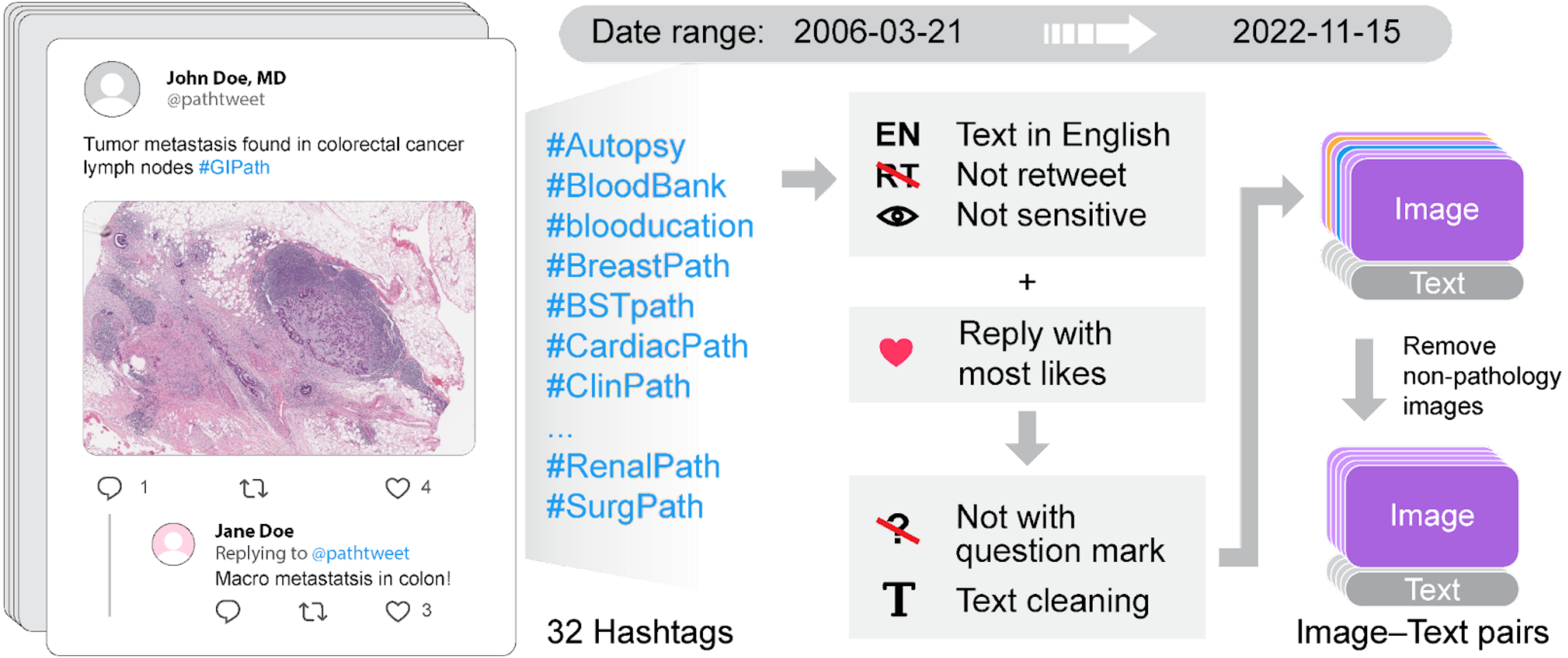
Article | Published: 17 August 2023

A visual–language foundation model for pathology image analysis using medical Twitter

Zhi Huang, Federico Bianchi, Mert Yuksekgonul, Thomas J. Montine & James Zou

Nature Medicine **29**, 2307–2316 (2023)

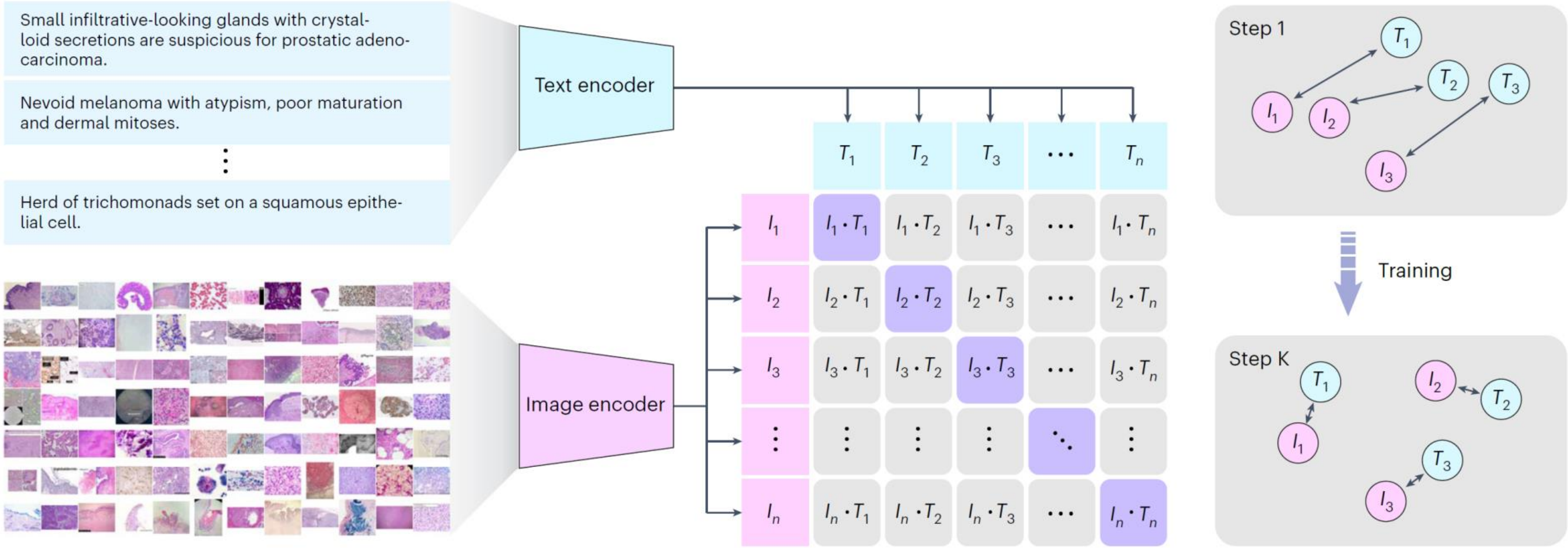
Creating OpenPath: >200K high-quality Twitter image-text pairs



Largest public dataset of pathology image + discussions.

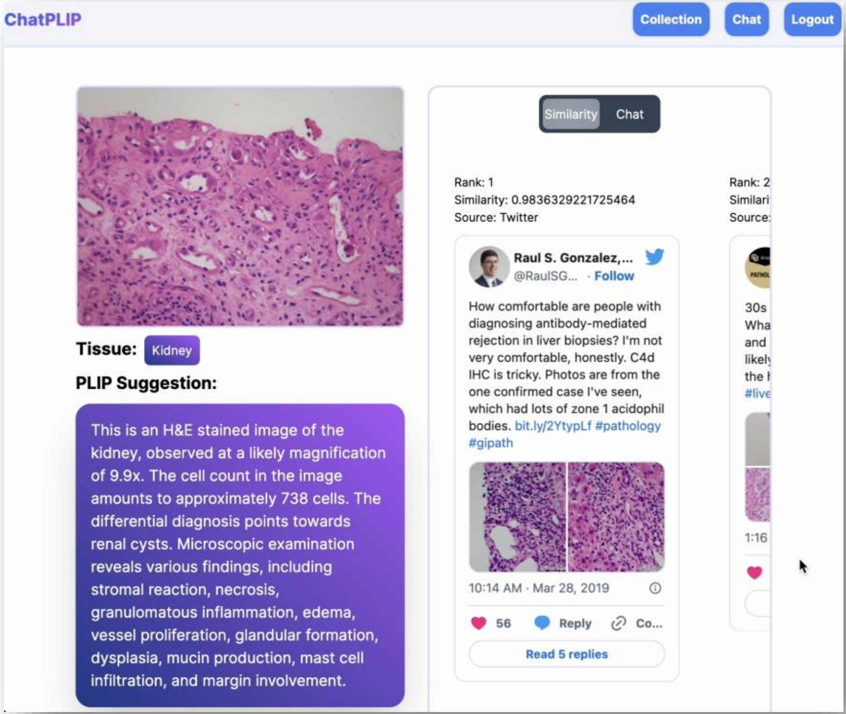
Slides courtesy of James Zou

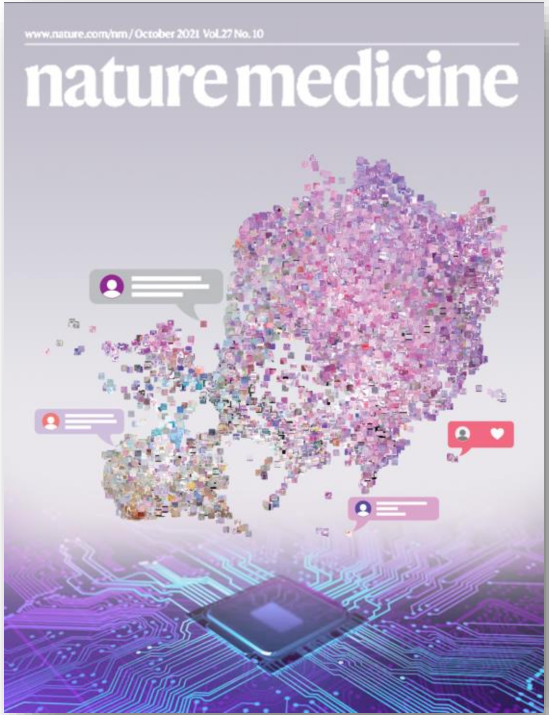
Pathology Language-Image Pretraining (PLIP)



Slides courtesy of James Zou

PLIP facilitates knowledge sharing and clinical decision support





👈 Search engine is available online now.

Slides courtesy of James Zou

Clinical Decision Support

Clinical Decision Support

Original Investigation | Health Informatics

September 25, 2023

APPRAISE-AI Tool for Quantitative Evaluation of AI Studies for Clinical Decision Support

Jethro C. C. Kwong, MD[1,2]; Adree Khondker, MD[1]; Katherine Lajkosz, MSc[1,3]; Matthew B. A. McDermott, PhD[4]; Xavier Borrat Frigola, MD[5,6]; Melissa D. McCradden, PhD[7,8,9]; Muhammad Mamdani, PharmD[2,10]; Girish S. Kulkarni, MD[1,11]; Alistair E. W. Johnson, DPhil[2,12,13]

JAMA Netw Open. 2023;6(9):e2335377. doi:10.1001/jamanetworkopen.2023.35377

Which AI model would you trust?

Study A	Study B
Abstract Data from 1 Academic Hospital	Abstract Data from 2 Community Hospitals
↓	↓
Train Random Forest Model	Train Random Forest Model
↓	↓
Validate Model on 1 other Academic Hospital	Validate Model on 2 other Academic Hospitals
↓	↓
AUROC = 0.80	AUROC = 0.80

APPRAISE-AI

Try APPRAISE-AI

0-19: Very low quality
20-39: Low quality
40-59: Moderate quality
60-79: High quality
80-100: Very high quality

	Article ID	
Methods		
4	**Source of Data: Describe how the dataset was obtained (e.g., single/multi-center, local/national database, etc.), and study period. If relevant, the diversity of the dataset is also described (e.g., inclusion of community hospitals, low/middle income populations, and institutions from other countries). [Max score 8]**	**4**
i	How many institutions were included in the dataset?	Multiple institutions
ii	Was the study period (start and end dates) reported?	Y
iii	Was the length of follow-up reported, if applicable?	Y
iv	What was the setting(s) of the institutions included in the data? *If not reported or unknown, select No.* Academic institutions	Y
	Institutions from multiple (> 1) countries	Y
	Community-based or rural hospital(s)	N
	Low/middle income patient populations	N
5	**Eligibility criteria: Specify all criteria for inclusion/exclusion of patients and features. Provide appropriate details (e.g., adults, age > 18) and rationale. [Max score 3]**	**3**
i	Inclusion criteria are provided	Y
ii	Exclusion criteria are provided	Y
iii	Details and rationale for criteria are provided	Y
6	**Ground truth: Define the ground truth of interest. Describe how it was collected (e.g., manual annotation by experts) and encoded (e.g., binary, categorical, dichotomized continuous, continuous variable, etc.). [Max score 6]**	**6**
i	Ground truth of interest is clearly defined *For unsupervised learning, describe what measure(s) will be used to correlate with the clusters (e.g., correlating disease-specific features with overall survival)*	Y
ii	How was the ground truth determined?	Multiple (>1), experts

Slides courtesy of Alistair Johnson

Potential applications of APPRAISE-AI

Researchers
Systematic reviews
Improve quality of AI studies

Hospitals
Implement AI into patient care

Journals/Conferences
Pre-screen submitted articles/abstracts

Funding Agencies
Grant allocation

Slides courtesy of Alistair Johnson

Clinical Decision Support

JAMIA

A SCHOLARLY JOURNAL OF INFORMATICS IN HEALTH AND BIOMEDICINE

JOURNAL ARTICLE FEATURED

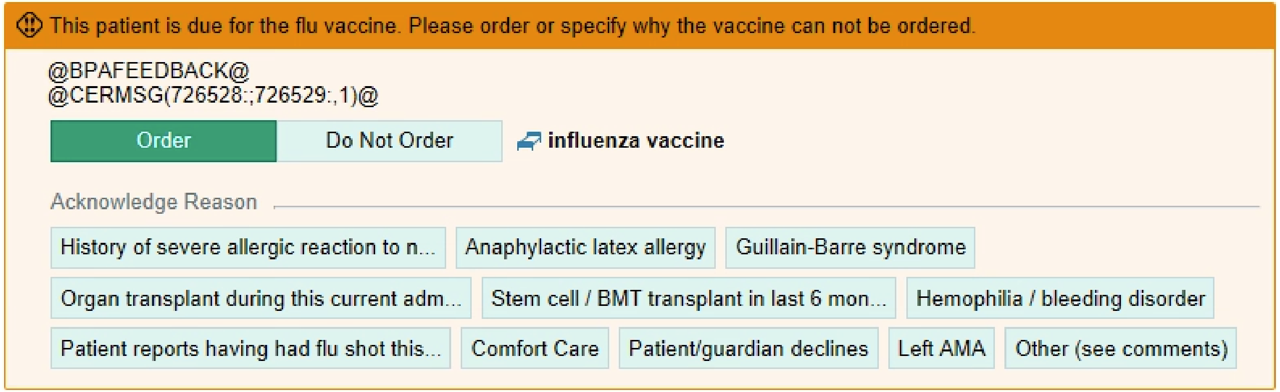
Using AI-generated suggestions from ChatGPT to optimize clinical decision support

Siru Liu, Aileen P Wright, Barron L Patterson, Jonathan P Wanderer, Robert W Turer, Scott D Nelson, Allison B McCoy, Dean F Sittig, Adam Wright

Journal of the American Medical Informatics Association, Volume 30, Issue 7, July 2023, Pages 1237–1245, https://doi.org/10.1093/jamia/ocad072

Leveraging Explainable Artificial Intelligence (XAI) to Optimize CDS

- 90% of alerts are overridden or ignored by clinicians with justifiable reasons (e.g., irrelevancy, poor timing, or incomplete clinical conditions)
- Alert fatigue

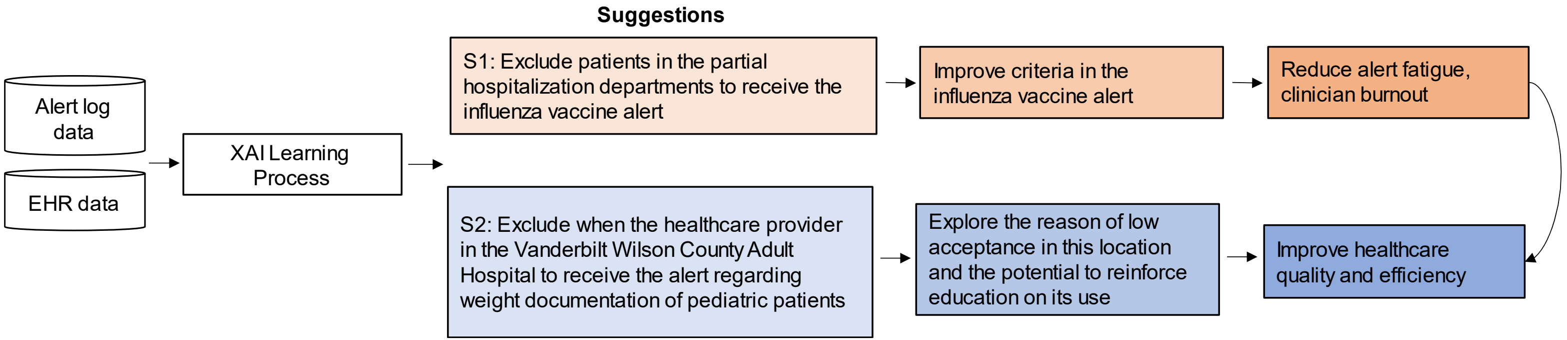


Patients with severe psychiatric disease are often hospitalized, while patients with less severe disease are treated in the outpatient setting. We have a "partial hospitalization" program, where patients spend the day at the hospital, but sleep at home. In our EHR, they are treated as inpatients, so they were receiving the inpatient flu vaccine alert. However, they program does not have the ability to administer flu shots, so they rarely accepted the alert. Our XAI tool identified that patients in this program were very unlikely to have the alert accepted. After discussion with our psychiatric leadership, we excluded these patients from the BPA, eliminating unhelpful alert firing.

Slides courtesy of Siru Liu

Leveraging Explainable Artificial Intelligence (XAI) to Optimize CDS

- Identified 96 helpful suggestions for 18 BPAs
- Among 2,991,823 firings, 278,807 firings (9.3%) could have been eliminated.
- Using XAI could identify improvements regarding CDS that might be overlooked or delayed in manual reviews.
- Unveils a secondary purpose for the XAI: to improve quality by discovering scenarios where CDS alerts are not accepted due to workflow, education, or staffing issues.



Slides courtesy of Siru Liu

Clinical Decision Support

npj | digital medicine

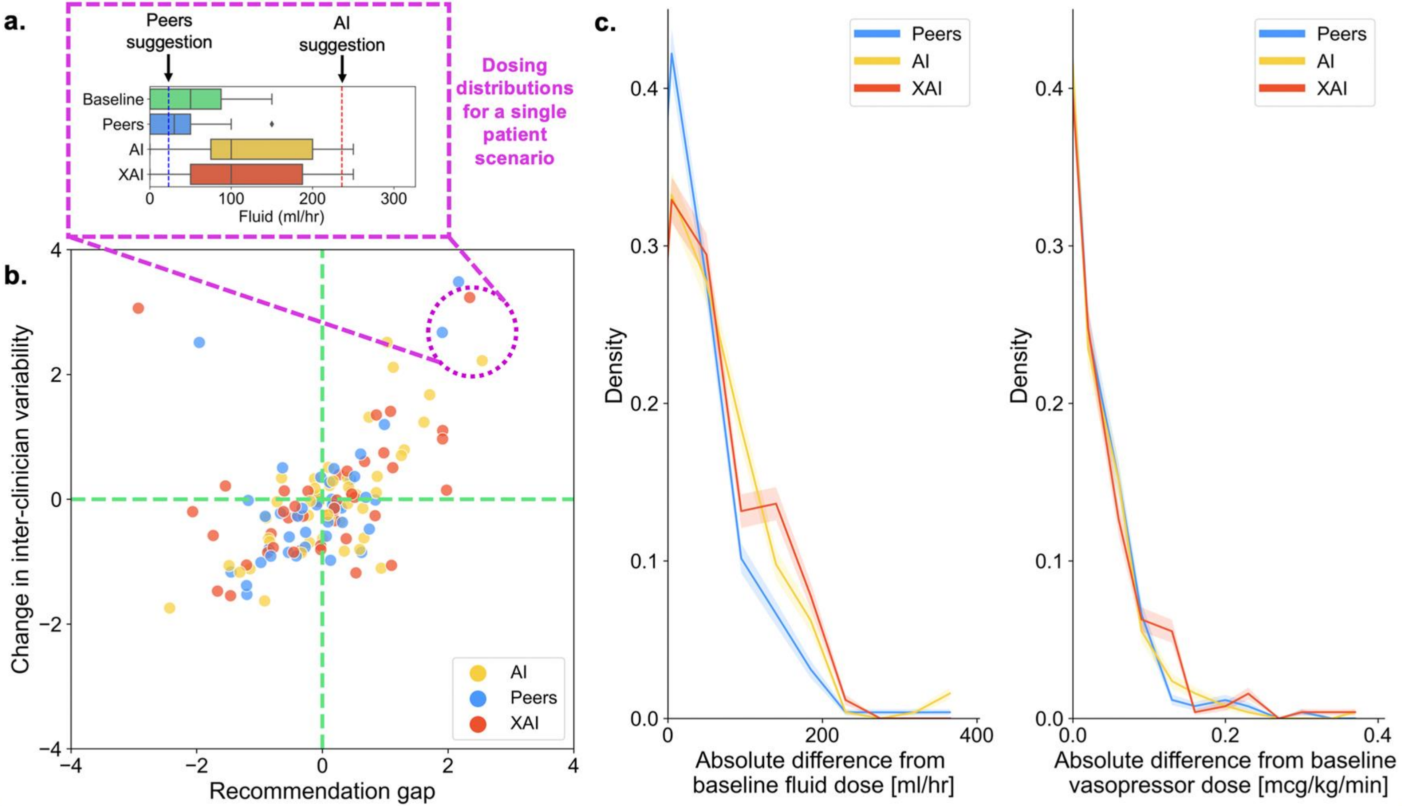
Article | Open access | Published: 07 November 2023

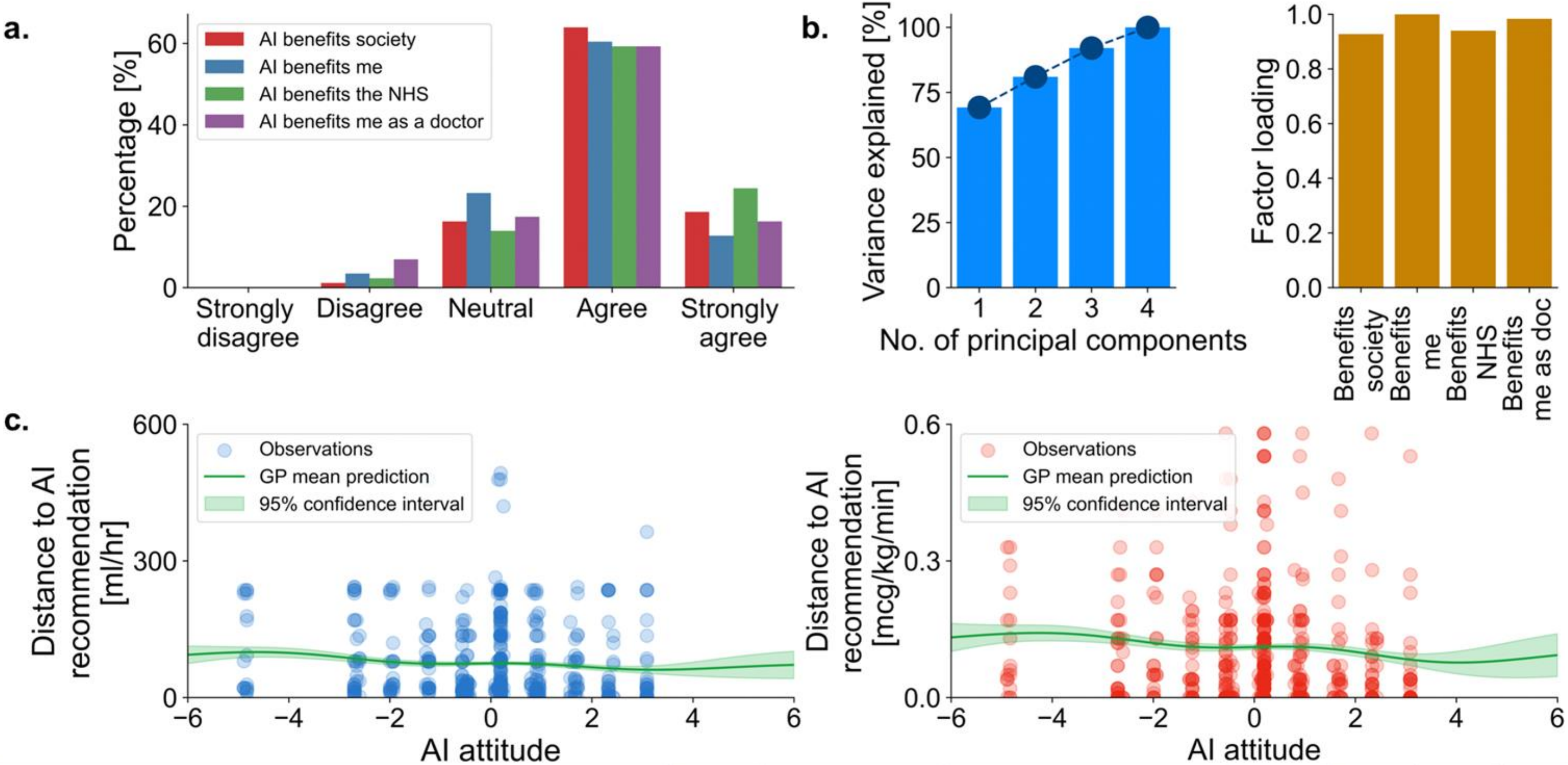
Quantifying the impact of AI recommendations with explanations on prescription decision making

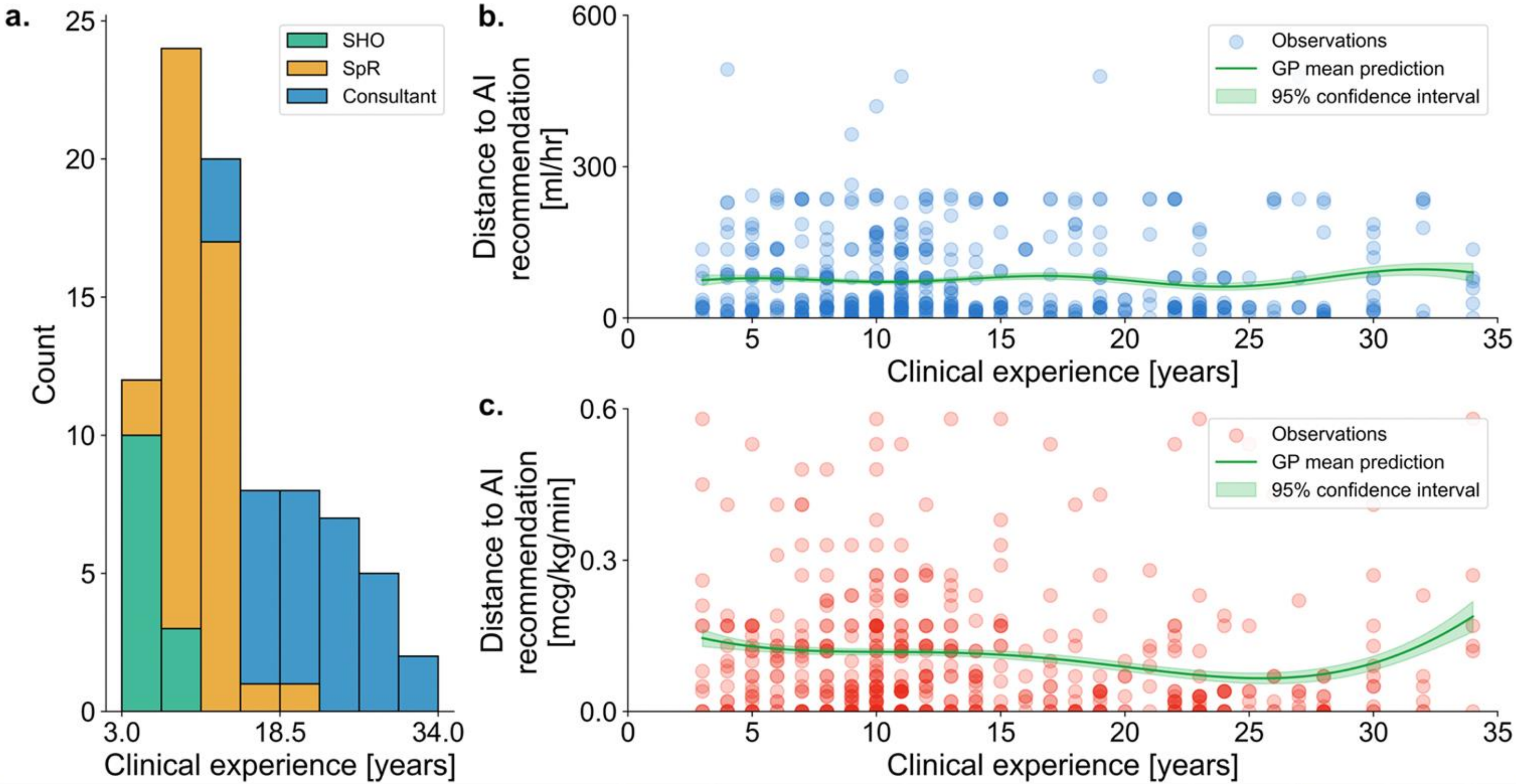
Myura Nagendran, Paul Festor, Matthieu Komorowski, Anthony C. Gordon & Aldo A. Faisal

npj Digital Medicine **6**, Article number: 206 (2023)

INFORMATICS PROFESSIONALS. LEADING THE WAY.




a.
Percentage [%]
AI benefits society
AI benefits me
AI benefits the NHS
AI benefits me as a doctor
Strongly disagree
Disagree
Neutral
Agree
Strongly agree
b.
Variance explained [%]
No. of principal components
Factor loading
Benefits society
Benefits me
Benefits NHS
Benefits me as doc
c.
Distance to AI recommendation [ml/hr]
Observations
GP mean prediction
95% confidence interval
AI attitude
Distance to AI recommendation [mcg/kg/min]
Observations
GP mean prediction
95% confidence interval
AI attitude

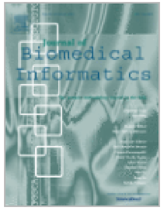


a.
Count
SHO
SpR
Consultant
0
5
10
15
20
25
3.0
18.5
34.0
Clinical experience [years]
b.
Distance to AI recommendation [ml/hr]
600
300
0
Observations
GP mean prediction
95% confidence interval
0
5
10
15
20
25
30
35
Clinical experience [years]
c.
Distance to AI recommendation [mcg/kg/min]
0.6
0.3
0.0
Observations
GP mean prediction
95% confidence interval
0
5
10
15
20
25
30
35
Clinical experience [years]

Clinical Decision Support

Journal of Biomedical Informatics

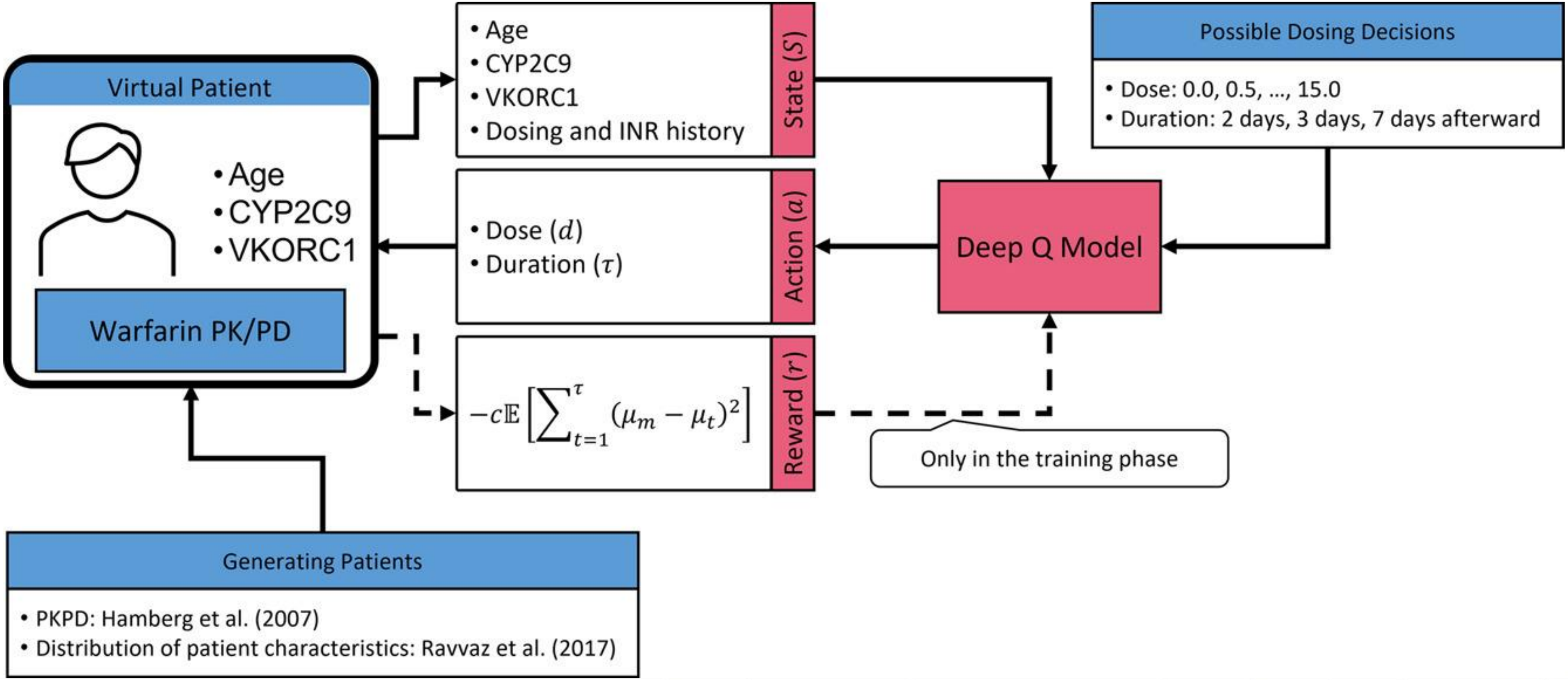
Volume 137, January 2023, 104267

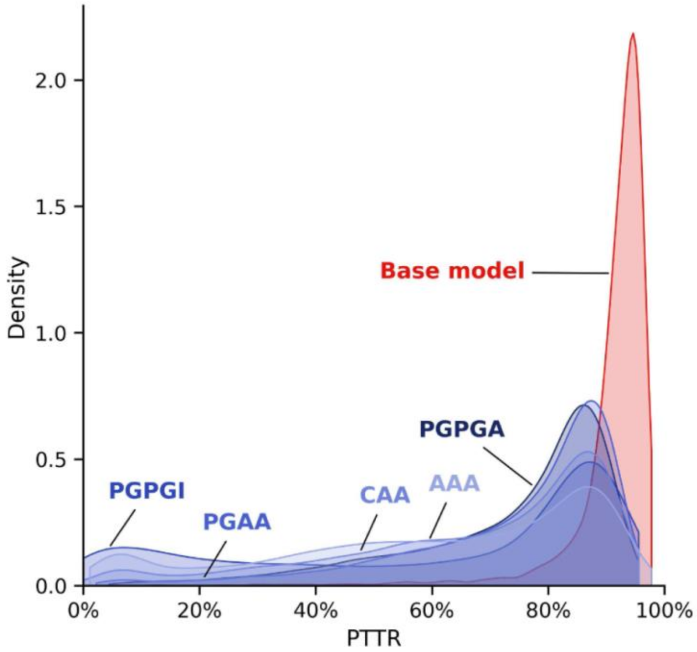


Original Research

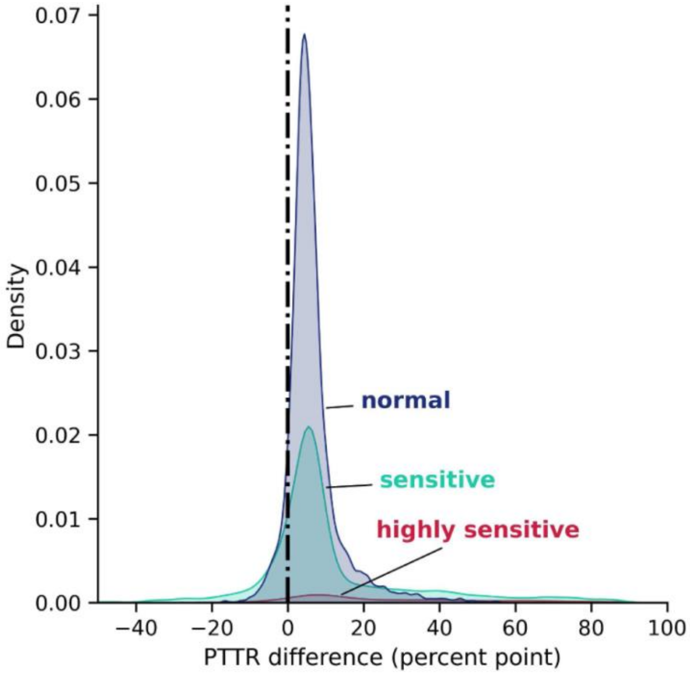
Optimizing warfarin dosing using deep reinforcement learning

Sadjad Anzabi Zadeh, W. Nick Street, Barrett W. Thomas


Virtual Patient
• Age
• CYP2C9
• VKORC1
Warfarin PK/PD
• Age
• CYP2C9
• VKORC1
• Dosing and INR history
State (S)
• Dose (d)
• Duration (τ)
Action (a)
$-c\mathbb{E}\left[\sum_{t=1}^{\tau}(\mu_m - \mu_t)^2\right]$
Reward (r)
Possible Dosing Decisions
• Dose: 0.0, 0.5, ..., 15.0
• Duration: 2 days, 3 days, 7 days afterward
Deep Q Model
Only in the training phase
Generating Patients
• PKPD: Hamberg et al. (2007)
• Distribution of patient characteristics: Ravvaz et al. (2017)



(a) The distribution of PTTR values for the base model and baseline protocols



(b) Percentage point gain/loss of individuals by adopting the proposed protocol vs their best respective baseline protocol

Imaging Informatics

Imaging Informatics

INFORMATICS PROFESSIONALS. LEADING THE WAY.

3524 IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 42, NO. 12, DECEMBER 2023

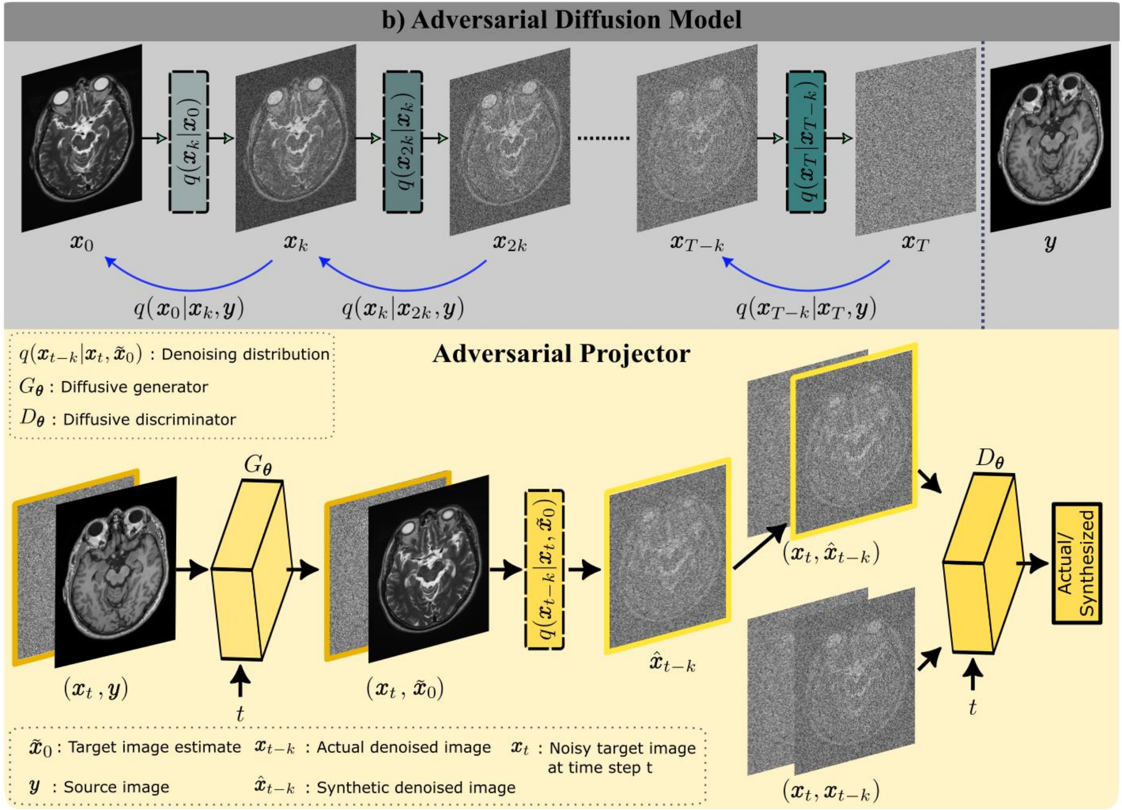
Unsupervised Medical Image Translation With Adversarial Diffusion Models

Muzaffer Özbey, Onat Dalmaz, Salman U. H. Dar, Hasan A. Bedel, Şaban Özturk, Alper Güngör, and Tolga Çukur, *Senior Member, IEEE*

SynDiff: Adversarial diffusion model for unsupervised translation

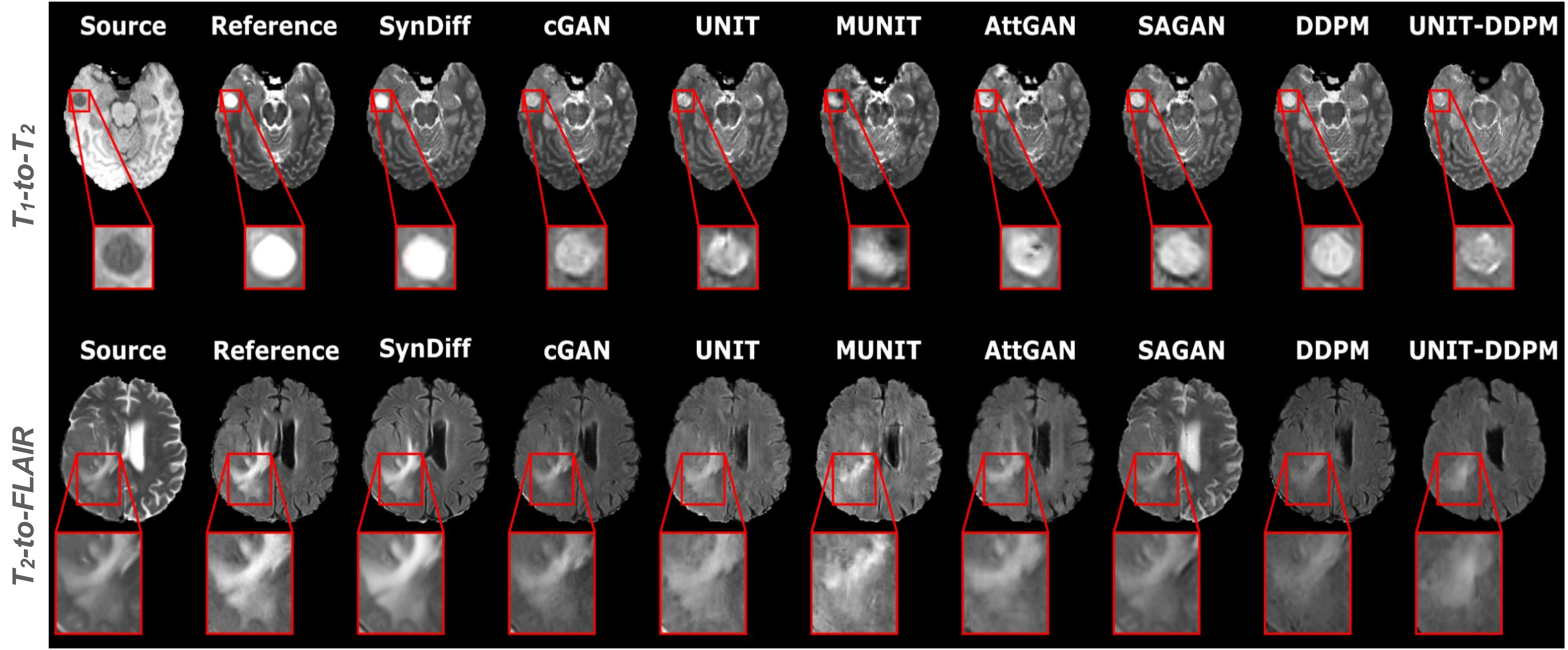
SynDiff:

- Source-conditional diffusion process to generate target images of a given anatomy
- Adversarial projector for fast image sampling in only 4 steps
- Cycle-consistent architecture for unsupervised training
- (T=4 steps, ~100 msec per 2D cross section)



Slides courtesy of Tolga Cukur

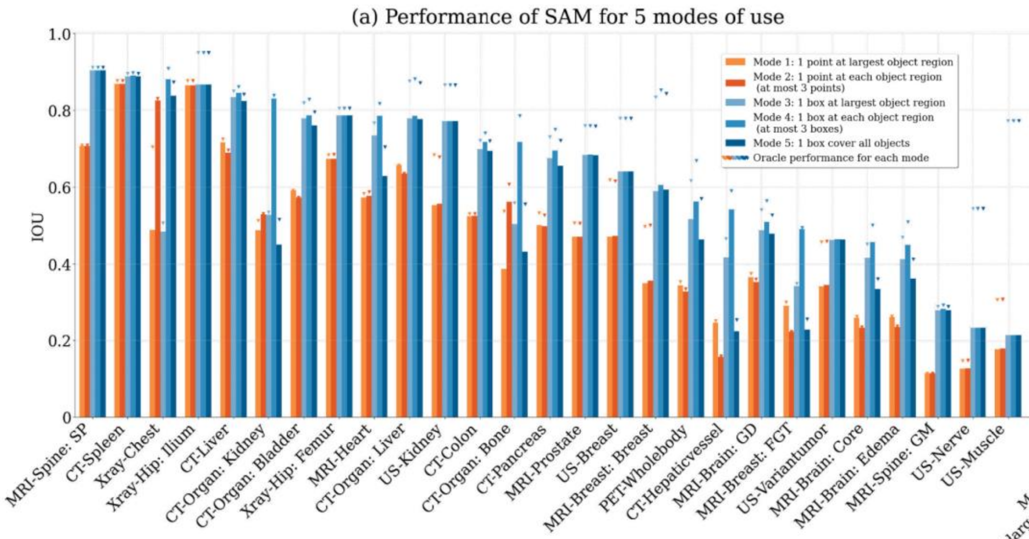
SynDiff: T1-to-T2, T2-to-FLAIR tasks

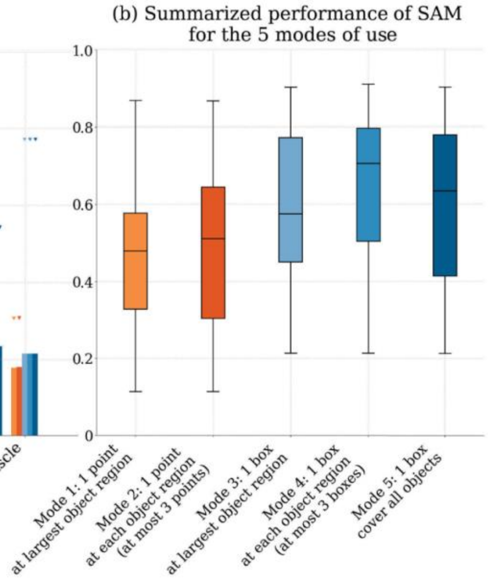


Medical Image Analysis
Volume 89, October 2023, 102918

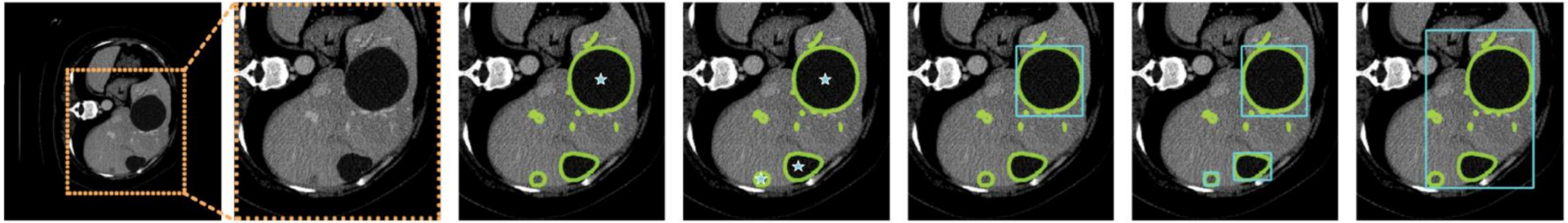
Segment anything model for medical image analysis: An experimental study

Maciej A. Mazurowski [a b c d], Haoyu Dong [b], Hanxue Gu [b], Jichen Yang [b], Nicholas Konz [b], Yixin Zhang [b]





INFORMATICS PROFESSIONALS. LEADING THE WAY.



Input Image

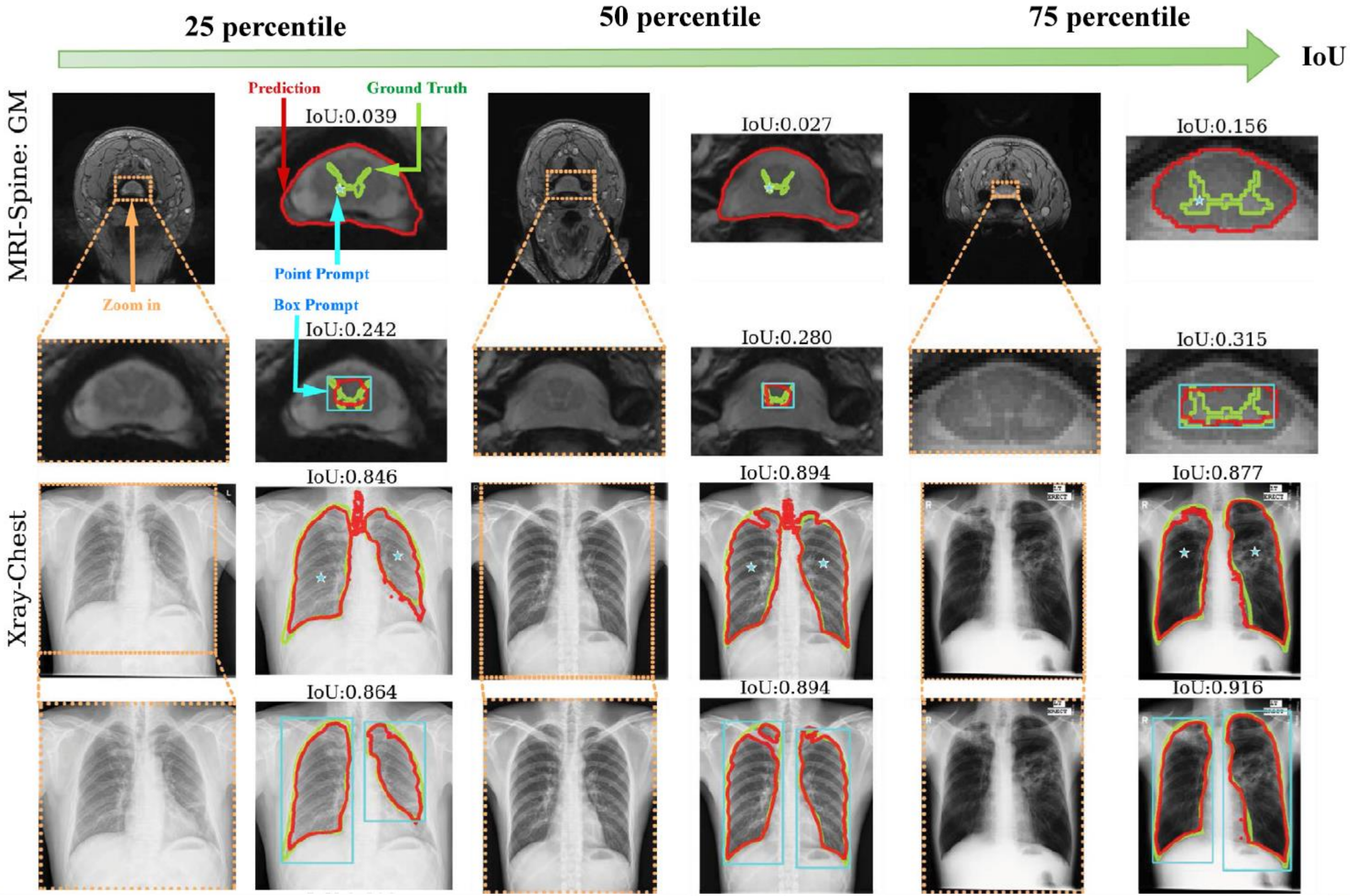
Prompt Mode 1: 1 **point** at **largest** object region

Prompt Mode 2: 1 **point** for **each** object region (at most 3 points)

Prompt Mode 3: 1 **box** around **largest** object region

Prompt Mode 4: 1 **box** around **each** object region (at most 3 boxes)

Prompt Mode 5: 1 **box** covers **all** objects


25 percentile
50 percentile
75 percentile
IoU
MRI-Spine: GM
Prediction
Ground Truth
IoU:0.039
Point Prompt
Zoom in
Box Prompt
IoU:0.242
IoU:0.027
IoU:0.280
IoU:0.156
IoU:0.315
Xray-Chest
IoU:0.846
IoU:0.864
IoU:0.894
IoU:0.894
IoU:0.877
IoU:0.916

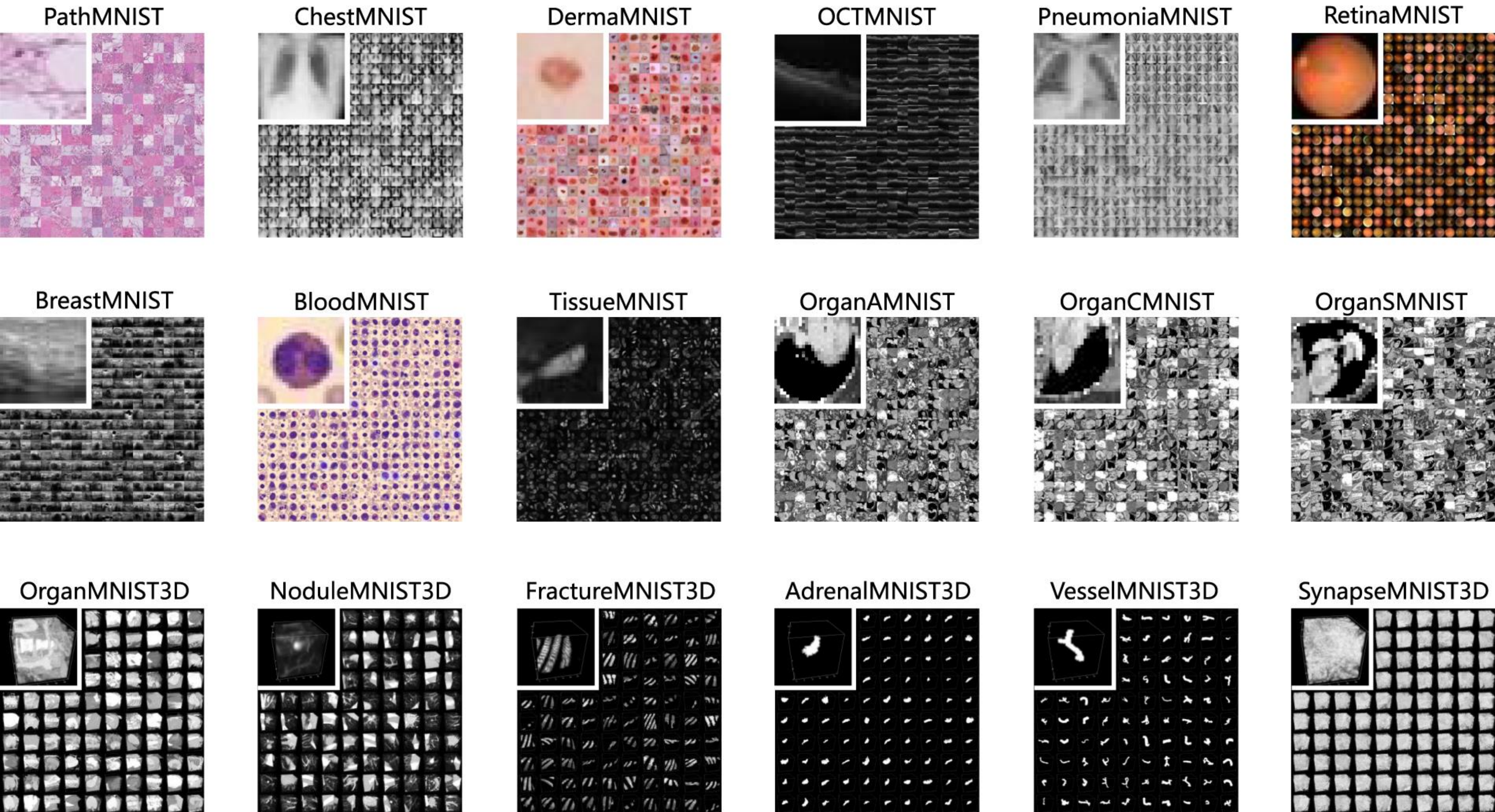
Imaging Informatics

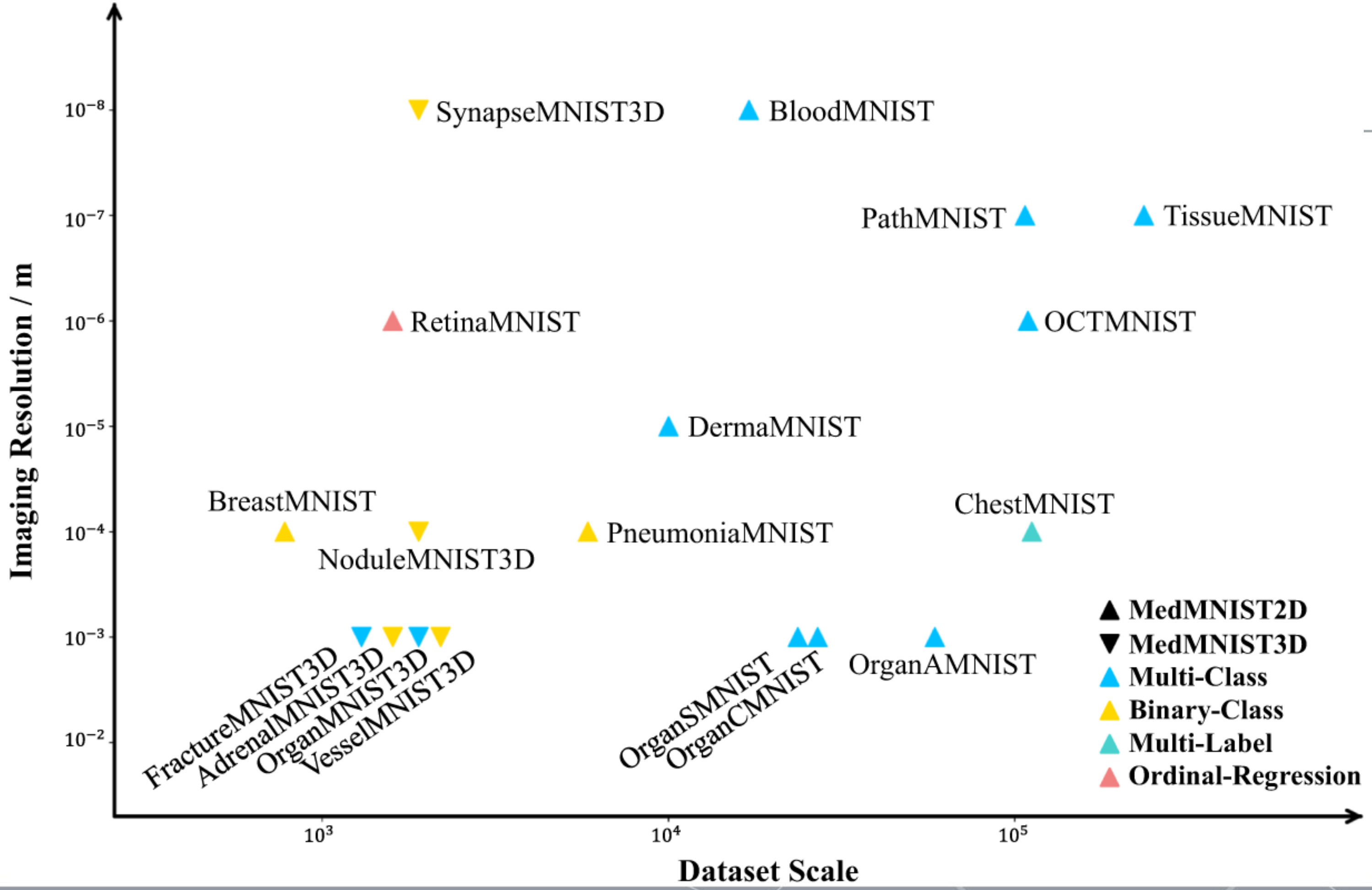
Data Descriptor | Open access | Published: 19 January 2023

MedMNIST v2 - A large-scale lightweight benchmark for 2D and 3D biomedical image classification

Jiancheng Yang, Rui Shi, Donglai Wei, Zequan Liu, Lin Zhao, Bilian Ke, Hanspeter Pfister & Bingbing Ni

Scientific Data **10**, Article number: 41 (2023) | Cite this article


PathMNIST
ChestMNIST
DermaMNIST
OCTMNIST
PneumoniaMNIST
RetinaMNIST
BreastMNIST
BloodMNIST
TissueMNIST
OrganAMNIST
OrganCMNIST
OrganSMNIST
OrganMNIST3D
NoduleMNIST3D
FractureMNIST3D
AdrenalMNIST3D
VesselMNIST3D
SynapseMNIST3D


SynapseMNIST3D
BloodMNIST
PathMNIST
TissueMNIST
RetinaMNIST
OCTMNIST
DermaMNIST
BreastMNIST
NoduleMNIST3D
PneumoniaMNIST
ChestMNIST
FractureMNIST3D
AdrenalMNIST3D
OrganMNIST3D
VesselMNIST3D
OrganSMNIST
OrganCMNIST
OrganAMNIST
MedMNIST2D
MedMNIST3D
Multi-Class
Binary-Class
Multi-Label
Ordinal-Regression
Imaging Resolution / m
Dataset Scale
10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^3
10^4
10^5

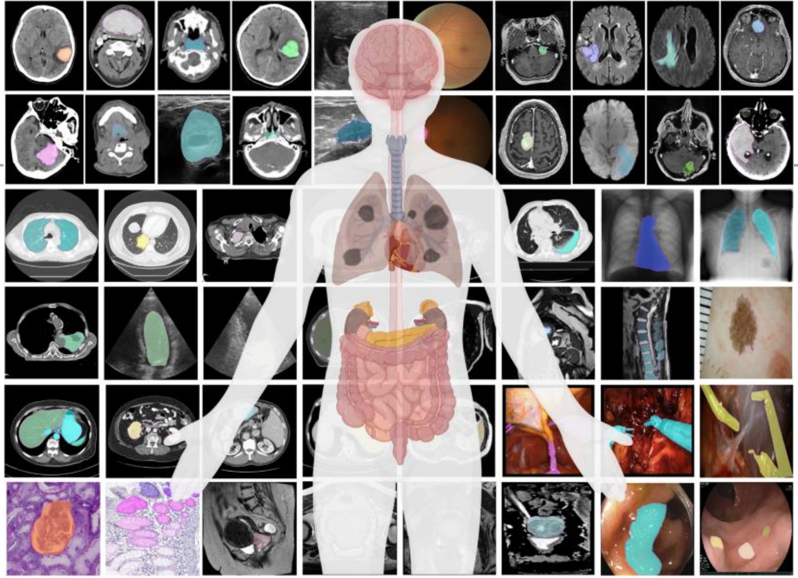
nature communications

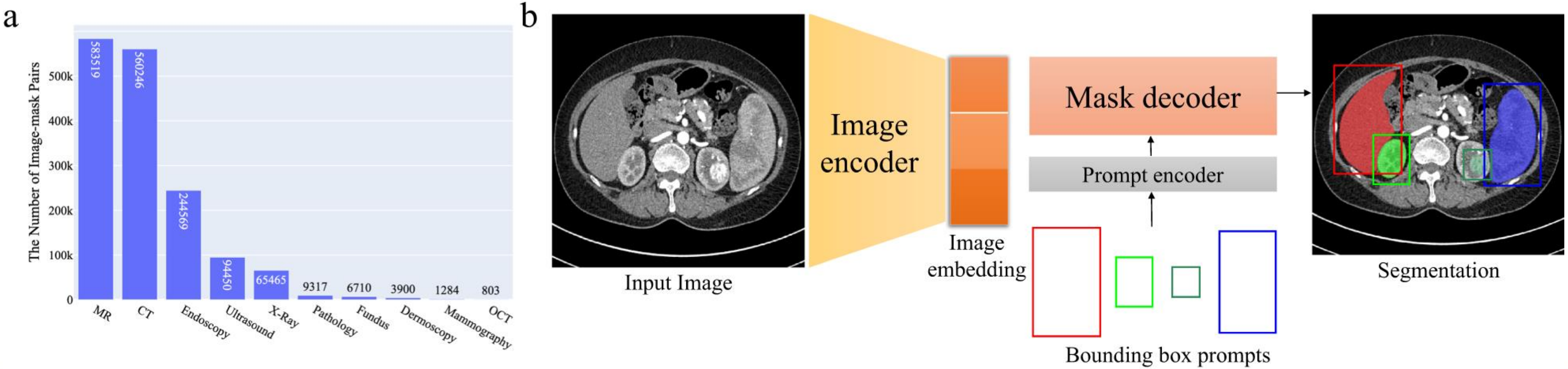
Article | Open access | Published: 22 January 2024

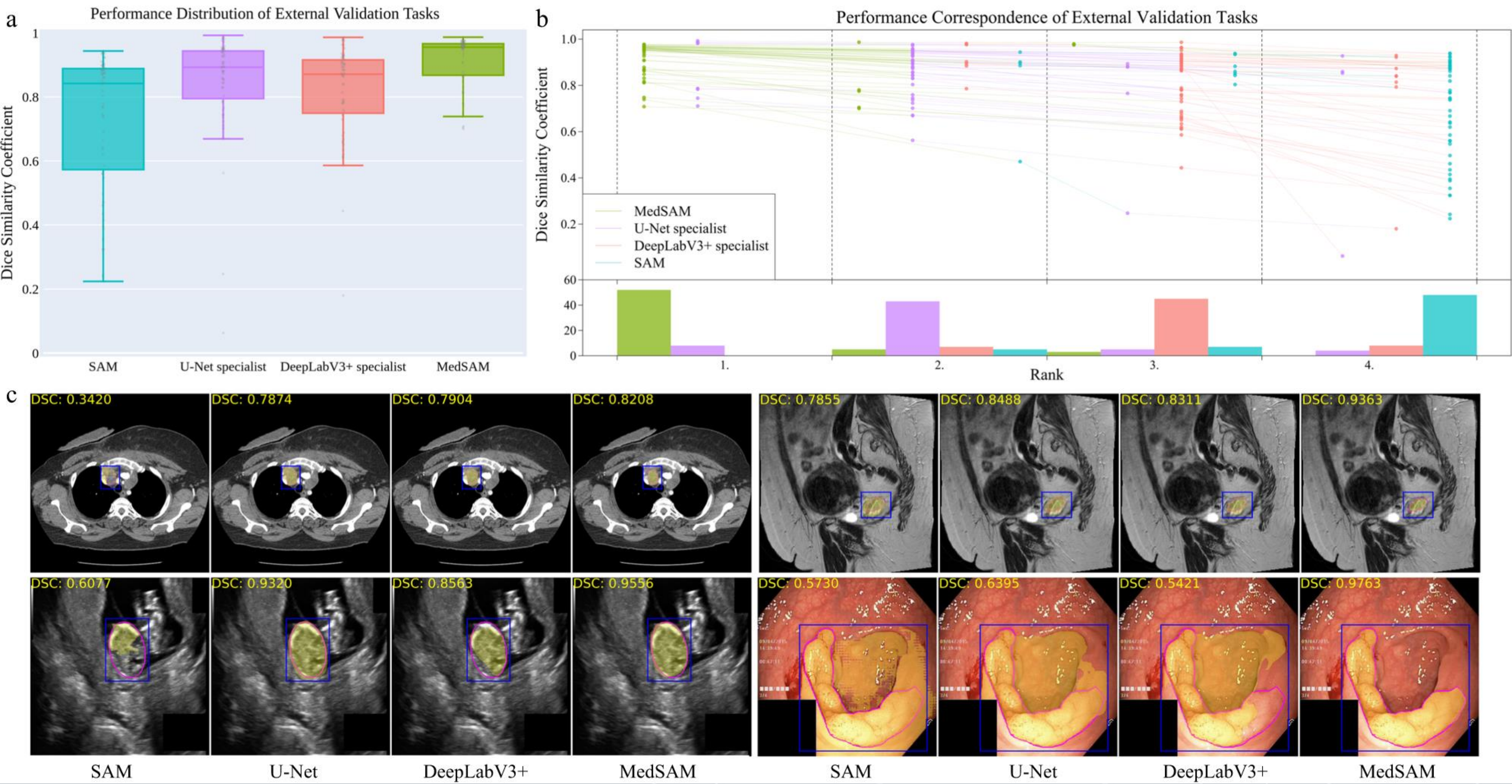
Segment anything in medical images

Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You & Bo Wang

Nature Communications **15**, Article number: 654 (2024) | Cite this article




a
The Number of Image-mask Pairs
0
100k
200k
300k
400k
500k
583519
560246
244569
94450
65465
9317
6710
3900
1284
803
MR
CT
Endoscopy
Ultrasound
X-Ray
Pathology
Fundus
Dermoscopy
Mammography
OCT
b
Input Image
Image encoder
Image embedding
Mask decoder
Prompt encoder
Bounding box prompts
Segmentation


a
Performance Distribution of External Validation Tasks
Dice Similarity Coefficient
SAM
U-Net specialist
DeepLabV3+ specialist
MedSAM
b
Performance Correspondence of External Validation Tasks
Dice Similarity Coefficient
MedSAM
U-Net specialist
DeepLabV3+ specialist
SAM
Rank
c
DSC: 0.3420
DSC: 0.7874
DSC: 0.7904
DSC: 0.8208
DSC: 0.7855
DSC: 0.8488
DSC: 0.8311
DSC: 0.9363
DSC: 0.6077
DSC: 0.9320
DSC: 0.8563
DSC: 0.9556
DSC: 0.5730
DSC: 0.6395
DSC: 0.5421
DSC: 0.9763
SAM
U-Net
DeepLabV3+
MedSAM
SAM
U-Net
DeepLabV3+
MedSAM

nature

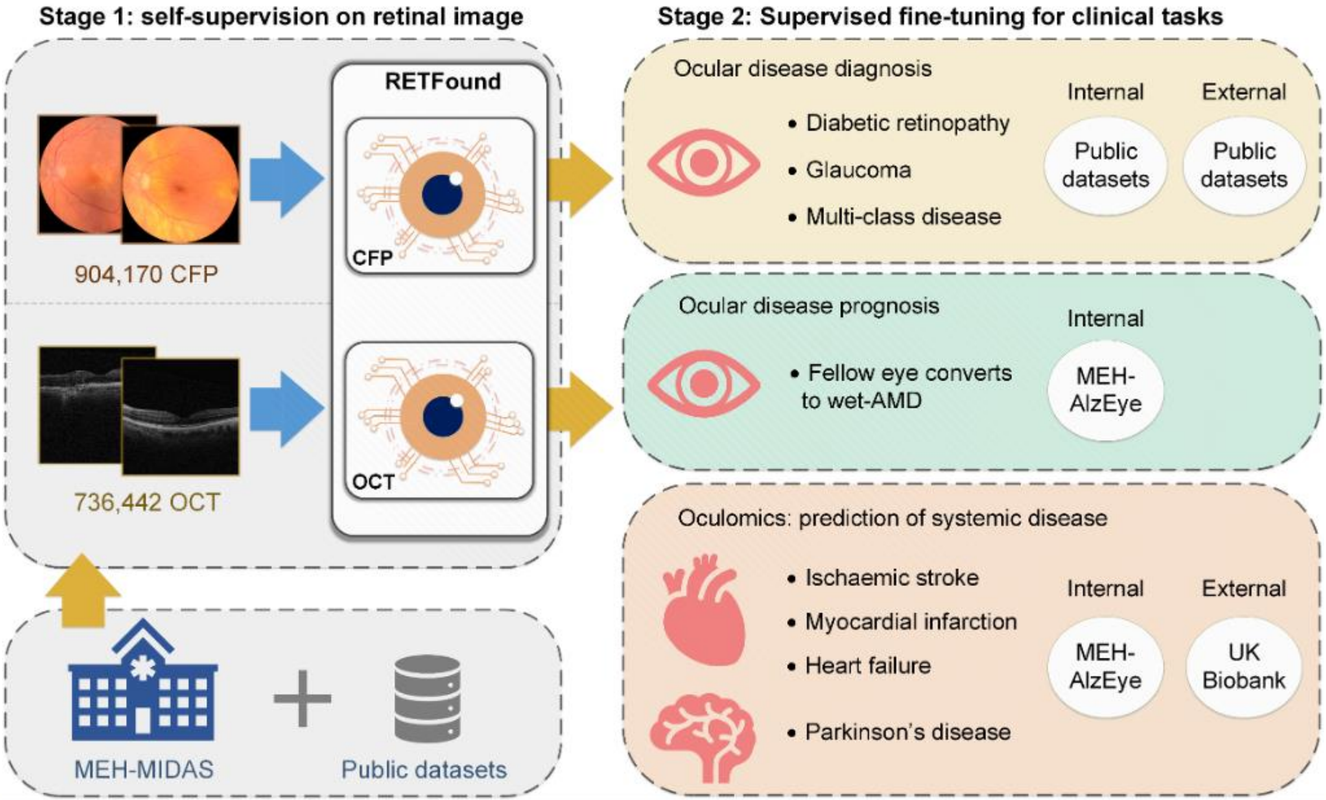
Article | Open access | Published: 13 September 2023

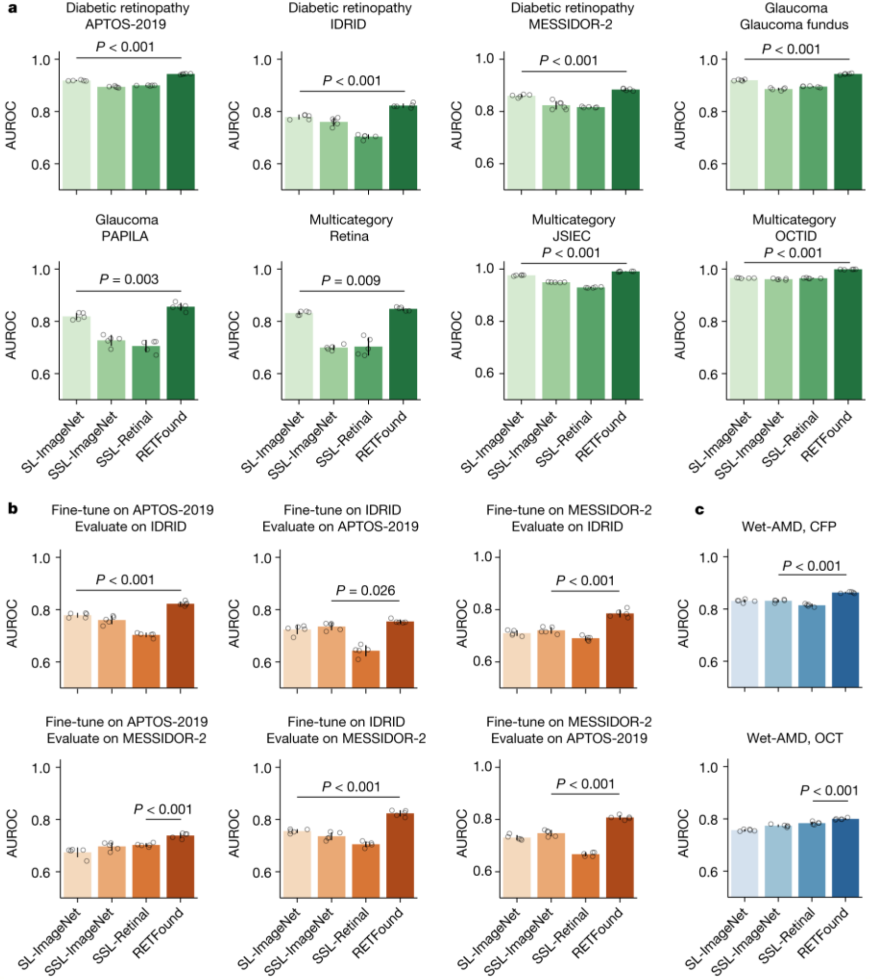
A foundation model for generalizable disease detection from retinal images

Yukun Zhou, Mark A. Chia, Siegfried K. Wagner, Murat S. Ayhan, Dominic J. Williamson, Robbert R. Struyven, Timing Liu, Moucheng Xu, Mateo G. Lozano, Peter Woodward-Court, Yuka Kihara, UK Biobank Eye & Vision Consortium, Andre Altmann, Aaron Y. Lee, Eric J. Topol, Alastair K. Denniston, Daniel C. Alexander & Pearse A. Keane

Nature **622**, 156–163 (2023) | Cite this article

Introducing RETFound…





Slides courtesy of Pearse Keane

Imaging Informatics

CVPR 2023

Dynamic Graph Enhanced Contrastive Learning for Chest X-ray Report Generation

Mingjie Li[1] Bingqian Lin[2] Zicong Chen[5] Haokun Lin[2] Xiaodan Liang[2,3,4] Xiaojun Chang[1*]
[1]ReLER, AAII, University of Technology Sydney [2]School of ISE, Sun Yat-Sen University
[3]Department of Computer Vision, Mohamed bin Zayed University of Artificial Intelligence
[4]Peng Cheng National Lab [5]The University of Hong Kong

Ground Truth Report

A single portable AP semi-upright view of the chest was obtained. Cardiomediastinal silhouette including moderate cardiomegaly is stable. An ICD device is also unchanged in position. Interval development of increased opacification in the right lower lung probably reflects asymmetric edema and a layering effusion…

Pre-constructed Graph

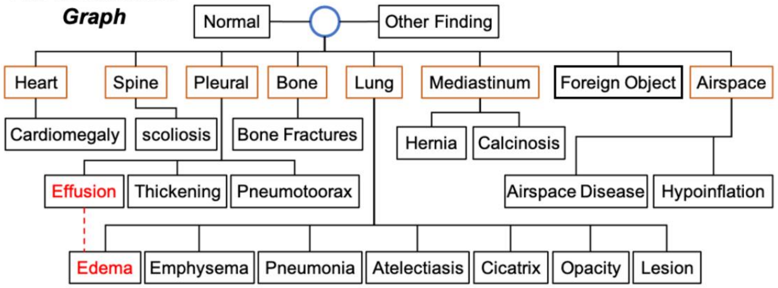
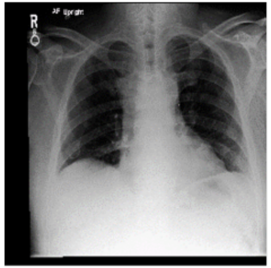
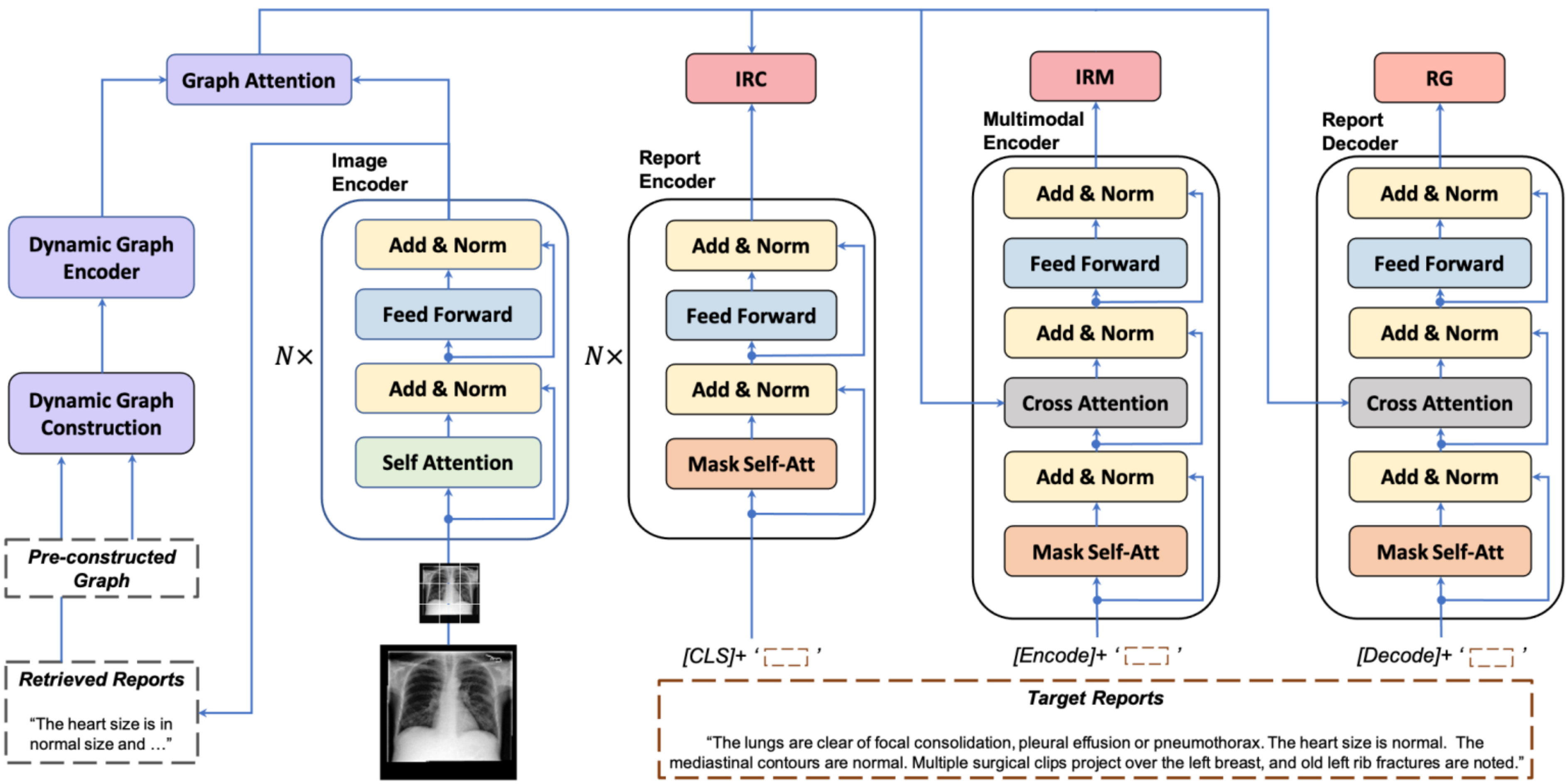


Image	*Ground Truth Report*	*Retrieved Specific Knowledge*	*Ours*	*R2Gen*
	lateral view somewhat limited due to overlying motion artifact lungs are low in volume there is no focal airspace the consolidation to suggest pneumonia a 1 2-cm calcified granuloma just below the medial aspect of the right hemidiaphragm is unchanged from prior study no pleural effusions or pulmonary edema there is no pneumothorax the inferior sternotomy wire is fractured but unchanged surgical clips and vascular markers in the thorax are related to prior cabg surgery.	effusion located _ at pleural,vascular modify pulmonary,sternotomy suggestive _ of cabg,consolidation suggestive _ of effusion,effusion suggestive _ of consolidation,low modify effusion,low suggestive _ of effusion, low suggestive _ of consolidation,vascular modify chest,effusion suggestive _ of atelectasis,atelectasis modify lung	lung volumes are low. there is no focal consolidation effusion or pneumothorax. airspace consolidation is noted within the left upper lobe compatible with pneumonia. borderline size of the cardiac silhouette without evidence of pulmonary edema. midline sternotomy wires and mediastinal clips are again noted.	ap upright and lateral views of the chest provided . lung volumes are low limiting assessment . allowing for this there is no focal consolidation effusion or pneumothorax . the cardiomediastinal silhouette is normal . imaged osseous structures are intact . no free air below the right hemidiaphragm is seen.


Graph Attention
IRC
IRM
RG
Multimodal Encoder
Report Decoder
Image Encoder
Report Encoder
Dynamic Graph Encoder
Dynamic Graph Construction
Add & Norm
Feed Forward
Add & Norm
Self Attention
N×
Add & Norm
Feed Forward
Add & Norm
Mask Self-Att
N×
Add & Norm
Feed Forward
Add & Norm
Cross Attention
Add & Norm
Mask Self-Att
Add & Norm
Feed Forward
Add & Norm
Cross Attention
Add & Norm
Mask Self-Att
Pre-constructed Graph
Retrieved Reports
"The heart size is in normal size and …"
[CLS]+ ' '
[Encode]+ ' '
[Decode]+ ' '
Target Reports
"The lungs are clear of focal consolidation, pleural effusion or pneumothorax. The heart size is normal. The mediastinal contours are normal. Multiple surgical clips project over the left breast, and old left rib fractures are noted."

Evaluation and Implementation

Evaluation and Implementation

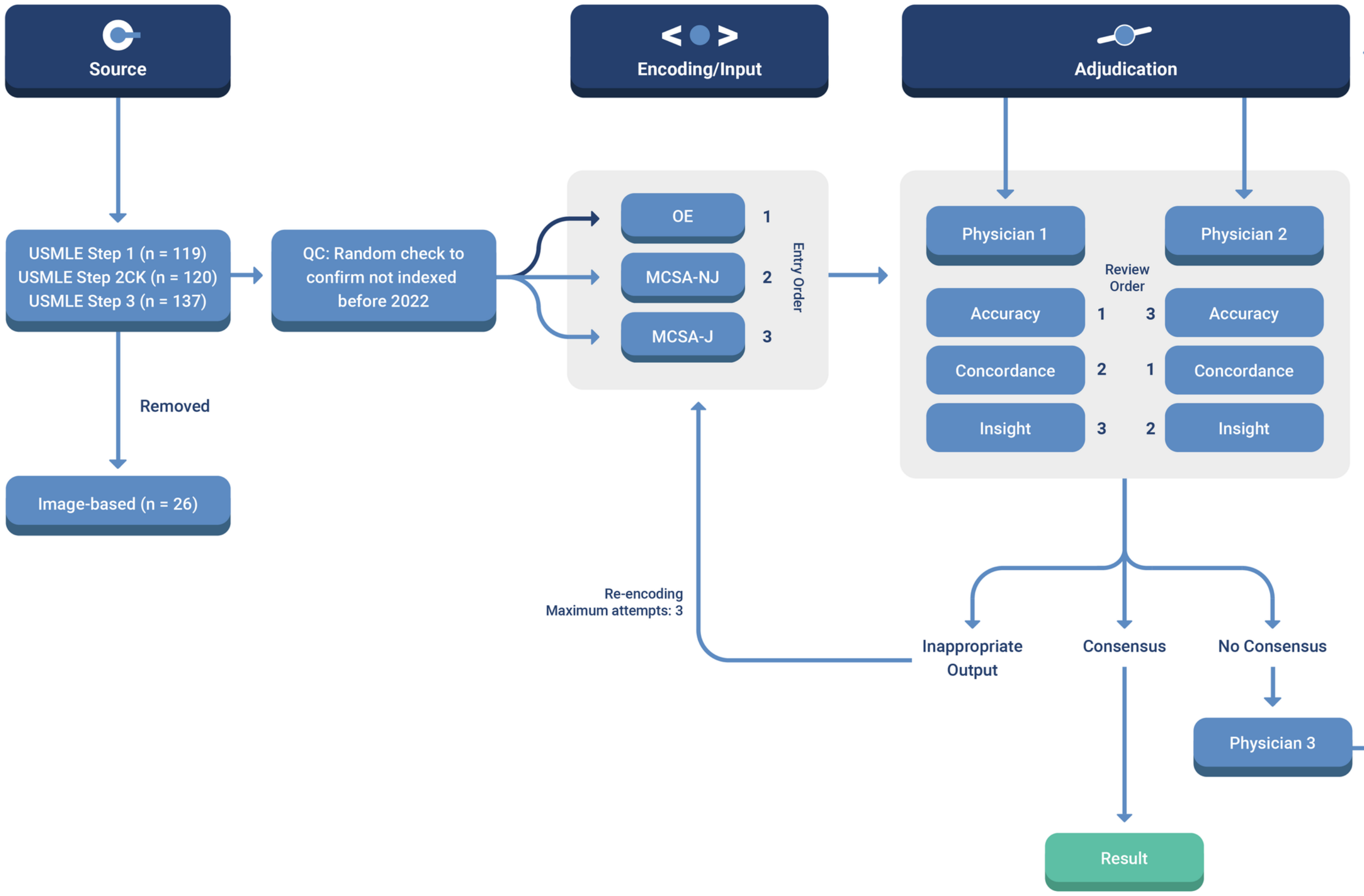
PLOS DIGITAL HEALTH

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

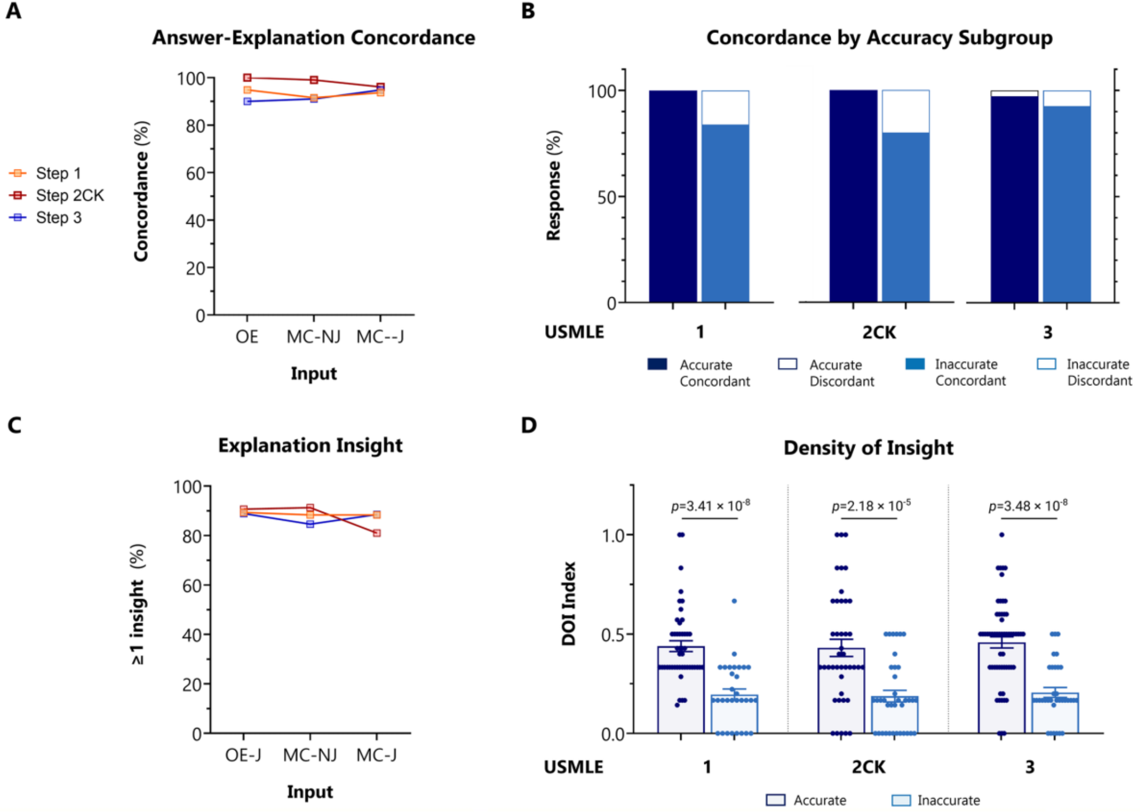
Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models

Tiffany H. Kung, Morgan Cheatham, Arielle Medenilla, Czarina Sillos, Lorie De Leon, Camille Elepaño, Maria Madriaga, Rimel Aggabao, Giezel Diaz-Candido, James Maningo, Victor Tseng



Schematic of workflow for sourcing, encoding, and adjudicating results.

Evaluation and Implementation



Evaluation and Implementation

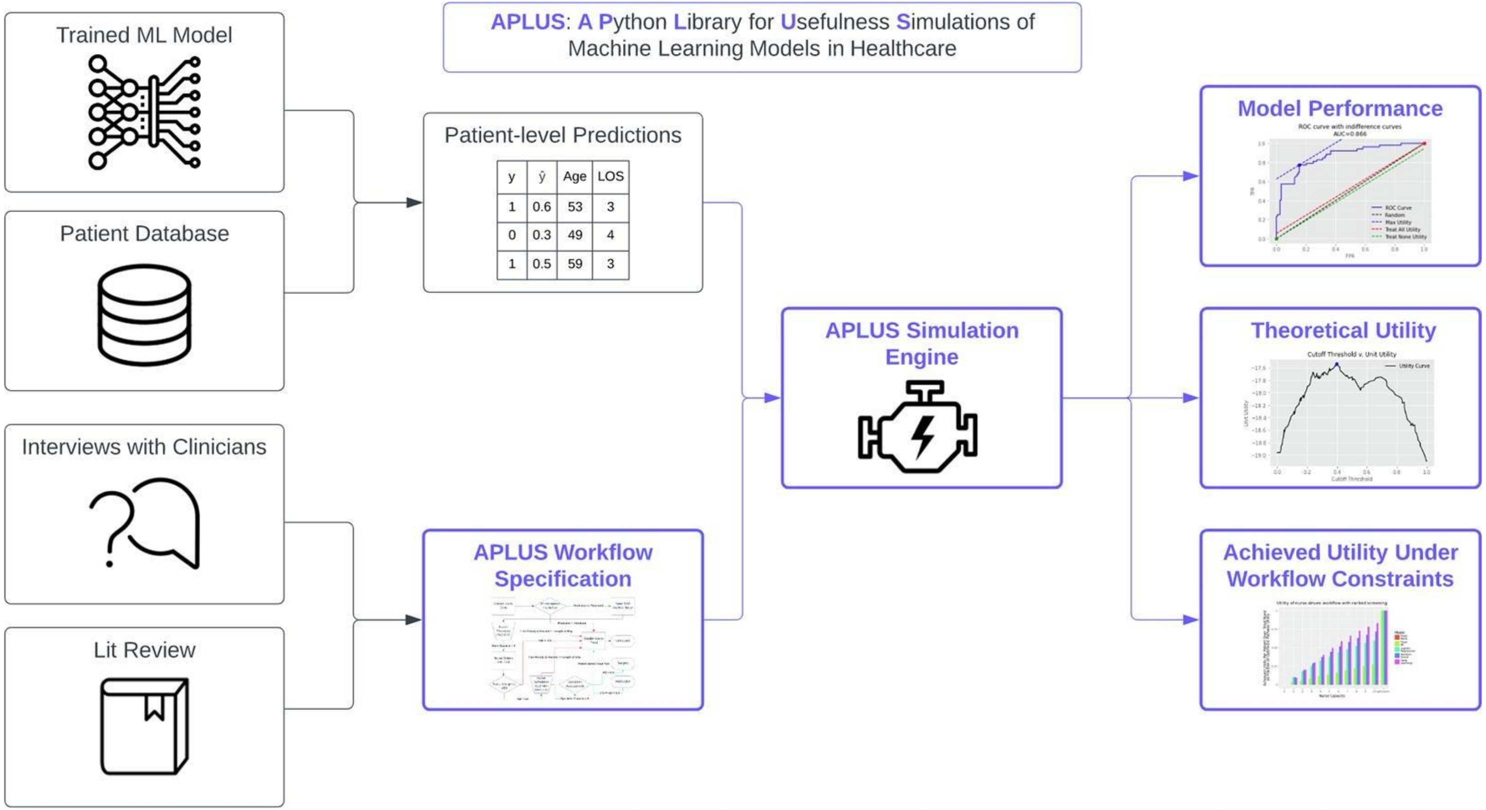
Journal of Biomedical Informatics
Volume 139, March 2023, 104319

Original Research

APLUS: A Python library for usefulness simulations of machine learning models in healthcare

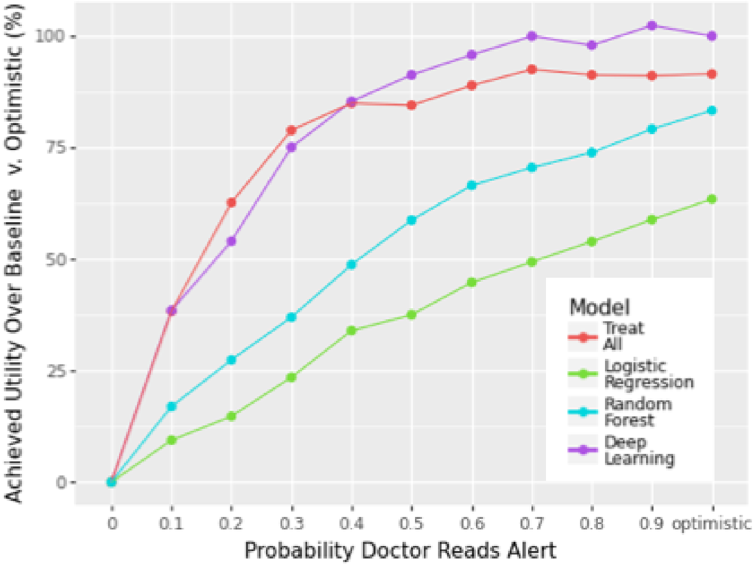
Michael Wornow [a], Elsie Gyang Ross [b c], Alison Callahan [b 1], Nigam H. Shah [b d e f 1]

Evaluation and Implementation



A **usefulness assessment** quantifies the **achievable utility** of an ML model under a **realistic** care delivery environment factoring in:

1. Model predictive performance
2. Workflow steps
3. Resource constraints
4. Utility of actions
5. Patient flow over time



AI Implementation and Evaluation

"Impact of AI-aided colonoscopy in clinical practice: a prospective randomised controlled trial." Scholer J, Alavanja M, de Lange T et al. *BMJ Open Gastroenterol*. 2024 Jan 30;11(1):e001247.

"Role of Artificial Intelligence in Colonoscopy Detection of Advanced Neoplasias : A Randomized Trial." Mangas-Sanjuan, de-Castro, Cubiella et al. *Ann Intern Med.* 2023 Sep;176(9):1145-1152. doi: 10.7326/M22-2619.

"Measuring the Impact of AI in the Diagnosis of Hospitalized Patients: A Randomized Clinical Vignette Survey Study." Jabbour S, Fouhey D, Shapard S, et al. *JAMA*. 2023 Dec 19;330(23):2275-2284. doi: 10.1001/jama.2023.22295.

"Effect of an Artificial Intelligence Decision Support Tool on Palliative Care Referral in Hospitalized Patients: A Randomized Clinical Trial." Wilson P, Ramear P, Philpot LM, et al. *J Pain Symptom Management*. 2023 Jul; 66(1) 24-32.

AI Implementation and Evaluation Best Practices

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

Artificial Intelligence in Health Care

The Hope, the Hype, the Promise, the Peril

Michael Matheny,
Sonoo Thadaney Israni, Mahnoor Ahmed,
and Danielle Whicher, *Editors*

(Technically 2022) NATIONAL ACADEMY of MEDICINE

OCTOBER 30, 2023

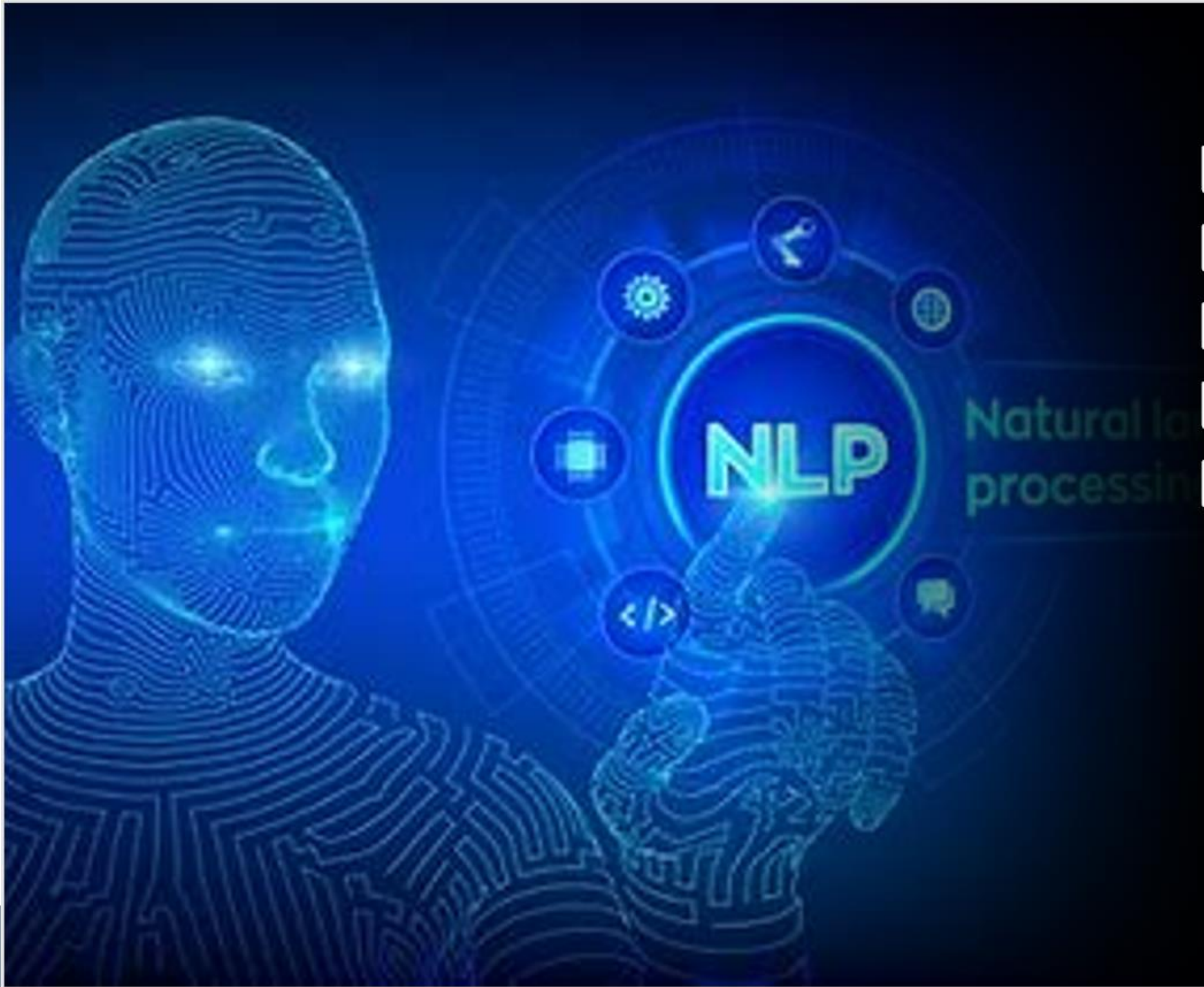
Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

BRIEFING ROOM › PRESIDENTIAL ACTIONS

Health Care Artificial Intelligence Code of Conduct

NATIONAL ACADEMY of MEDICINE

"A Nationwide Network of Health AI Assurance Laboratories." N Shah, JD Halamka, S Saria, M Pencina, T Tazbaz, M Tripathi, A Callahan, H Hildahl, B Anderson. *JAMA.* 2024 Jan 16; 331(3): 245-9. Doi:10.1001/jama.2023.26930. *More to come on this....*



Large Language Models and Natural Language Processing

LLMs and NLP

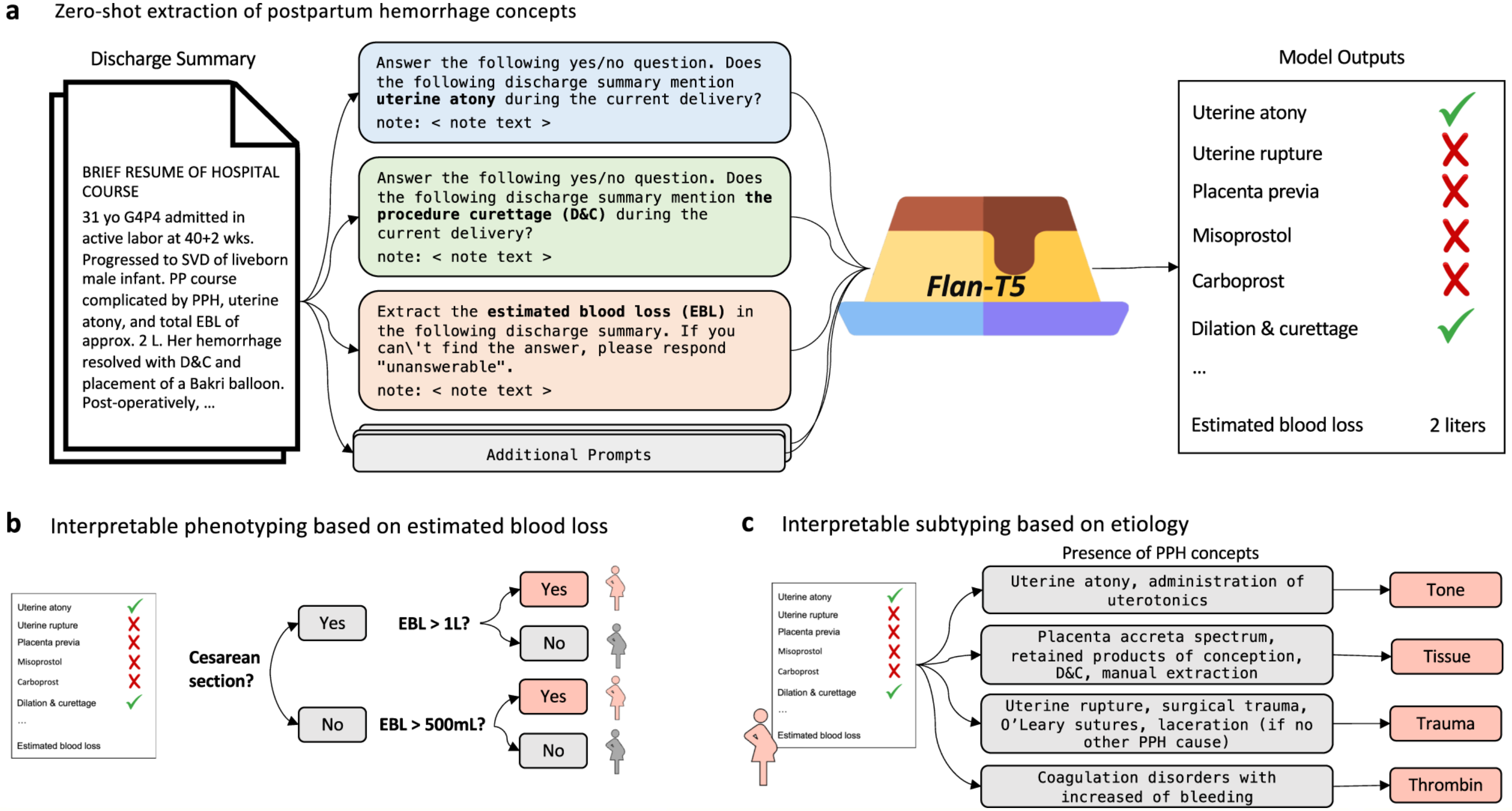
Article | Open access | Published: 30 November 2023

Zero-shot interpretable phenotyping of postpartum hemorrhage using large language models

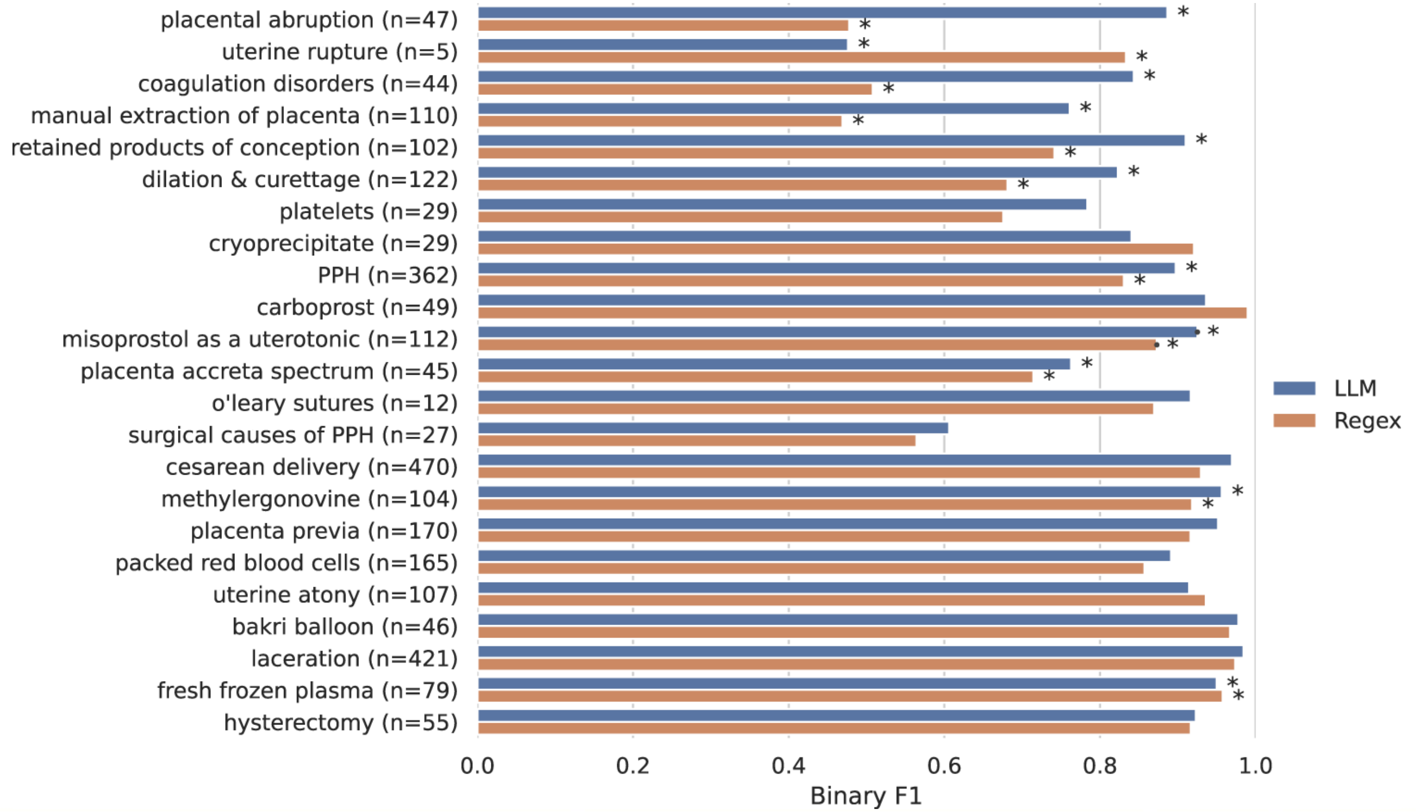
Emily Alsentzer, Matthew J. Rasmussen, Romy Fontoura, Alexis L. Cull, Brett Beaulieu-Jones, Kathryn J. Gray, David W. Bates & Vesela P. Kovacheva

npj Digital Medicine **6**, Article number: 212 (2023) | Cite this article

LLMs to identify granular disease phenotypes in clinical notes without supervising examples



Flan-T5 language model and regular expression



LLMs and NLP

Article | Open access | Published: 12 July 2023

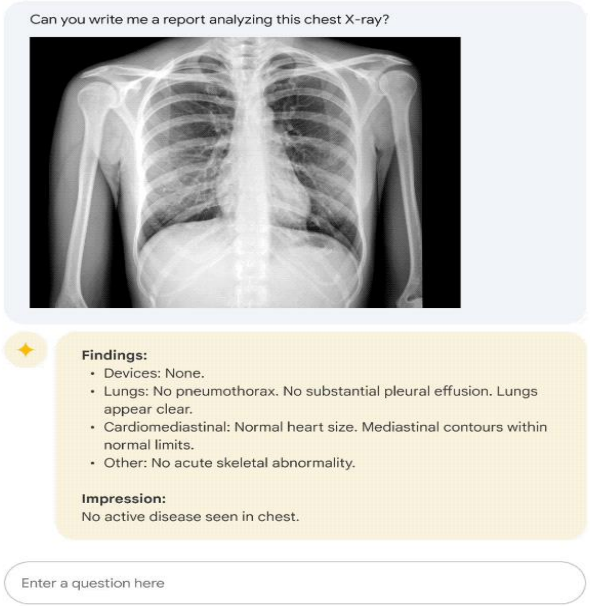
Large language models encode clinical knowledge

Karan Singhal, Shekoofeh Azizi, Tao Tu, S. Sara Mahdavi, Jason Wei, Hyung Won Chung, Nathan Scales, Ajay Tanwani, Heather Cole-Lewis, Stephen Pfohl, Perry Payne, Martin Seneviratne, Paul Gamble, Chris Kelly, Abubakr Babiker, Nathanael Schärli, Aakanksha Chowdhery, Philip Mansfield, Dina Demner-Fushman, Blaise Agüera y Arcas, Dale Webster, Greg S. Corrado, Yossi Matias, Katherine Chou, Juraj Gottweis, Nenad Tomasev, Yun Liu, Alvin Rajkomar, Joelle Barral, Christopher Semturs, Alan Karthikesalingam & Vivek Natarajan

— Show fewer authors

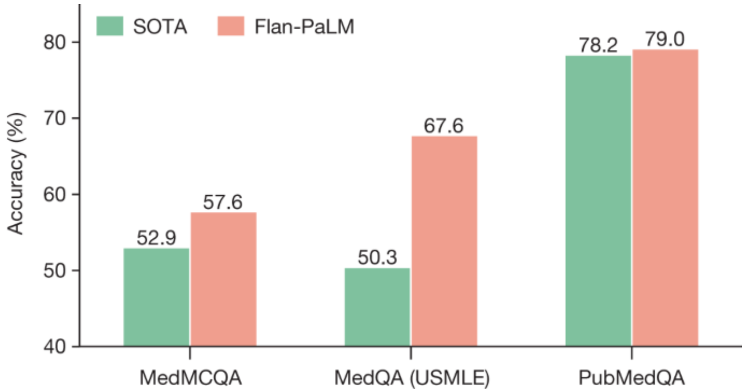
Nature **620**, 172–180 (2023) | Cite this article

Med-PaLM Multimodal

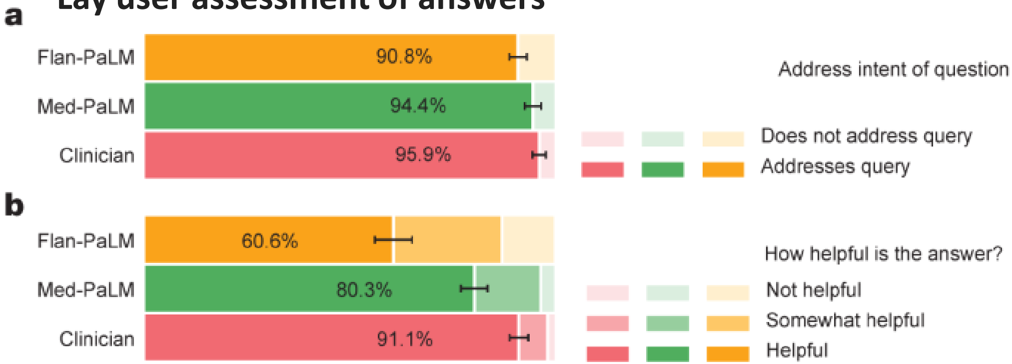


First demonstration of a **generalist biomedical AI system** that can encode and interpret modalities spanning text, imaging, EHR and genomics with a **single AI agent** and perform multiple clinical and biomedical tasks at **state-of-the-art**

Clinician evaluation of answers



Lay user assessment of answers



LLMs and NLP

Article | Open access | Published: 24 August 2023

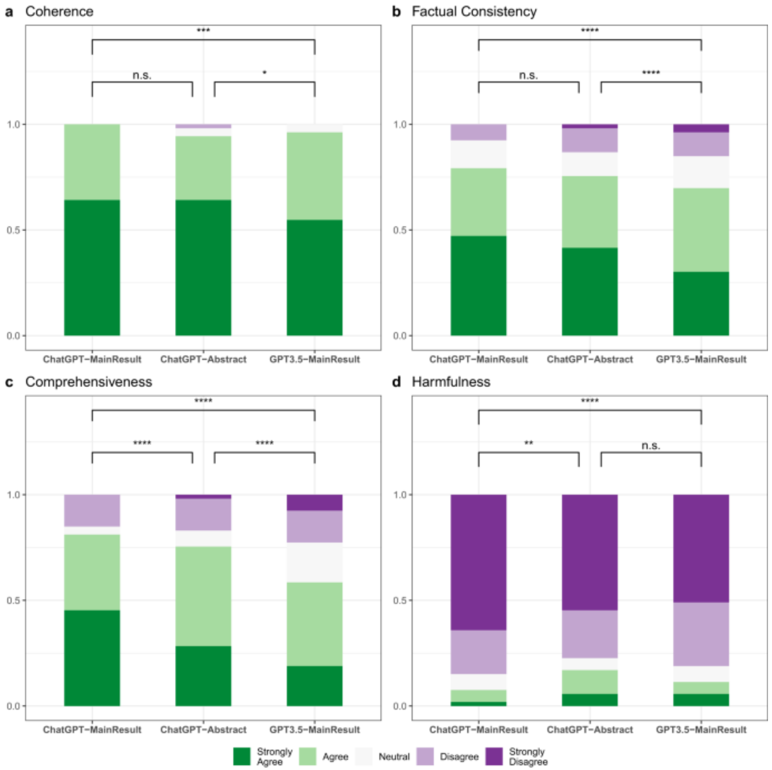
Evaluating large language models on medical evidence summarization

Liyan Tang, Zhaoyi Sun, Betina Idnay, Jordan G. Nestor, Ali Soroush, Pierre A. Elias, Ziyang Xu, Ying Ding, Greg Durrett, Justin F. Rousseau, Chunhua Weng & Yifan Peng

npj Digital Medicine **6**, Article number: 158 (2023) | Cite this article

Slides courtesy of Roxana_Daneshjou

Evaluating LLMs on medical evidence summarization



Human evaluation is essential to properly assess the quality and factuality of medical evidence summaries generated by LLMs

Four dimensions of summary quality are defined: Coherence; Factual Consistency; Comprehensiveness; and Harmfulness.

LLMs could be susceptible to generating factually inconsistent summaries and making overly convincing or uncertain statements.

More information: https://github.com/ebmlab

LLMs and NLP

THE LANCET
Digital Health

ARTICLES | VOLUME 5, ISSUE 12, E882-E894, DECEMBER 2023

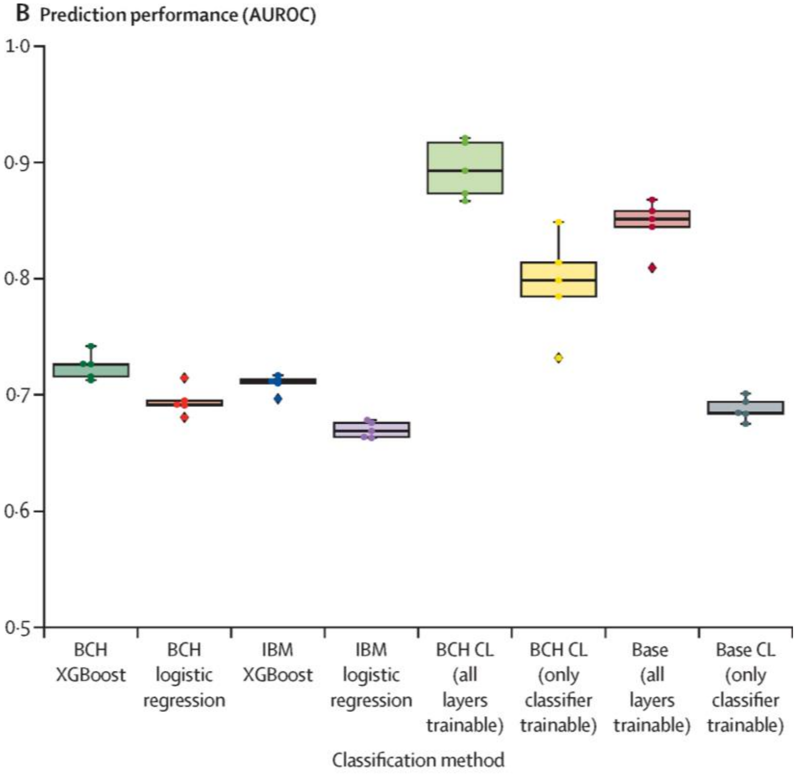
Download Full Issue

Predicting seizure recurrence after an initial seizure-like episode from routine clinical notes using large language models: a retrospective cohort study

Brett K Beaulieu-Jones, PhD • Mauricio F Villamar, MD • Phil Scordis, PhD • Ana Paula Bartmann, MD PhD • Waqar Ali, PhD • Benjamin D Wissel, MD PhD • Emily Alsentzer, PhD • Johann de Jong, PhD • Arijit Patra, PhD • Prof Isaac Kohane, MD PhD • Show less

Open Access • Published: December, 2023 • DOI: https://doi.org/10.1016/S2589-7500(23)00179-6 •

Finetuning LLMs to Predict Seizure Recurrence



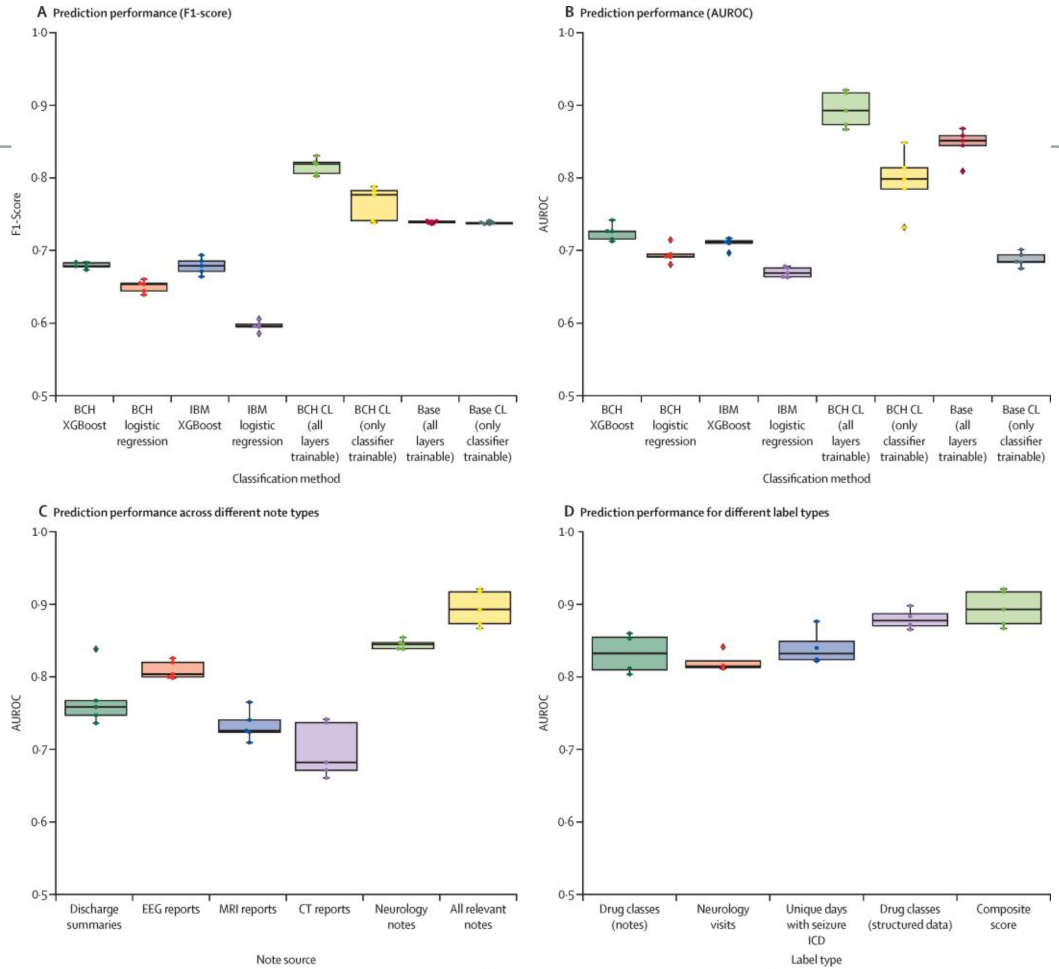
Guidelines suggest starting an anti-seizure medication after an initial event if there is a 60% or higher risk of additional seizures

Clinical intuition is hard to measure with a prospective survey, can we measure whether we are recording the right information in notes?

Base LLM (Clinical Longformer) trained on adult ICU data outperforms structured data

Domain- and Location-specific finetuning further improves performance

BUT there are many more steps before considering deployment of such a system



A Prediction performance (F1-score)
F1-Score
Classification method
BCH XGBoost
BCH logistic regression
IBM XGBoost
IBM logistic regression
BCH CL (all layers trainable)
BCH CL (only classifier trainable)
Base (all layers trainable)
Base CL (only classifier trainable)
B Prediction performance (AUROC)
AUROC
C Prediction performance across different note types
Discharge summaries
EEG reports
MRI reports
CT reports
Neurology notes
All relevant notes
Note source
D Prediction performance for different label types
Drug classes (notes)
Neurology visits
Unique days with seizure ICD
Drug classes (structured data)
Composite score
Label type

LLMs and NLP

JOURNAL ARTICLE

An open natural language processing (NLP) framework for EHR-based clinical research: a case demonstration using the National COVID Cohort Collaborative (N3C)

Sijia Liu, Andrew Wen, Liwei Wang, Huan He, Sunyang Fu, Robert Miller, Andrew Williams, Daniel Harris, Ramakanth Kavuluru, Mei Liu, Noor Abu-el-Rub, Dalton Schutte, Rui Zhang, Masoud Rouhizadeh, John D Osborne, Yongqun He, Umit Topaloglu, Stephanie S Hong, Joel H Saltz, Thomas Schaffter, Emily Pfaff, Christopher G Chute, Tim Duong, Melissa A Haendel, Rafael Fuentes, Peter Szolovits, Hua Xu, Hongfang Liu ✉ on behalf of National COVID Cohort Collaborative (N3C) Natural Language Processing (NLP) Subgroup, National COVID Cohort Collaborative (N3C)

Author Notes

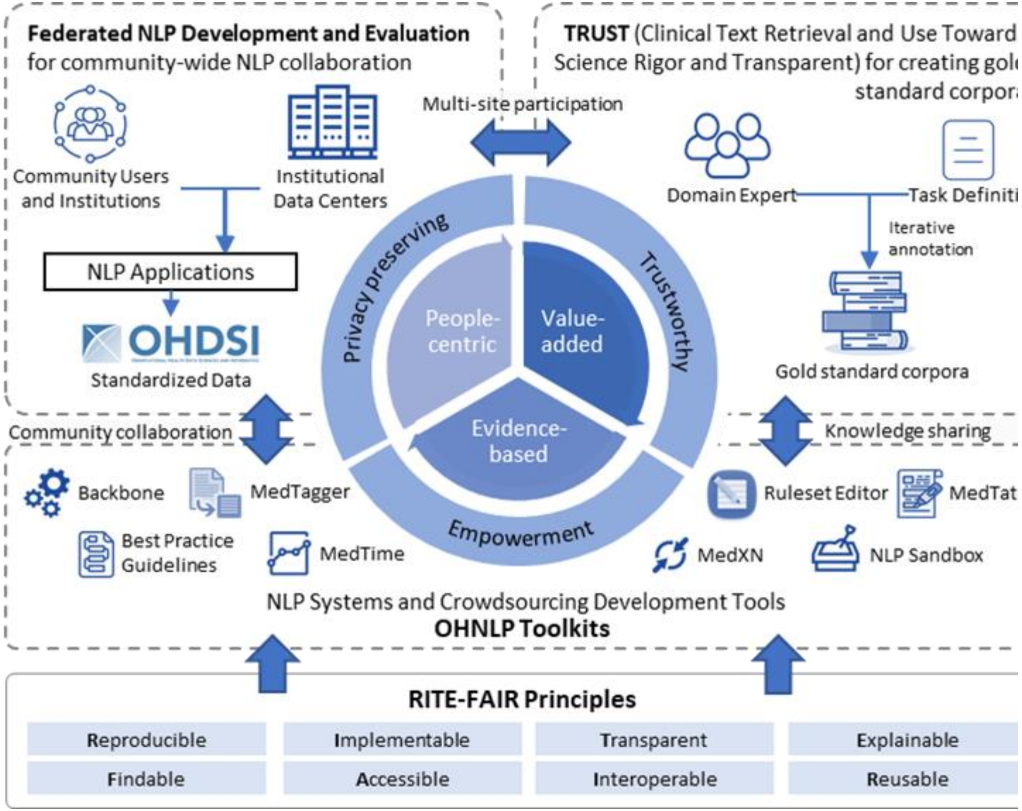
Journal of the American Medical Informatics Association, Volume 30, Issue 12, December 2023, Pages 2036–2040, https://doi.org/10.1093/jamia/ocad134

Published: 09 August 2023 **Article history** ▾

OHNLP Toolkit

- Despite Rapid Advances in Clinical NLP Methodologies, Practical Adoption of Clinical NLP is Greatly Hindered by Process Heterogeneity and Human Factor Variations
- This Paper Presents a Use Case Demonstration of the Open Health Natural Language Processing Toolkit (**OHNLPTK**), Developed in Large Part to Help Address These Issues, to Support COVID-19 Sign/Symptom Extraction as part of the National COVID Cohort Collaborative (**N3C**)
- Demonstrates that Multi-Site Engagement and Data Contribution During the Development Process Substantially Increases Algorithm Robustness
- Highlights Necessity of Federated Annotation and Evaluation Efforts in NLP Development

OHNLP Toolkit



- Focused around the RITE-FAIR principles throughout the NLP Algorithm Development and Evaluation Process
- Consists of:
  - Federated NLP Development and Evaluation Framework
  - Privacy-Preserving Scientifically Rigorous and Transparent Gold Standard Development Toolchain
  - Framework for Algorithm Provenance and Alteration Retention

Slides courtesy of Hongfang Liu

Original Investigation | Statistics and Research Methods

March 15, 2023

Associations Between Natural Language Processing–Enriched Social Determinants of Health and Suicide Death Among US Veterans

Avijit Mitra, MSc[1]; Richeek Pradhan, MD[2]; Rachel D. Melamed, PhD[3]; Kun Chen, PhD[4,5]; David C. Hoaglin, PhD[6]; Katherine L. Tucker, PhD[7]; Joel I. Reisman, AB[8]; Zhichao Yang, MS[1]; Weisong Liu, PhD[9,10]; Jack Tsai, PhD[11,12]; Hong Yu, PhD[1,8,9,10]

» Author Affiliations | Article Information

JAMA Netw Open. 2023;6(3):e233079. doi:10.1001/jamanetworkopen.2023.3079

NLP-extracted SBDH Categories

SBDH	Example tokens	SDOH
Social isolation	Alone, lonely, divorce, widow etc.	Yes
Transition of care	Discharge, admission, change in medication, transfer etc.	Yes
Barriers to care	Transportation issues, garbled speech, communication problems etc.	Yes
Financial insecurity	Unemployed, poor, unemployment, rehabilitation etc.	Yes
Housing instability	Eviction, homeless, homelessness etc.	Yes
Food insecurity	Hungry, pantry, starvation, food voucher etc.	Yes
Violence	Firearms, violence, assault, weapon, abuse, homicidal, racism etc.	Yes
Legal problems	Imprisonment, parole, arrested, felony, investigation, prison etc.	Yes
Substance abuse	Alcohol, tobacco, heroin, cocaine, smoking, overdose etc.	No
Psychiatric symptoms	PTSD, depression, anxiety, schizophrenia, insomnia, hallucination etc.	No
Pain	Pain, suffering, hurting, discomfort etc.	No
Patient disability	Disabled, blind, hearing loss, wheelchair etc.	No

Results

AMIA INFORMATICS PROFESSIONALS. LEADING THE WAY.

- All 8 NLP-extracted SDOH were significantly associated with increased risk of death by suicide
- All structured and combined SDOHs also showed significant associations
- Legal problems had the largest estimated effect size across all three groups of SDOHs
- Structured SDOHs consistently showed higher aORs for suicide than NLP-extracted SDOHs

Table 3. Associations of SDOH With Veterans' Death by Suicide

SDOH factors	aOR (95% CI)[a] NLP-extracted	Structured	Combined
Social problems[b]	1.94 (1.83-2.06)	2.11 (1.94-2.29)	1.95 (1.84-2.07)
Financial problems[c]	1.91 (1.79-2.04)	2.18 (1.97-2.42)	1.92 (1.80-2.05)
Housing instability	1.90 (1.78-2.03)	2.28 (2.06-2.53)	1.93 (1.80-2.06)
Legal problems	2.62 (2.38-2.89)	2.63 (2.37-2.91)	2.66 (2.46-2.89)
Violence	2.34 (2.17-2.52)	1.96 (1.77-2.16)	2.12 (1.98-2.27)
Barriers to care	1.86 (1.74-1.99)	NA	1.86 (1.74-1.98)
Transition of care	1.53 (1.44-1.62)	NA	1.51 (1.43-1.60)
Food insecurity	1.85 (1.62-2.11)	NA	1.85 (1.62-2.11)
Nonspecific psychosocial needs	NA	2.09 (1.94-2.25)	2.07 (1.92-2.23)

Abbreviations: aOR, adjusted odds ratio; NA, not applicable; NLP, natural language processing; SDOH, social determinants of health.

[a] Each model was adjusted for sociodemographic variables, psychiatric symptoms, substance abuse, pain, patient disability, clinical comorbidities, and all SDOH in its group.

[b] Social problems indicates social or familial problems from structured data with social isolation from NLP-extracted data.

[c] Financial problems indicates employment or financial problems from structured data with job or financial insecurity from NLP-extracted data.

Ethical, Legal, and Societal Issues

Ethical, Legal and Social Issues

Journal of Medical Internet Research | Journal Information | Browse Journal

Published on 16.11.2023 in Vol 25 (2023)

Preprints (earlier versions) of this paper are available at https://preprints.jmir.org/preprint/47609, first published March 26, 2023.

Developer Perspectives on Potential Harms of Machine Learning Predictive Analytics in Health Care: Qualitative Analysis

Ariadne A Nichol [1]; Pamela L Sankar [2]; Meghan C Halley [1]; Carole A Federico [1]; Mildred K Cho [1]

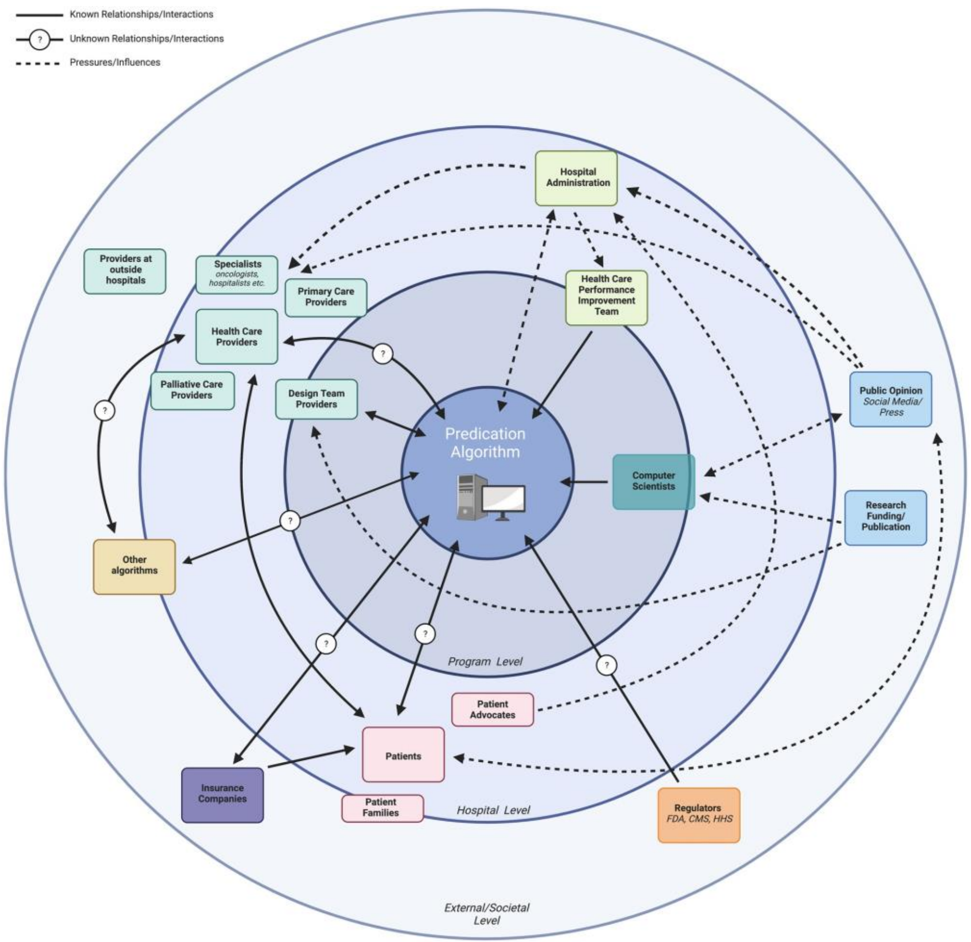
Ethical, Legal and Social Issues

JOURNAL ARTICLE FEATURED

A framework to identify ethical concerns with ML-guided care workflows: a case study of mortality prediction to guide advance care planning

Diana Cagliero, Natalie Deuitch, Nigam Shah, Chris Feudtner, Danton Char

Journal of the American Medical Informatics Association, Volume 30, Issue 5, May 2023, Pages 819–827, https://doi.org/10.1093/jamia/ocad022



AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

Ethical, Legal and Social Issues

JOURNAL ARTICLE EDITOR'S CHOICE FEATURED

Integrating patient voices into the extraction of social determinants of health from clinical notes: ethical considerations and recommendations FREE

Andrea L Hartzler ✉, Serena Jinchen Xie, Patrick Wedgeworth, Carolin Spice, Kevin Lybarger, Brian R Wood, Herbert C Duber, Gary Hsieh, Angad P Singh, SDoH Community Champion Advisory Board Author Notes

Journal of the American Medical Informatics Association, Volume 30, Issue 8, August 2023, Pages 1456–1462, https://doi.org/10.1093/jamia/ocad043

Ethical considerations - "AI4People"

Extracting social drivers of health (SDoH) from clinical notes could have unintended consequences! Automated SDoH systems should promote ...

Beneficence

Benefit and empower people by promoting well-being and preserving dignity. Could automation:

- *Alleviate patient discomfort with SDoH screening?*
- *Reduce clinician burden of identifying actionable SDoH?*

Nonmaleficence

Avoid harm by preserving privacy, confidentiality, security, preventing data misuse, and mitigating inaccurate data generated by AI. Could automation:

- *Harm patients with stigmatizing labels?*
- *Inhibit patient-clinician trust?*
- *Lead to false positive or false negative SDoH screens?*

Autonomy

Provide people freedom to make decisions for themselves, including how much agency to delegate to AI. Could automation:

- *Override a patient's desire not to receive assistance?*
- *Allow patients the freedom to opt in or opt out?*

Justice

Be fair and equitable by promoting prosperity and preserving solidarity, without disempowering people when imperfect mining leads to the unintentional creation of biased data that can reinforce inequities. Could automation:

- *Mislabel patients in stigmatizing ways?*
- *Generate biased SDoH that reinforce inequities?*

Explicability

Exhibit transparency and accountability while supporting other 4 traditional bioethics principles.

Automated SDoH systems should:

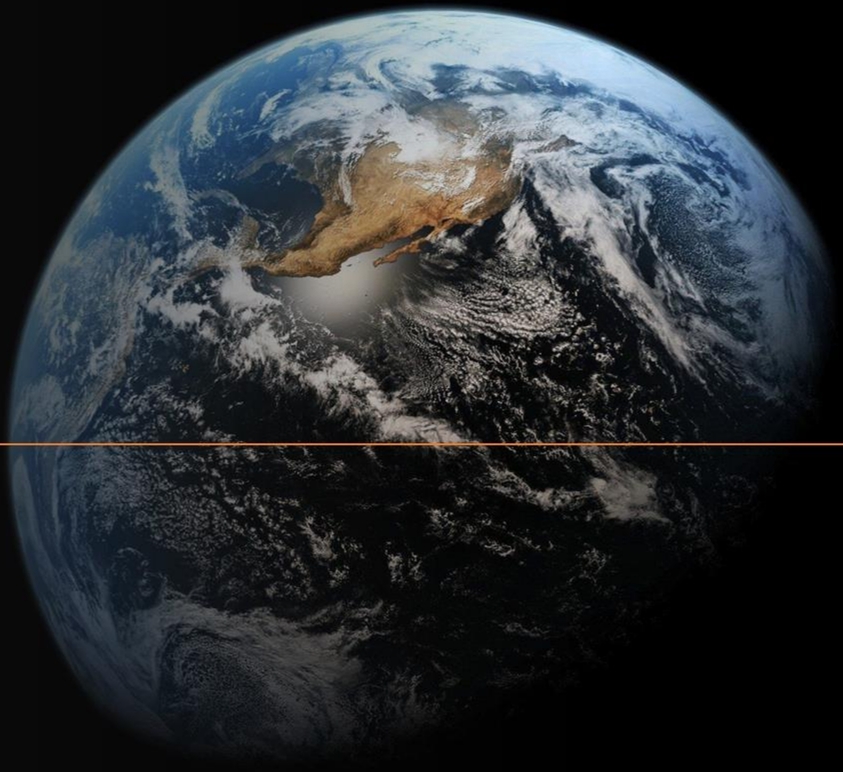
- *Inform patients of potential benefits and harms*
- *Allow patients to choose whether/how extracted data is used*
- *Prevent the perpetuation of bias*
- *Ensure organizational accountability for unintended consequences*

Recommendations to integrate patient voice

Patient acceptability of automated SDoH systems is contingent on mitigating patient concerns, particularly for individuals with social needs who may be impacted most.

1. **Engage patients inclusively**
   - Find individuals with lived experience
   - Work with people from traditionally marginalized and underrepresented groups
   - Diversity comes in many forms!
2. **Engage patients with transparency**
   - Make participation accessible
   - Consider literacy, both health and digital
3. **Engage patients cooperatively**
   - Treat as partners in co-production
   - Fair compensation for effort and expertise
   - Democratic participation

Public and Global Health Informatics



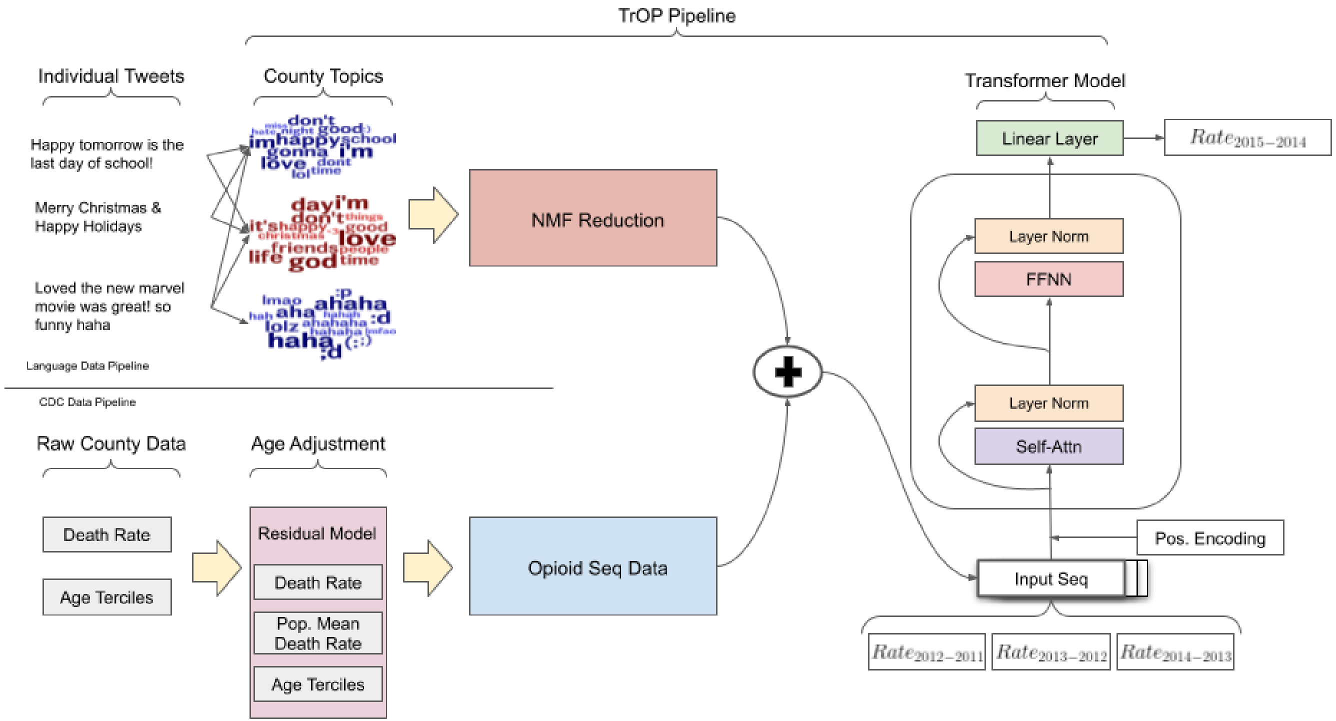
Article | Open access | Published: 08 March 2023

Opioid death projections with AI-based forecasts using social media language

Matthew Matero, Salvatore Giorgi, Brenda Curtis, Lyle H. Ungar & H. Andrew Schwartz

npj Digital Medicine **6**, Article number: 35 (2023) | Cite this article

TROP – TRansformer for Opioid Prediction



Results – Overall Performance and Language

AMIA INFORMATICS PROFESSIONALS. LEADING THE WAY.

Model (history)	MAPE	MAE
Baselines		
Last (1)	16.16	5.76
Mean (4)	36.61	11.67
Linear		
Ridge AR(3)	6.31	2.63
Deep Learning		
Recurrent Neural Net (3)	3.99	1.64
TROP (3)	**2.92***	**1.15***

Model (history)	MAPE	MAE
With Language		
Ridge AR(3)	6.31	2.63
Recurrent Neural Net(3)	3.99	1.64
TROP (3)	**2.92***	**1.15***
Without Language		
Ridge AR (3)	7.09	2.93
Recurrent Neural Net (3)	6.97	2.84
TROP (3)	**6.93***	**2.81***

Public and Global Health Informatics

THE LANCET
Global Health

Wastewater monitoring can anchor global disease surveillance systems

Aparna Keshaviah, ScM * • Megan B Diamond, SM * • Matthew J Wade, PhD • Prof Samuel V Scarpino, PhD •
on behalf of the Global Wastewater Action Group • Show footnotes

Open Access • Published: June, 2023 • DOI: https://doi.org/10.1016/S2214-109X(23)00170-5 • Check for updates

Slides courtesy of Samuel Scarpino. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of The Rockefeller Foundation

Wastewater Surveillance can be a powerful public health tool

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

Why it Matters

With limited insights on what pathogens are circulating and where, policymakers are forced to make decisions with a partial picture of community health. Continuous monitoring of wastewater, complemented by other data sources, will give clarity to on ongoing and impending threats, resulting in more effective interventions.

200 Million

people are predicted to be displaced by climate hazards by 2050, often making it harder for them to access health care facilities

>25 Pathogens

hundreds of chemicals, and antimicrobial-resistant genes can be detected in wastewater

3-7 Days

advance warning provided by wastewater surveillance can stop disease spread

The Rockefeller Foundation

https://www.rockefellerfoundation.org/initiative/wastewater-surveillance/

Wide adoption of wastewater surveillance

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

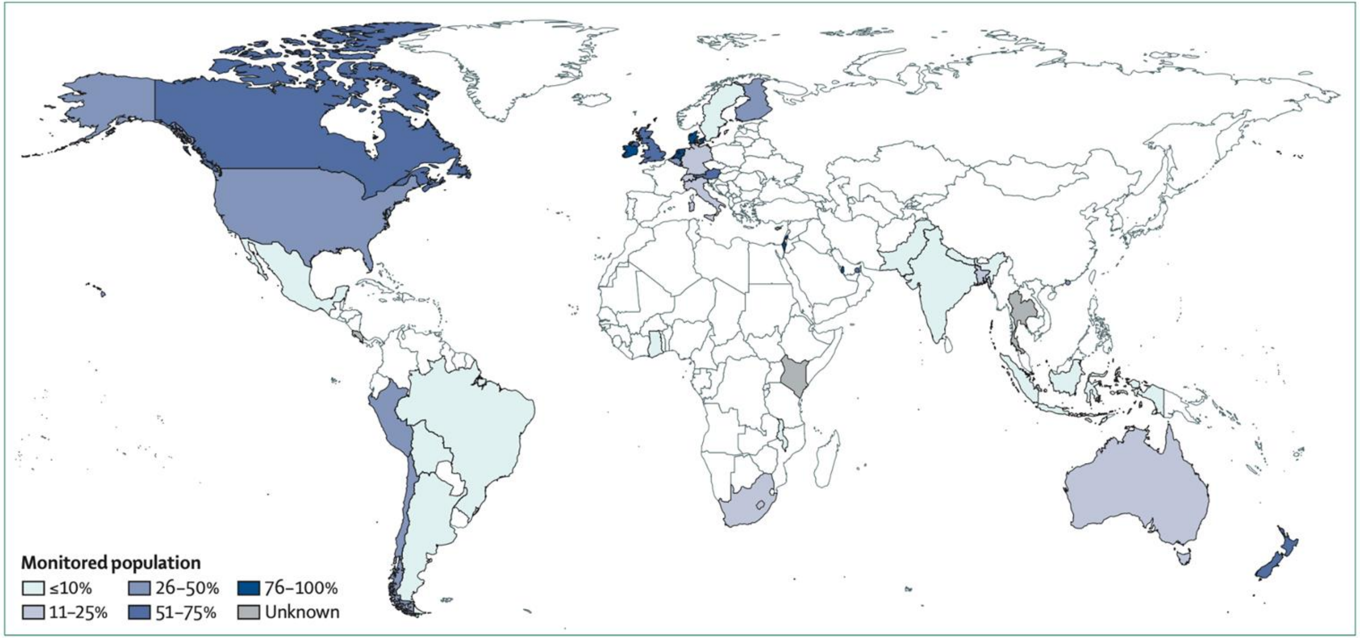


***Figure:* Population coverage of wastewater monitoring programmes in surveyed countries**
This map shows the percentage of a country's population that was monitored through wastewater testing for the 43 (shaded) countries that responded to the wastewater survey.

	LMICs (n=16)	HICs (n=27)
Genomic targets measured by lab		
N1	69%	70%
N2	50%	48%
N unspecified	13%	19%
E	31%	30%
S	0	15%
Orf	19%	26%
Other	13%	19%
Laboratory method for SARS-CoV-2 measurement		
qPCR	94%	96%
ddPCR	6%	37%
Other	13%	7%
Method used to monitor SARS-CoV-2 variants		
PCR	50%	74%
Genomic sequencing	75%	85%
Not applicable (variants not monitored)	50%	0
Frequency of testing for SARS-CoV-2 variants		
Just a few samples	13%	7%
More than few samples, but not regularly	13%	15%
Regularly	13%	59%
Not reported	13%	19%
Not applicable (variants not monitored)	50%	0

This table shows the variety of approaches that wastewater monitoring programmes use to detect or quantify the SARS-CoV-2 virus in wastewater, or both. Respondents could select more than one response for questions about the genomic targets measured and laboratory methods used. For frequency of testing, percentages might add up to more than 100% due to rounding. ddPCR=digital droplet PCR. HICs=high-income countries. LMICS=low-income and middle-income countries. qPCR=quantitative PCR.

***Table:* Characteristics of wastewater testing methods**

Slides courtesy of Samuel Scarpino. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of The Rockefeller Foundation

Public and Global Health Informatics

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

PNAS

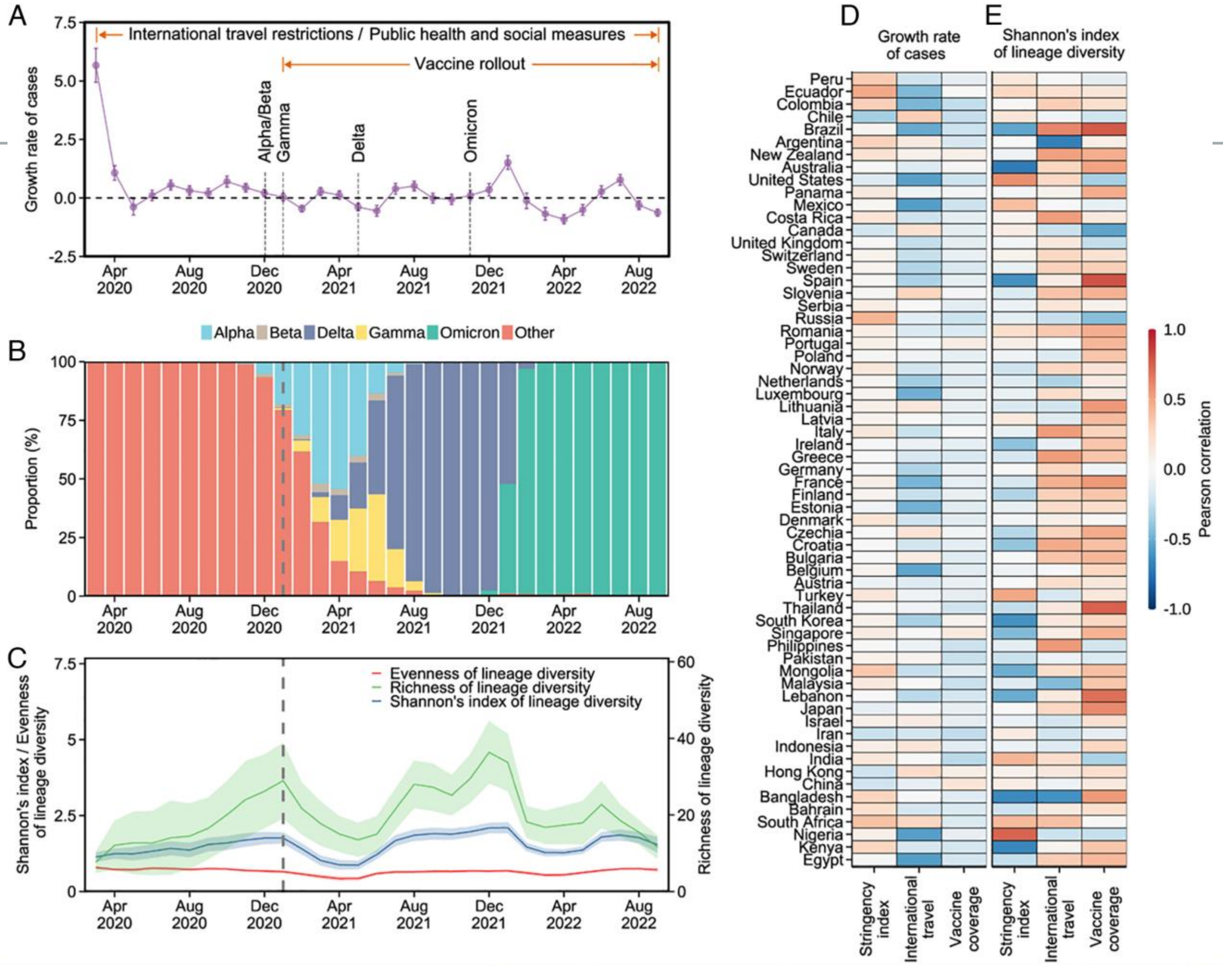
RESEARCH ARTICLE | ENVIRONMENTAL SCIENCES

OPEN ACCESS

Check for updates

Association of vaccination, international travel, public health and social measures with lineage dynamics of SARS-CoV-2

Lingyue Yang[a,1], Zengmiao Wang[a,1], Lin Wang[b,1], Bram Vrancken[c,d], Ruixue Wang[a], Yuanlong Wei[a], Benjamin Rader[e,f], Chieh-Hsi Wu[g], Yuyang Chen[a], Peiyi Wu[a], Bingying Li[a], Qiushi Lin[a], Lu Dong[h], Yujun Cui[i], Mang Shi[j], John S. Brownstein[d,k], Nils Chr. Stenseth[l,m,n,2], Ruifu Yang[i,2], and Huaiyu Tian[a,2]


A
International travel restrictions / Public health and social measures
Vaccine rollout
Alpha/Beta
Gamma
Delta
Omicron
Growth rate of cases
B
Alpha Beta Delta Gamma Omicron Other
Proportion (%)
C
Evenness of lineage diversity
Richness of lineage diversity
Shannon's index of lineage diversity
Shannon's index / Evenness of lineage diversity
Richness of lineage diversity
D
Growth rate of cases
E
Shannon's index of lineage diversity
Stringency index
International travel
Vaccine coverage
Pearson correlation

A Year of Remarkable Progress in AI and Data Science for Healthcare

- Transformative advancements across multiple healthcare domains
- From the emulation of large-scale trials to federated causal inference, the field has pushed the boundaries of real-world evidence and causal inference
- Advances in multi-modal AI—from notes to images to omics—have enhanced diagnostic accuracy and our understanding of complex diseases
- Innovations like the generative models for synthesizing complex proteins have demonstrated the AI's impact on improving diagnostics and treatment
- The evolution of LLMs has revolutionized natural language processing capabilities within the healthcare domain

Ahead: Collaborative AI and Data Science

- Model-model collaboration
  - The integration of multimodal data sources promise to unlock deeper insights into patient health and disease progression
- Human-model collaboration
  - The advancement of large language models opens the door wider for human-model collaboration
  - Foundation model as a colleague, ethical considerations
- Human-human collaboration
  - Collaborative efforts between clinicians, researchers, and technologists
  - Partnerships between academia and industry will be crucial in overcoming existing challenges and unlocking the full potential of AI and data science to revolutionize healthcare