

**eHealth Lab**  
ehealthlab.cci.fsu.edu



# Harnessing Explainable, Equitable, and Actionable Data Science to Improve Health

**Zhe He**, PhD, FAMIA

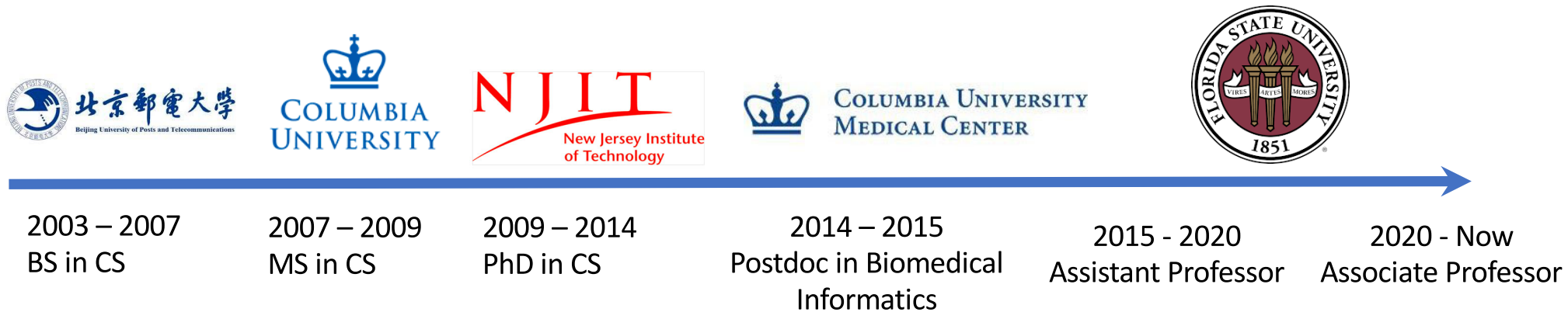
Associate Professor, School of Information  
Director of Biostatistics, Informatics and Research Design Program, UF-FSU CTSA  
Director of eHealth Lab  
Florida State University

November 20, 2023 at Moffitt Cancer Center

 DrZheHe



# Academic Journey

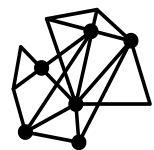


2019 – 2023, Program Chair of MSIT Program (#5 USNW)

2019 –, Director of Biostatistics, Informatics, and Research Design (BIRD) Program, UF-FSU CTSA Hub

Chair-Elect/Chair (2024-2026), Vice Chair (2022-2023), Secretary (2019–2022), AMIA Knowledge Discovery and Data Mining Working Group





- **Overarching goal:** Improving population health and advancing biomedical research through the application of informatics and electronic health data from heterogeneous sources

R01AG064529  
FSU Intramural

ARTIFICIAL  
INTELLIGENCE FOR  
AGING

CLINICAL  
RESEARCH  
INFORMATICS

R21AG061431  
R21AG061431-02S1

R21LM013911  
P01AA029547

EXPLAINABLE AI IN  
MEDICINE

CONSUMER HEALTH  
INFORMATICS

FSU ISL Grant



# Outline

- Promises of Data Science in Healthcare
- Challenges of Data Science in Healthcare and Possible Pathways Forward



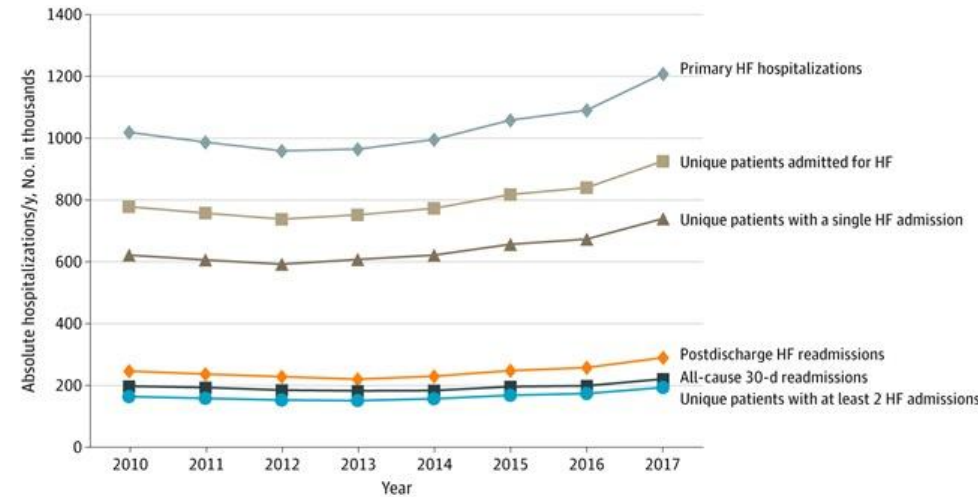
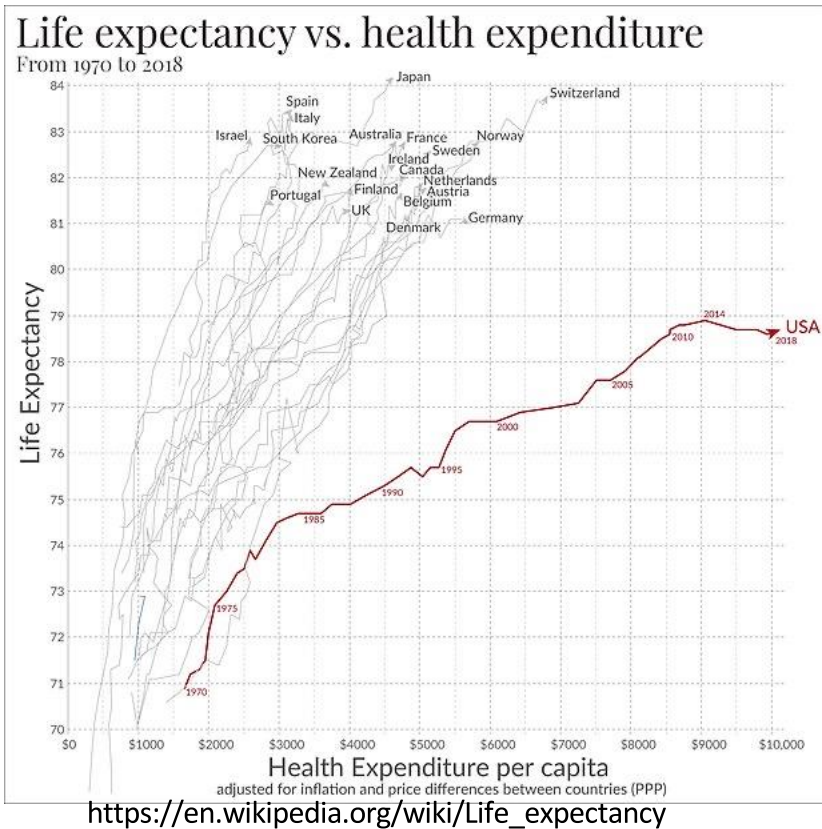
# Outline

- Promises of Data Science in Healthcare
- Challenges of Data Science in Healthcare and Possible Pathways Forward



# Health Expenditure vs. Outcomes

US Healthcare Spending in 2022: \$4.5 trillion (\$13,677 per capita)



No. at risk	2010	2011	2012	2013	2014	2015	2016	2017
Primary HF hospitalizations	1017475	985034	957509	962819	994099	1057652	1090348	1208334
Unique patients admitted for HF	776307	754532	735372	748456	769795	815499	837707	924066
Unique patients with a single HF admission	618462	602688	588451	603915	618575	653953	669875	736707
Postdischarge HF readmissions	241168	230503	222137	214362	224303	242152	252641	284269
All-cause 30-d readmissions	191819	187197	178884	176494	177448	189633	192684	214803
Unique patients with at least 2 HF admissions	157845	151844	146921	144541	151220	161546	167832	187359

Agarwal (2021). JAMA. 10.1001/jamacardio.2020.7472



# Precision Medicine Initiative



BRIEFING ROOM

ISSUES

THE ADMINISTRATION

1600 PENN

Search



## THE PRECISION MEDICINE INITIATIVE



The NEW ENGLAND  
JOURNAL of MEDICINE

### Perspective

#### A New Initiative on Precision Medicine

Francis S. Collins, M.D., Ph.D., and Harold Varmus, M.D.

N Engl J Med 2015; 372:793-795 | February 26, 2015 | DOI: 10.1056/NEJMp1500523

Comments open through March 4, 2015

Share:

Article   References   Citing Articles (784)   Comments (7)   Metrics

"Tonight, I'm launching a new Precision Medicine Initiative to bring us closer to curing diseases like cancer and diabetes — and to give all of us access to the personalized information we need to keep ourselves and our families healthier."

— President Barack Obama, State of the Union Address, January 20, 2015

President Obama has long expressed a strong conviction that science offers great potential for improving health. Now, the President has announced a research initiative that aims to accelerate progress toward a new era of precision medicine ([www.whitehouse.gov/precisionmedicine](http://www.whitehouse.gov/precisionmedicine)). We believe that the time is right for this visionary initiative, and the National Institutes of Health (NIH) and other partners will work to achieve this vision.

#### AUDIO INTERVIEW



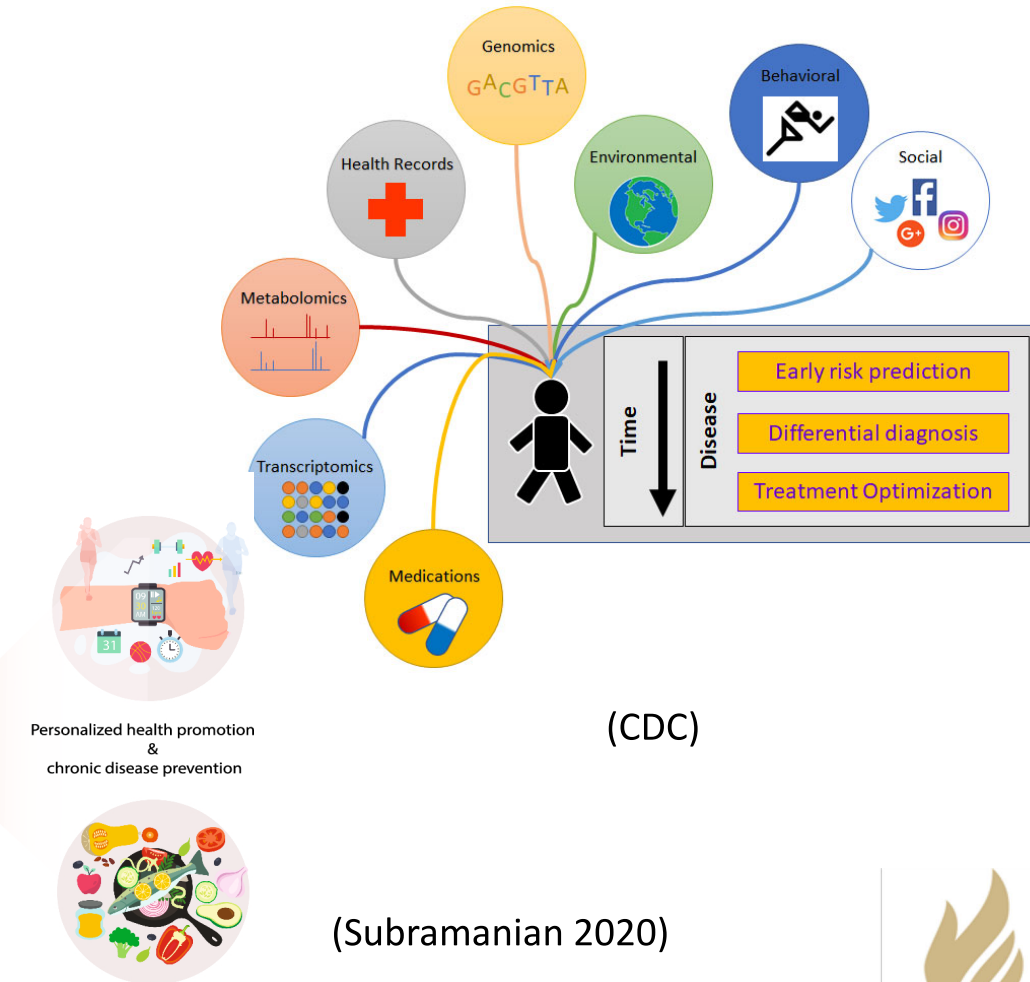
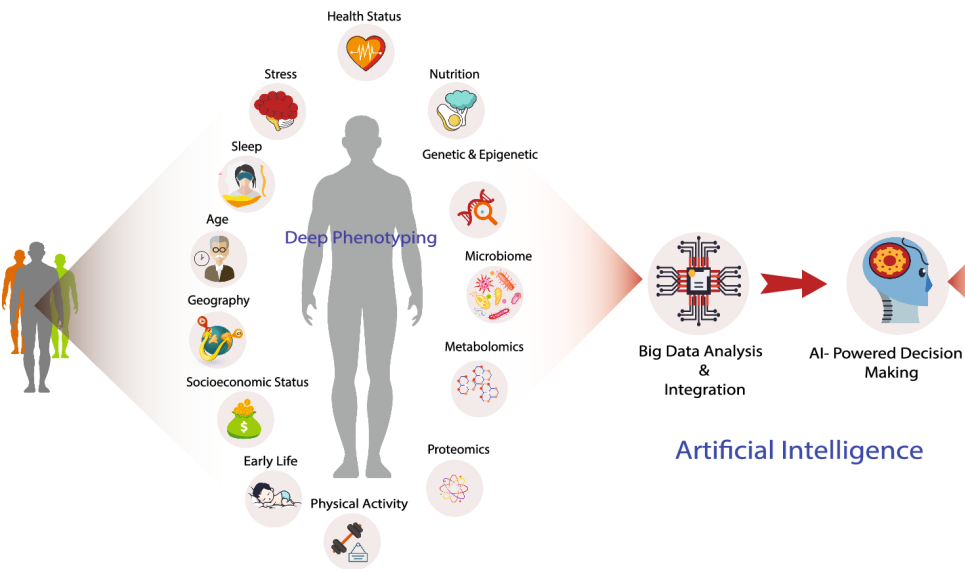
Interview with Dr. Francis Collins on what to expect from the recently announced Precision Medicine Initiative. (10:07)

Listen  
 Download

"The initiative will encourage and support the next generation of scientists to develop creative new approaches for detecting, measuring, and analyzing a wide range of biomedical information — including molecular, genomic, cellular, clinical, behavioral, physiological, and environmental parameters"

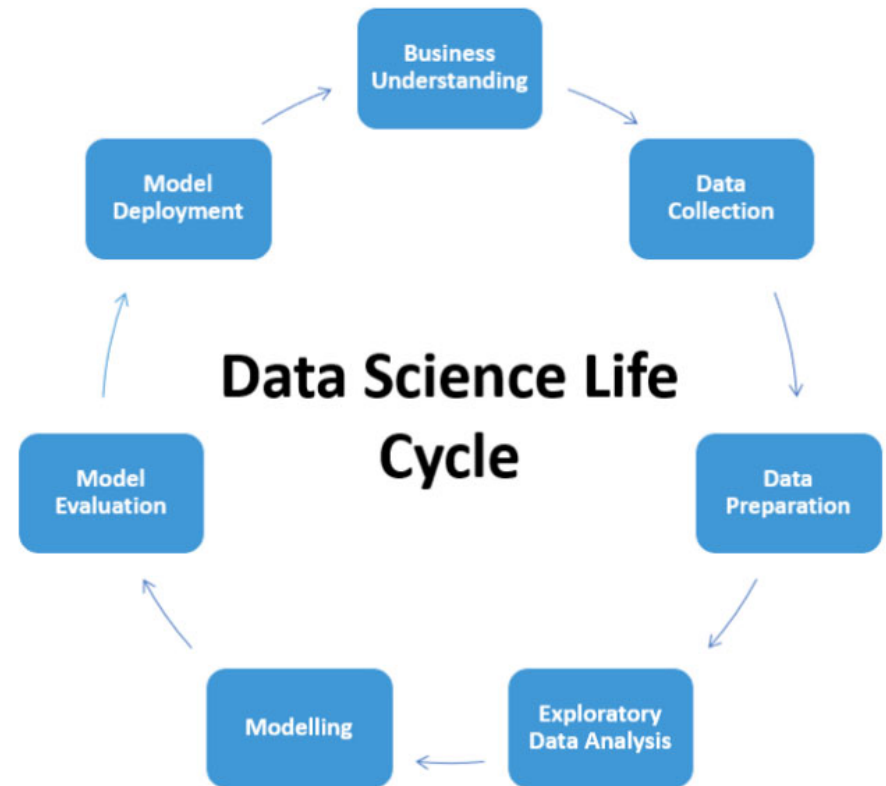
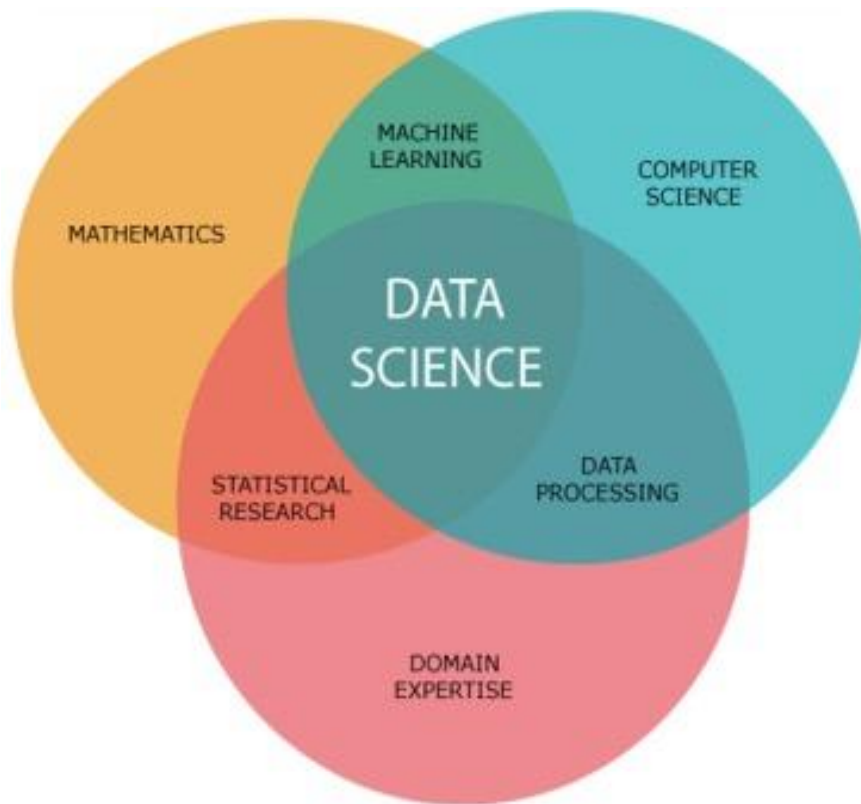
# AI-enabled Precision Health

- AI is an important enabler of precision health
- Primarily for big data analytics
  - Analyze large medical datasets
  - Find new correlations based on existing precedence
  - Support clinical decision making



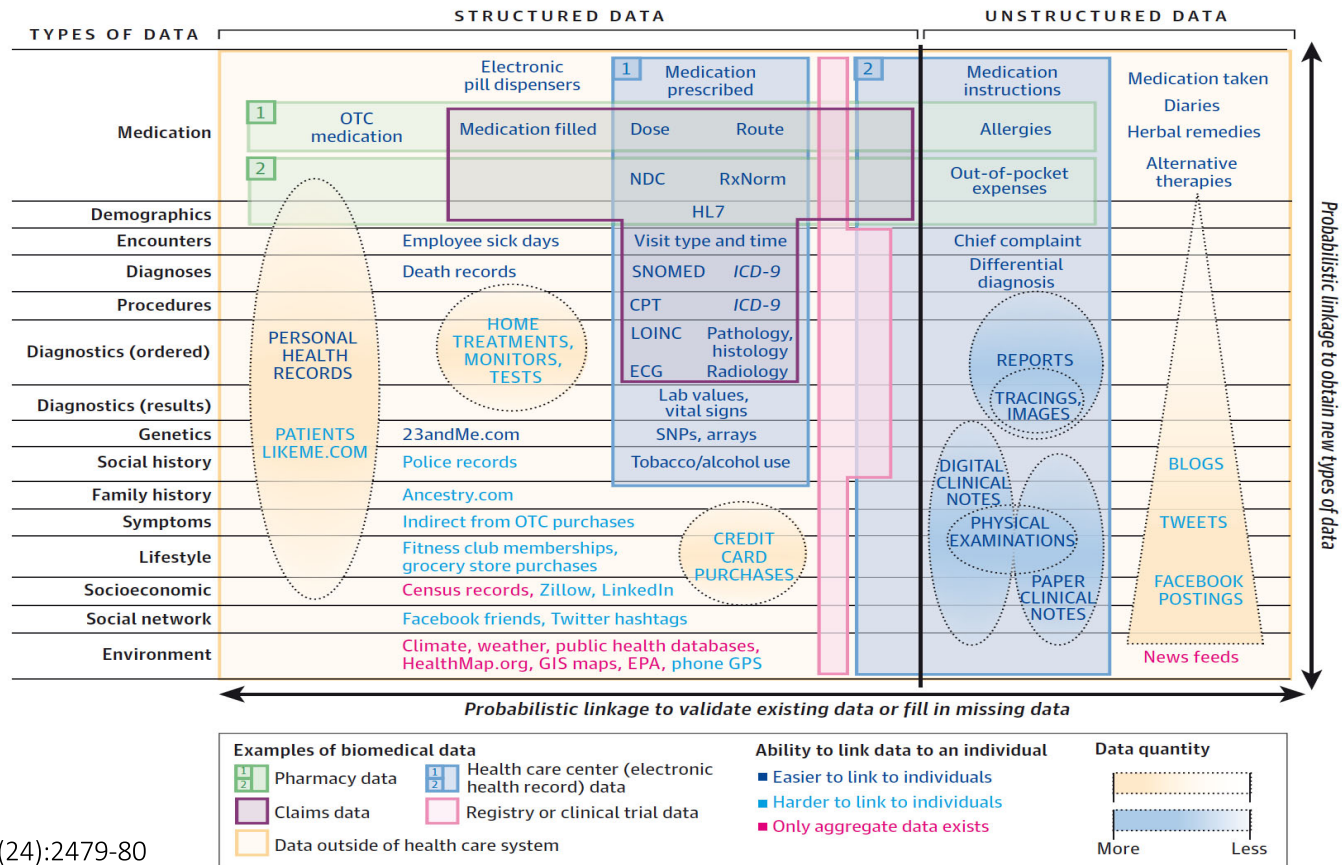


# Data Science



# Healthcare Big Data

Figure. The Tapestry of Potentially High-Value Information Sources That May be Linked to an Individual for Use in Health Care



Weber. JAMA 2014;311(24):2479-80



# Big Data for Pharmaceutical Field



Individual Case Safety Reports



Regulatory AERS



Medical Literature



Big Data Analytics



Medical Claims



Clinical EHRs



Clinical Trial Data



Social Media



Future Data Types

GUIDANCE DOCUMENT

## Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products

*Draft Guidance for Industry*

DECEMBER 2021

[Download the Draft Guidance Document](#)

[Read the Federal Register Notice](#)

Draft

Level 1 Guidance

Not for implementation. Contains non-binding recommendations.

[Share](#)

[Tweet](#)

[LinkedIn](#)

[Email](#)

[Print](#)

GUIDANCE DOCUMENT

## Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products

*Draft Guidance for Industry*

SEPTEMBER 2021

[Download the Draft Guidance Document](#)

[Read the Federal Register Notice](#)

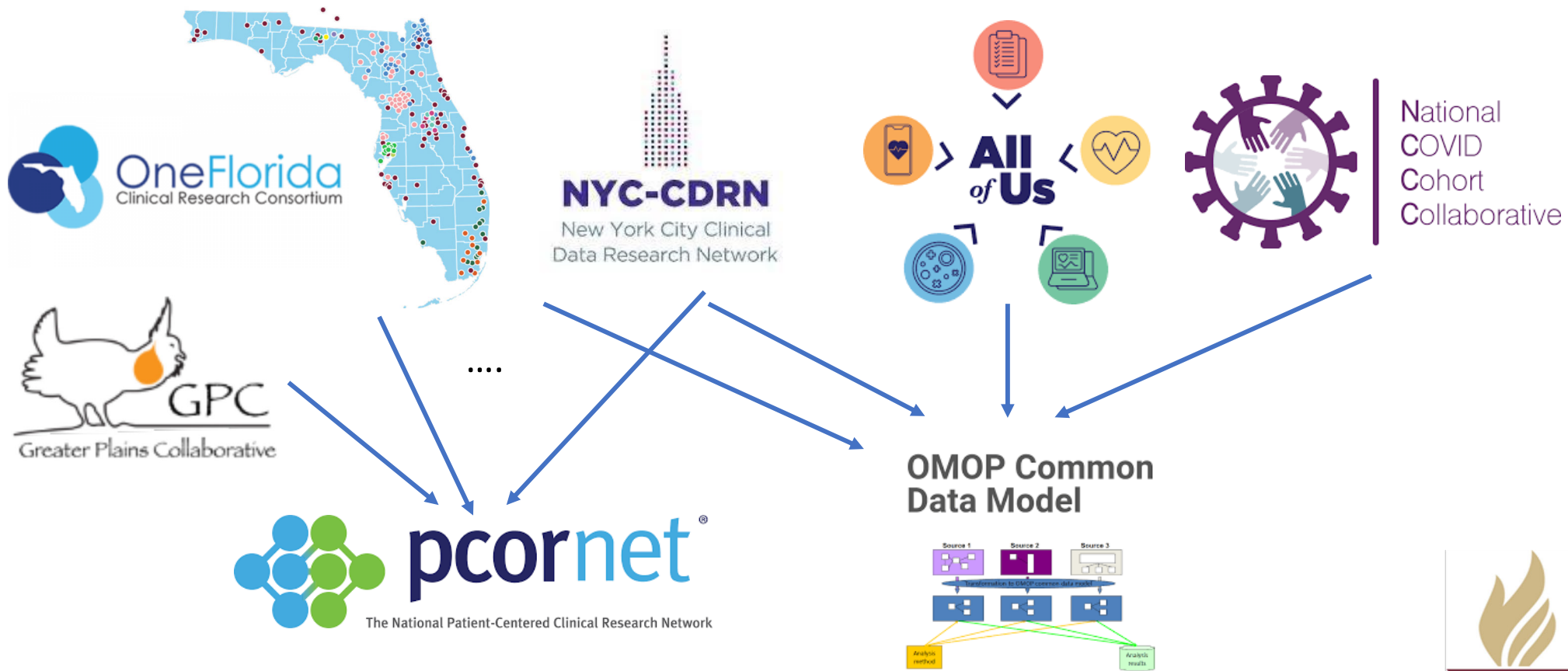
Draft

Level 1 Guidance

Not for implementation. Contains non-binding recommendations.

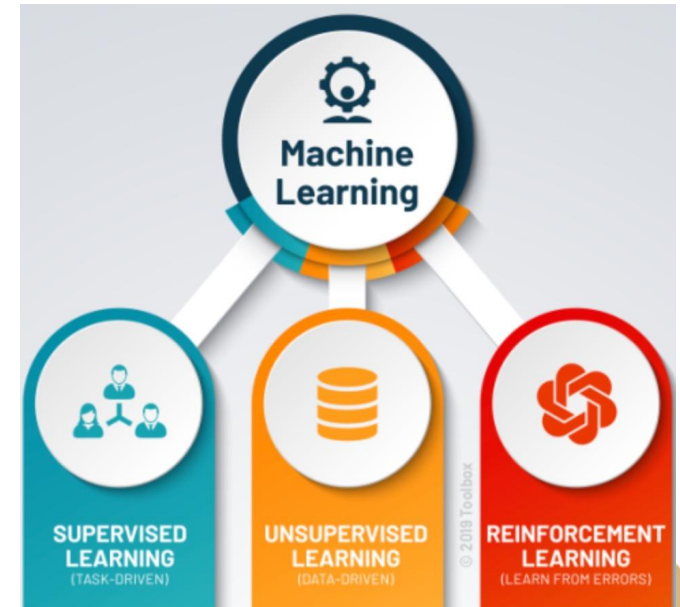


# Clinical Data Warehouses



# Applications of Data Science

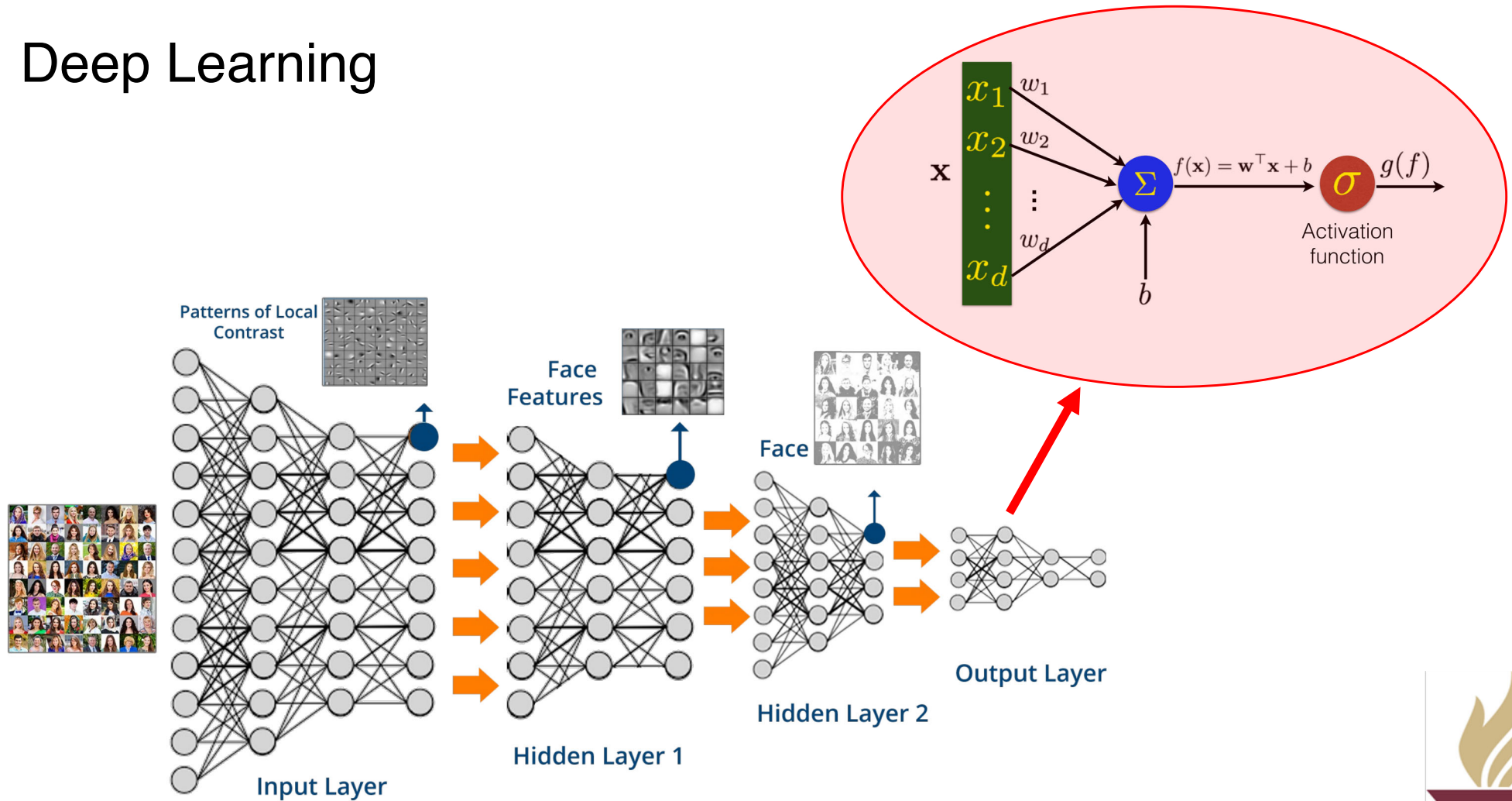
- Health outcome prediction
- Drug-drug interaction prediction
- Treatment adherence promotion
- Clinical trial patient selection optimization
- Deep phenotyping for precision medicine
- ....



<https://www.potentiaco.com/what-is-machine-learning-definition-types-applications-and-examples/>

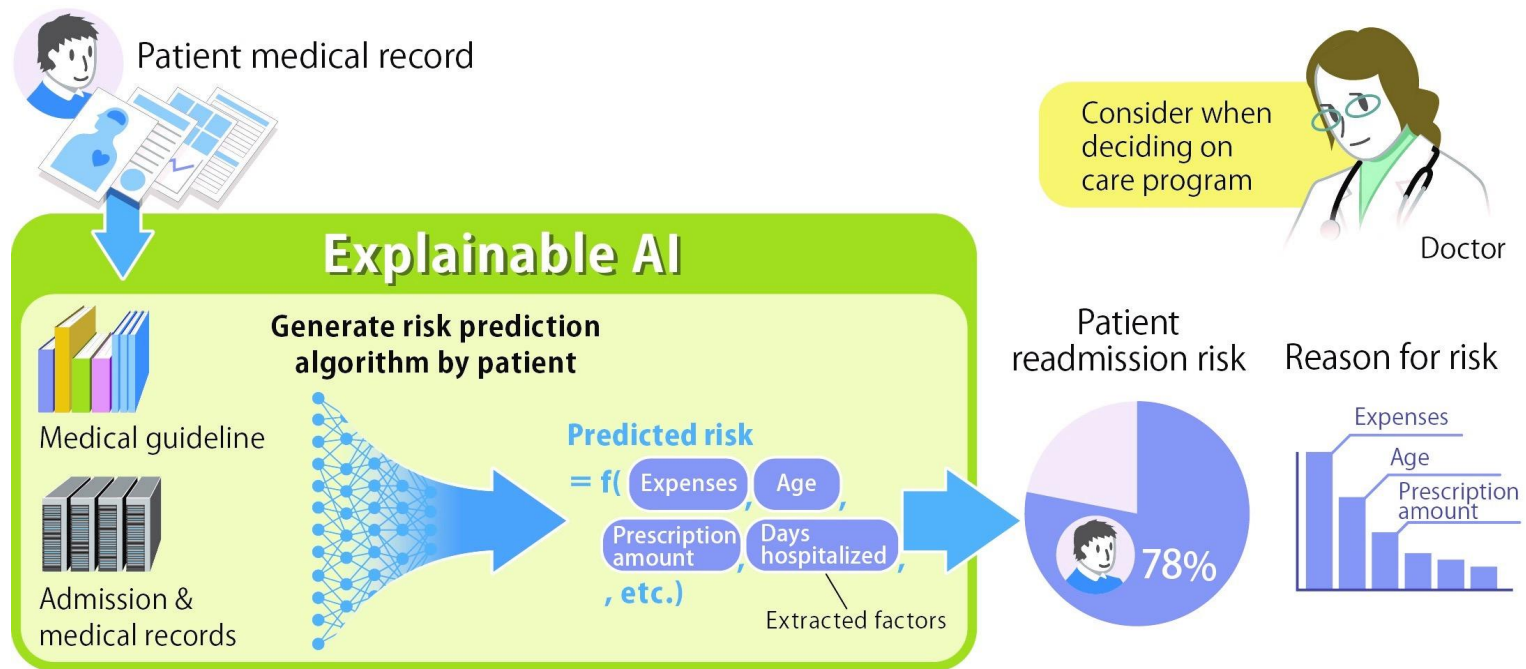


# Deep Learning

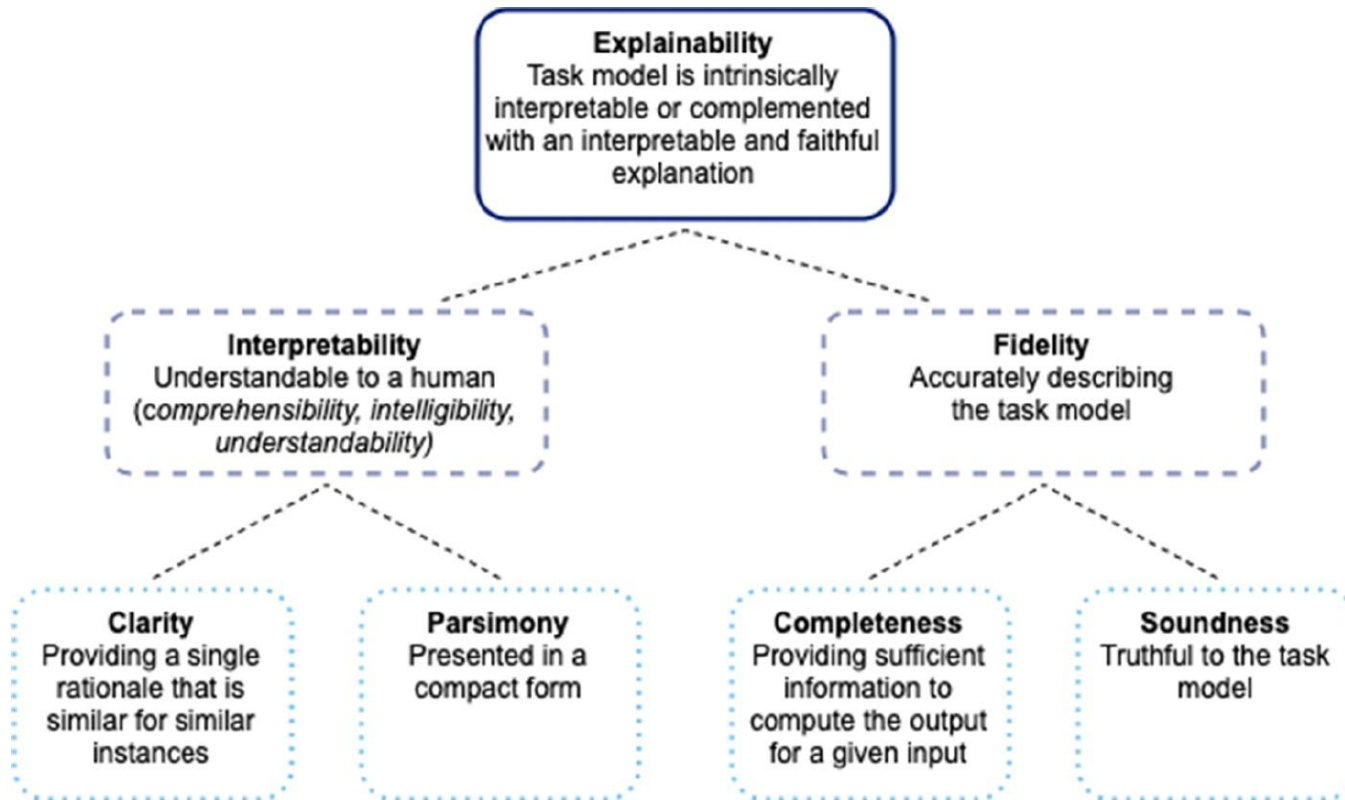


# Explainable AI

- **Explainable AI:** Enabling machines to explain their decisions and actions to human users



# Explainable AI Concept Map



(Markus 2021)





# Explainable AI Methods

- **By explainable models:**
  - Intrinsically interpretable models (e.g., logistic regression, decision tree)
  - Post hoc explanation (e.g., SHAP)
- **Scope of explanation:**
  - Local explanation (e.g., LIME, SHAP)
  - Global explanation (e.g., SHAP)
- **Types of explanation:**
  - Model-based explanation (e.g., Mimic learning)
  - Attribution-based explanations (e.g., SHAP)
  - Example-based explanations (e.g., counterfactual explanation)

*Journal of the American Medical Informatics Association*, 27(7), 2020, 1173–1185  
doi: 10.1093/jamia/ocaa053  
Advance Access Publication Date: 17 May 2020  
Review



Review

## Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review

Seyedeh Neelufar Payrovnaziri,<sup>1</sup> Zhaoyi Chen,<sup>2</sup> Pablo Rengifo-Moreno,<sup>3,4</sup> Tim Miller,<sup>5</sup> Jiang Bian,<sup>2</sup> Jonathan H. Chen,<sup>6,7</sup> Xiuwen Liu,<sup>8</sup> and Zhe He<sup>1</sup>

<sup>1</sup>School of Information, Florida State University, Tallahassee, Florida, USA, <sup>2</sup>Department of Health Outcomes and Biomedical Informatics, University of Florida, Gainesville, Florida, USA, <sup>3</sup>College of Medicine, Florida State University, Tallahassee, Florida, USA, <sup>4</sup>Tallahassee Memorial Hospital, Tallahassee, Florida, USA, <sup>5</sup>School of Computing and Information Systems, The University of Melbourne, Melbourne, Victoria, Australia, <sup>6</sup>Center for Biomedical Informatics Research, Department of Medicine, Stanford University, Stanford, California, USA, <sup>7</sup>Division of Hospital Medicine, Department of Medicine, Stanford University, Stanford, California, USA and <sup>8</sup>Department of Computer Science, Florida State University, Tallahassee, Florida, USA

**Corresponding Author:** Zhe He, PhD, School of Information, Florida State University, 142 Collegiate Loop, Tallahassee, FL 32306, USA (zhe@fsu.edu)

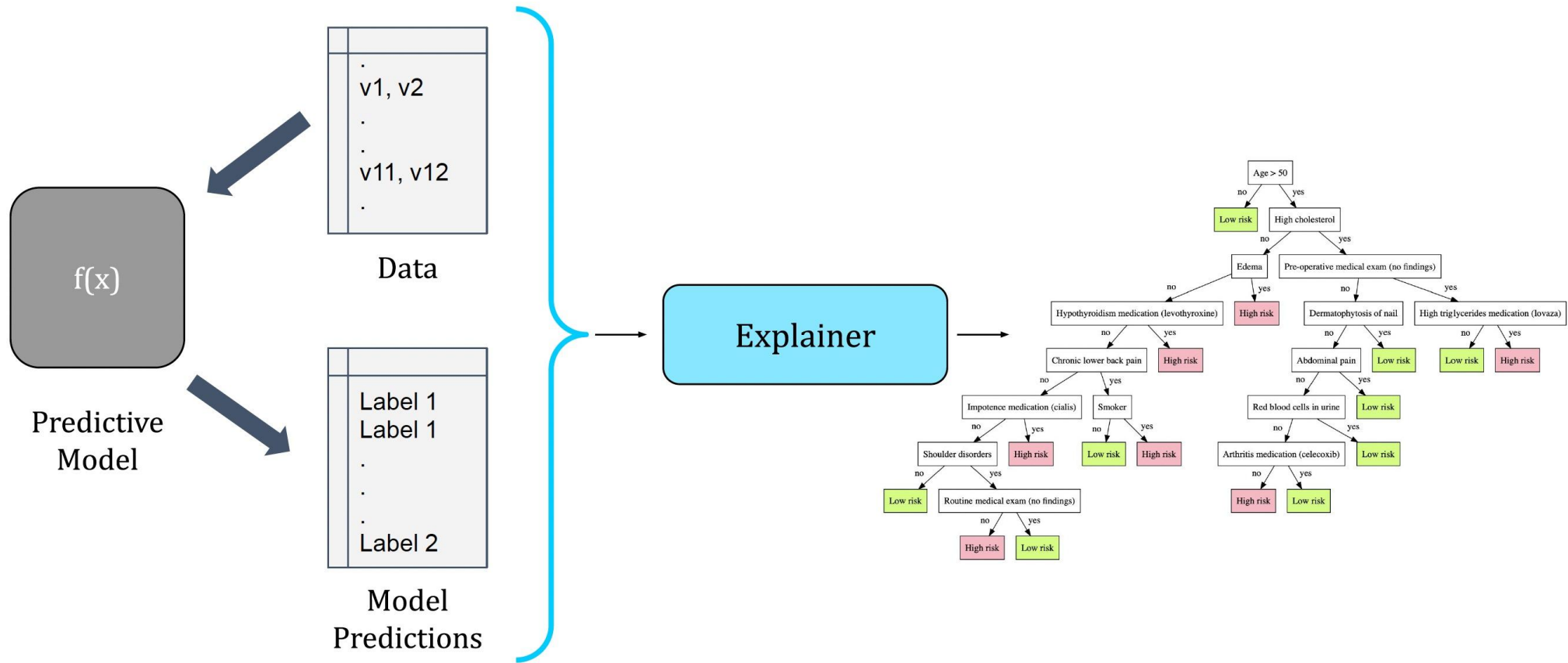
Received 22 January 2020; Revised 1 April 2020; Editorial Decision 4 April 2020; Accepted 7 April 2020

(Payrovnaziri 2020)

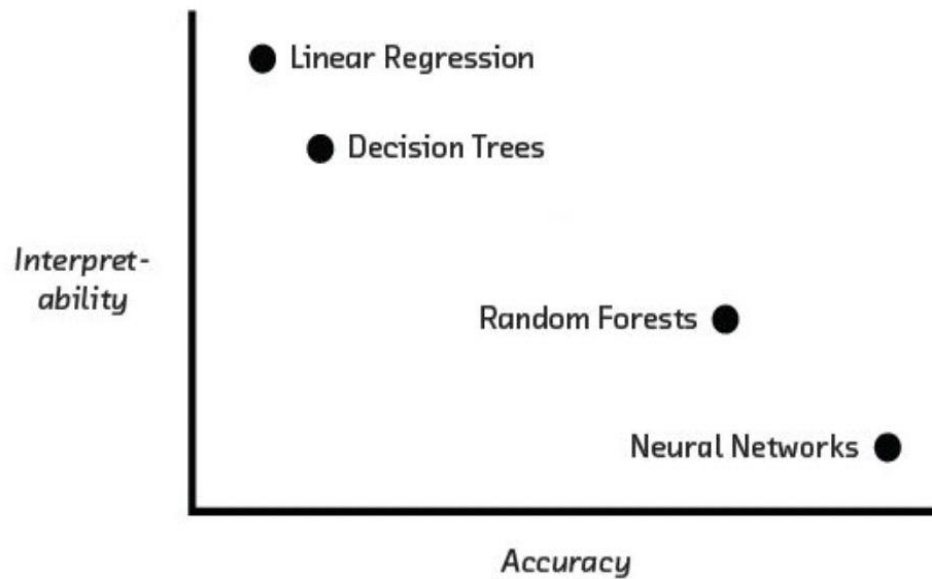
AAAI 21 Tutorial: <https://explainml-tutorial.github.io/aaai21?ref=recsyslabs.com>



# Example of Global Interpretation



# Trade Offs of Accuracy and Interpretability



nature machine intelligence

Perspective | Published: 13 May 2019

## Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead

Cynthia Rudin

*Nature Machine Intelligence* 1, 206–215(2019) | [Cite this article](#)

2794 Accesses | 17 Citations | 188 Altmetric | [Metrics](#)

**Annals of Internal Medicine**<sup>®</sup>

[LATEST](#) | [ISSUES](#) | [CHANNELS](#) | [CME/MOC](#) | [IN THE CLINIC](#) | [JOURNAL CLUB](#) | [WEB EXCLUSIVES](#) | [AUTHOR INFO](#)

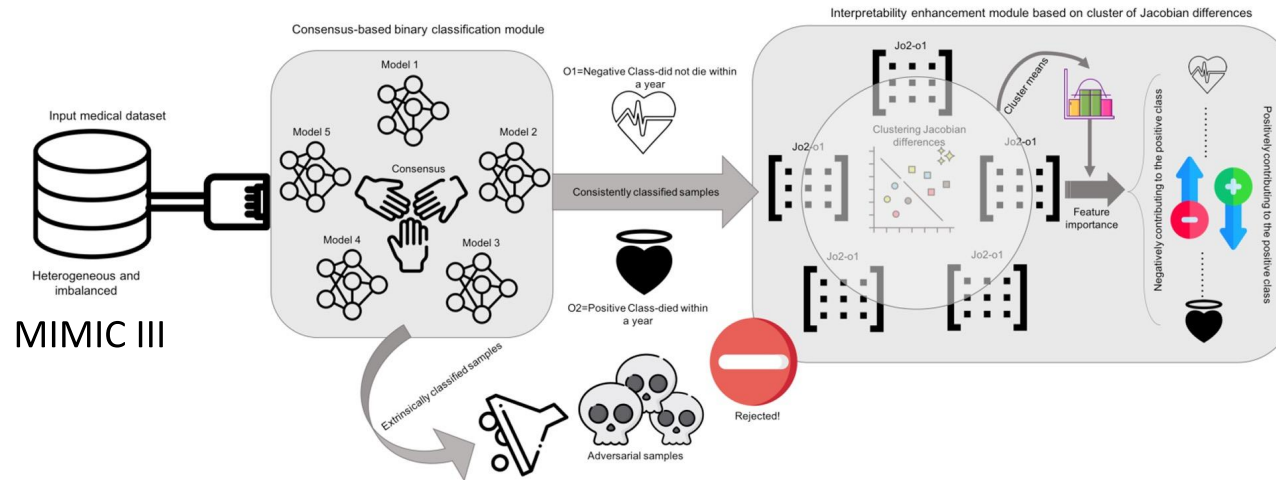
[PREV ARTICLE](#) | [THIS ISSUE](#) | [NEXT ARTICLE](#) | [IDEAS AND OPINIONS](#) | 7 JANUARY 2020

## Should Health Care Demand Interpretable Artificial Intelligence or Accept “Black Box” Medicine?

Fei Wang, PhD; Rainu Kaushal, MD, MPH; Dhruv Khullar, MD, MPP



# DeepConcensus (Salman 2020)



## Evaluation Result of Five Individual Models.

Model	Accuracy	ROC	Precision	Recall	F-measure
1	0.7421	0.6906	0.5849	0.5568	0.5705
2	0.7679	0.7259	0.6242	0.6167	0.6204
3	0.7348	0.6953	0.5657	0.5928	0.5789
4	0.7513	0.7006	0.6012	0.5846	0.5846
5	0.7495	0.6993	0.5974	0.5688	0.5828

5037 subjects & → 5436 admissions with applicable

Model	Accuracy	ROC	Precision	Recall	F-measure
LR	0.7845	0.7162	0.6923	0.5389	0.6060
SVM	0.7826	0.7066	0.7024	0.5089	0.5902
CSVM	0.7794	0.7257	0.6577	0.5868	0.6202
Consensus	0.8623	0.87	0.7631	0.6516	0.7030

Gastroenterology, and genitourinary

percutaneous cardiovascular procedure



# Contributions of Features

Contribution	Features List (mean, min, max, SD) **
Positive	sodium (138.63, 118.18, 139.66, 3.32), alkaline phosphatase (ALP) (98.81, 13, 172, 65.80), alanine aminotransferase (ALT) (83.93, 44, 5509, 226.65), creatinine (1.37, 0.15, 15.70, 0.78), blood urea nitrogen (BUN) (26.11, 5.10, 140.66, 10.82), lactate dehydrogenase (LD) (456.39, 100, 5664, 176.72), age of admission (70.62, 18.70, 100.76, 13.34), heart rate (82.34, 36.84, 132.66, 12.31), aspartate aminotransferase (AST) (138.81, 2, 13511.7, 72.70), troponin T (2.30, 2.30, 24.80, 1.70), respiratory rate (19.35, 8.73, 42.69, 3.23), potassium (4.18, 4.23, 4.24, 0.36), cancer_Positive, cortisol (33.01, 22, 238.2, 4.45), cholestrol ratio (4.03, 4.03, 6.8, 1.95)
Neutral	C-Reactive Protein (CRP) (52.03, 0.09, 273.75, 7.65), no genitourinary, marital_MARRIED, orthopaedic, other cardiac pacemaker implantation, no coronary bypass with cardiac catheterization or percutaneous Cardiac Procedure, endocrinology, hematological, marital_SINGLE
Negative	white blood cell count (11.30, 0.45, 107.67, 4.78), hemoglobin (10.96, 4.31, 18.7, 1.53), chloride (103.81, 80.42, 125.61, 4.29), triglycerides (139.54, 1, 1983, 69.16), bilirubin (0.9131, 0.1, 31.13, 0.07), bicarbonate (24.82, 7, 47.57, 3.58), albumin (3.20, 3.20, 3.30, 0.47), systolic blood pressure (106.67, 20, 334.78, 21.49), creatine kinase (1.37, 0.5, 29579, 931.52), cancer_negative, cardiac valve and other major cardiothoracic procedures with cardiac catheterization, Brain natriuretic peptide (BNP), coronary bypass without cardiac catheterization, coronary bypass with PTCA, cardiac defibrillator implant without cardiac catheterization, coronary bypass without cardiac catheterization or percutaneous cardiac procedure, cardiac valve and other major cardiothoracic procedures without cardiac catheterization

# Prediction of Early Mortality of Hemorrhagic Stroke Patients (AMIA Summit 2023)

- According to a recent survey, 35% of stroke patients die within 7 days of the stroke and about 50% of intracerebral hemorrhagic stroke patients died within 30 days.
- Hemorrhagic stroke occurs due to bleeding into the brain which is caused by rupture of the blood vessel.
- Most research focuses on ischemic stroke mortality instead of hemorrhagic stroke patients.

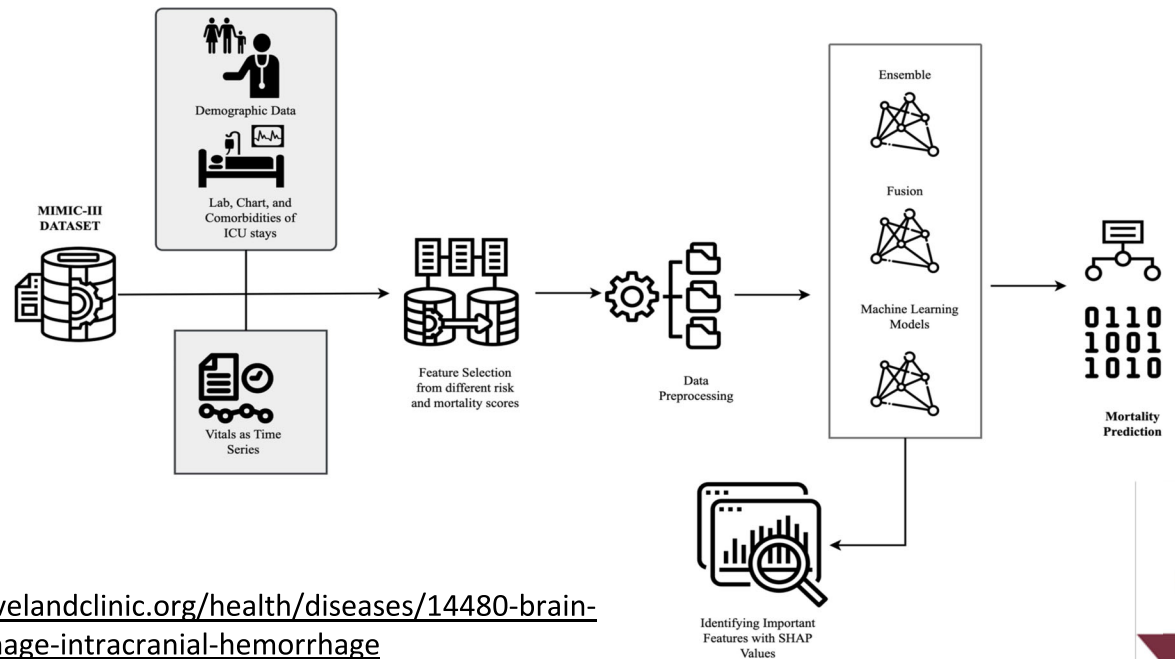
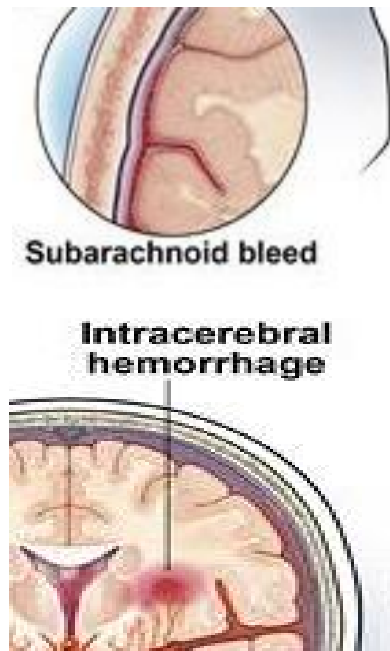
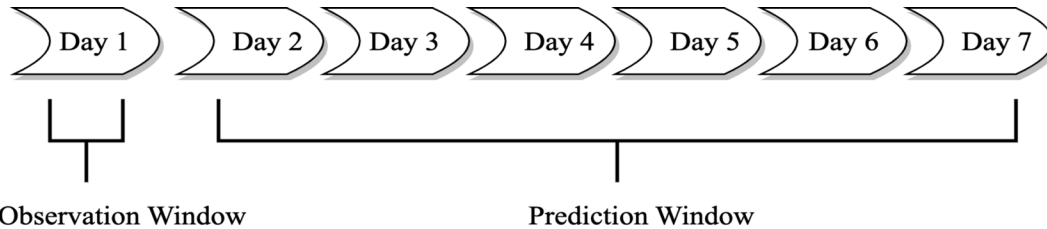


Image Source:

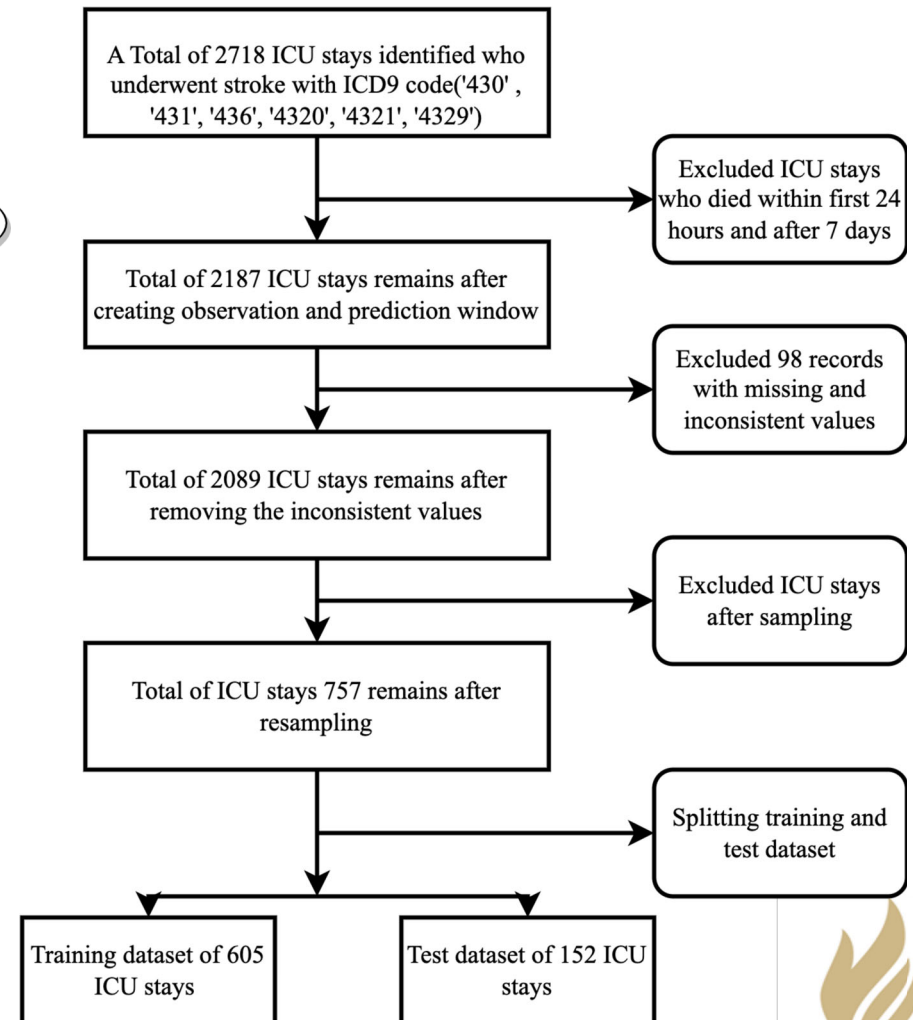
<https://my.clevelandclinic.org/health/diseases/14480-brain-bleed-hemorrhage-intracranial-hemorrhage>



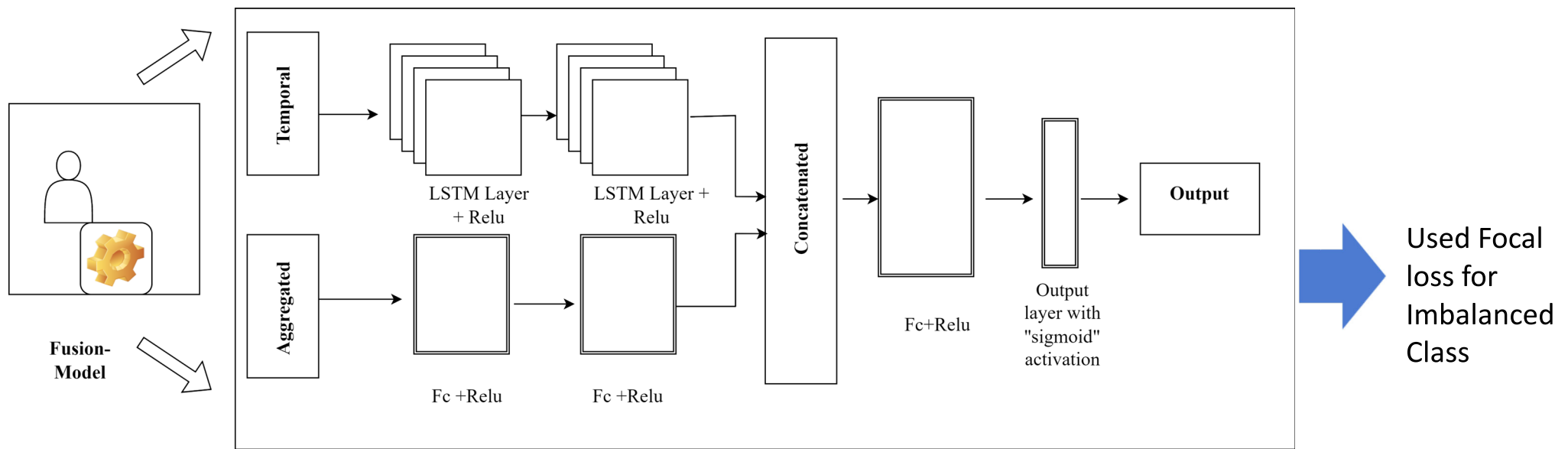
# Model Setup and Cohort



Age	Positive Instances		Negative Instances	
	Number	Mean (Standard Deviation)	Number	Mean (Standard Deviation)
0-19	0	0	0	0
20-45	21	36.6 (6.29)	62	37.9 (5.63)
45-65	87	57.16 (5.59)	169	57.02 (5.12)
65-75	67	70.53 (3.32)	90	71.78 (2.68)
75 and above	128 (excluding 4 inconsistent age values (age > 300))	82.06 (3.1)	124 (excluding 4 inconsistent age values (age > 300))	81.5 (3.43)
Gender	Number	Percentage	Number	Percentage
Male	158	51.47	244	54.22
Female	149	48.53	206	45.78
Race				
White	198	64.4	321	71.33
Black	18	5.9	38	8.44
Hispanic	15	4.9	13	2.9
Asian	15	4.9	11	2.44
Other Race	61	19.9	67	14.89



# Fusion Model

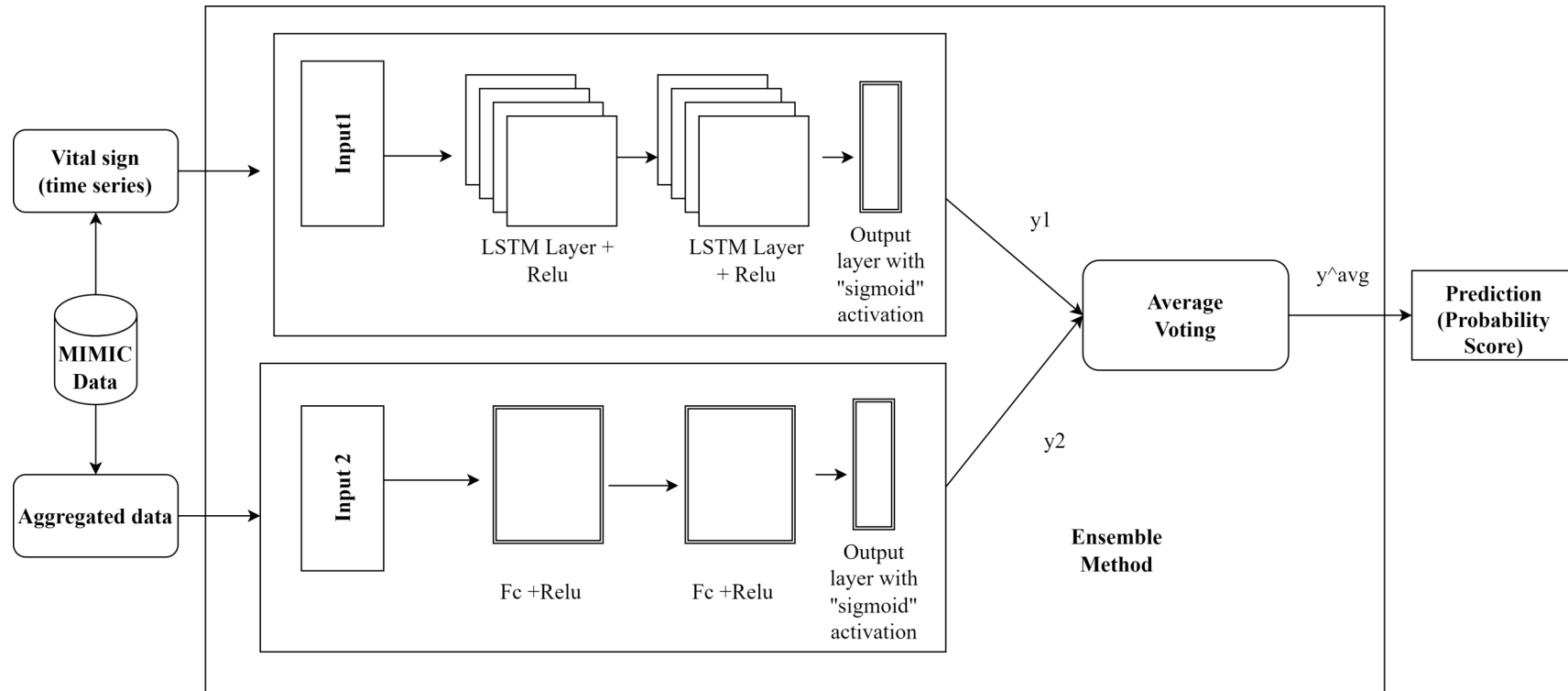


128 and 64 Neurons, Batch size 32,  
Epoch 100, 345,025 parameters





# Ensemble Model



128 and 64 Neurons, Batch size 32, Epoch 50 for model 1 and 100 for model 2, 29,057 parameters

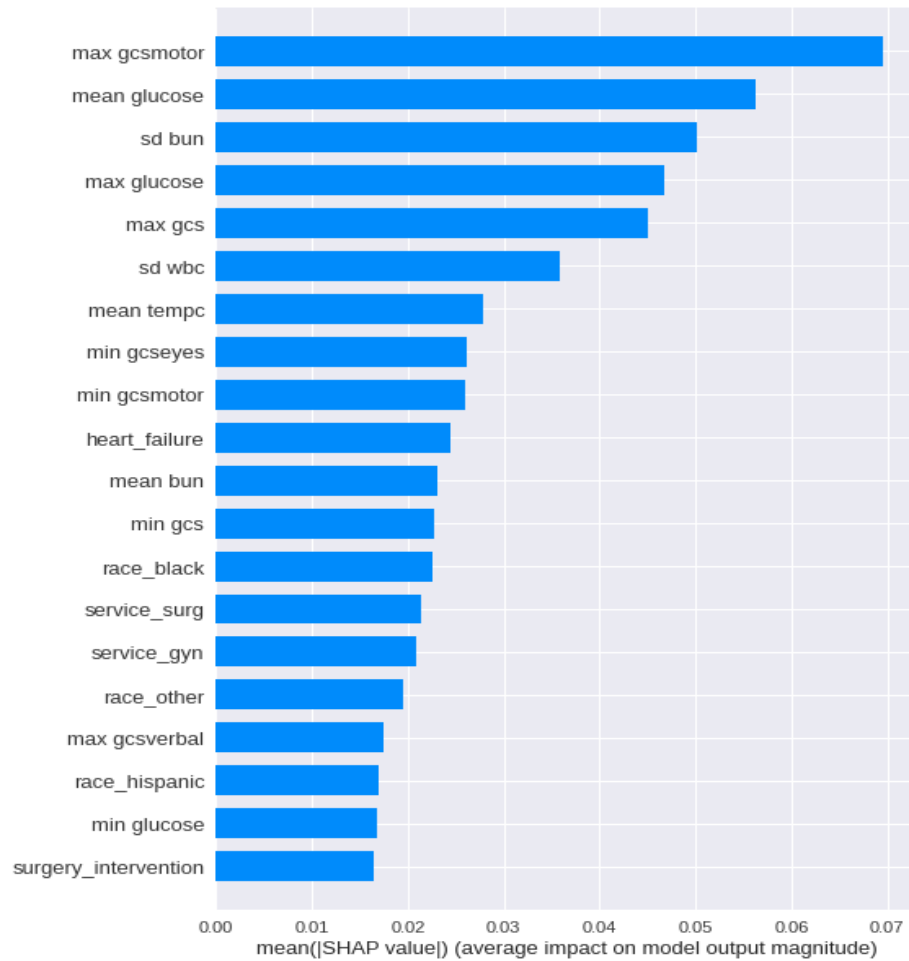


# Performance of Models

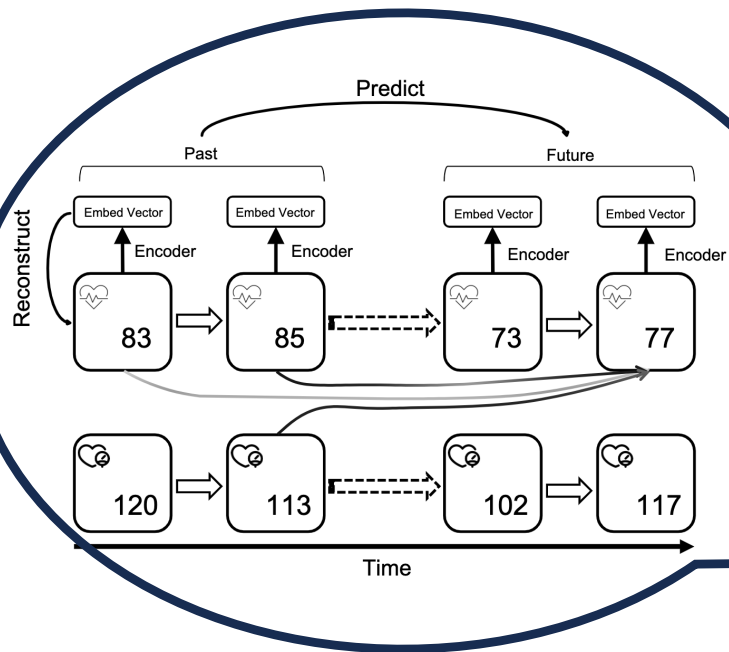
Models	Precision (STD)	Recall (STD)	F1 (STD)	Accuracy (STD)	AUROC (STD)
Fusion model	0.79	0.61	0.69	0.78	0.75
Ensemble (avg)	<b>0.8</b>	<b>0.77</b>	<b>0.79</b>	<b>0.83</b>	<b>0.82</b>
Logistic Regression	0.77 (0.04)	0.71 (0.05)	0.74 (0.04)	0.8 (0.03)	0.87 (0.03)
Decision Tree	0.72 (0.06)	0.68 (0.06)	0.7 (0.04)	0.76 (0.03)	0.80 (0.05)
Random Forest	0.76 (0.05)	0.74 (0.06)	0.75 (0.05)	0.8 (0.04)	0.87 (0.03)
XGBoost	0.75 (0.06)	0.74 (0.05)	0.75 (0.05)	0.8 (0.04)	0.87 (0.03)



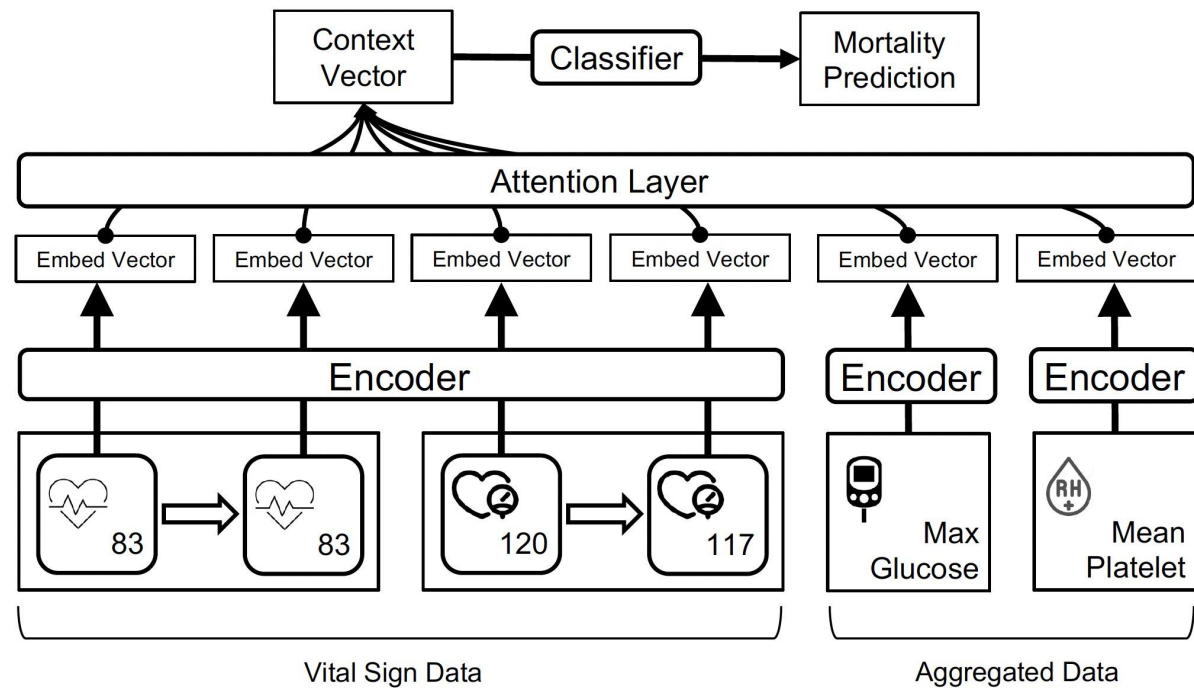
# Explanation with SHAP on Aggregate Features

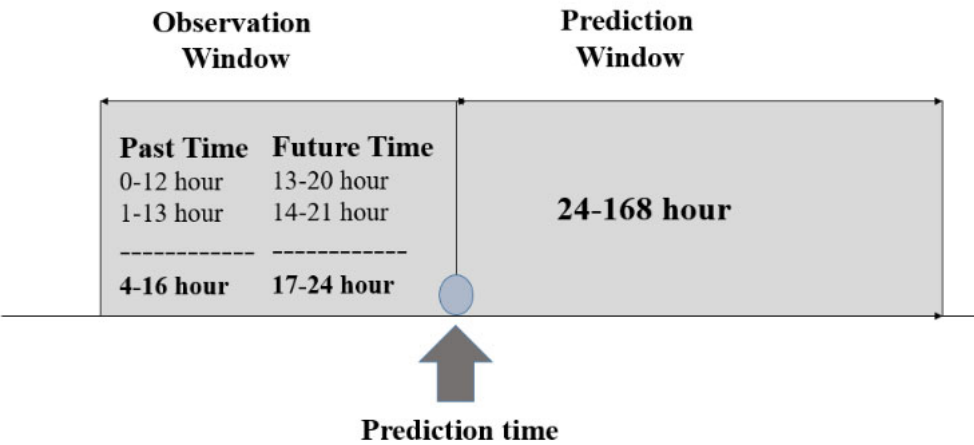


# Local Explanation with Attention (BCB 2023)



Self-supervised training framework  
(Transformer-based encoder)





**Table 3.** Utility Performance Result

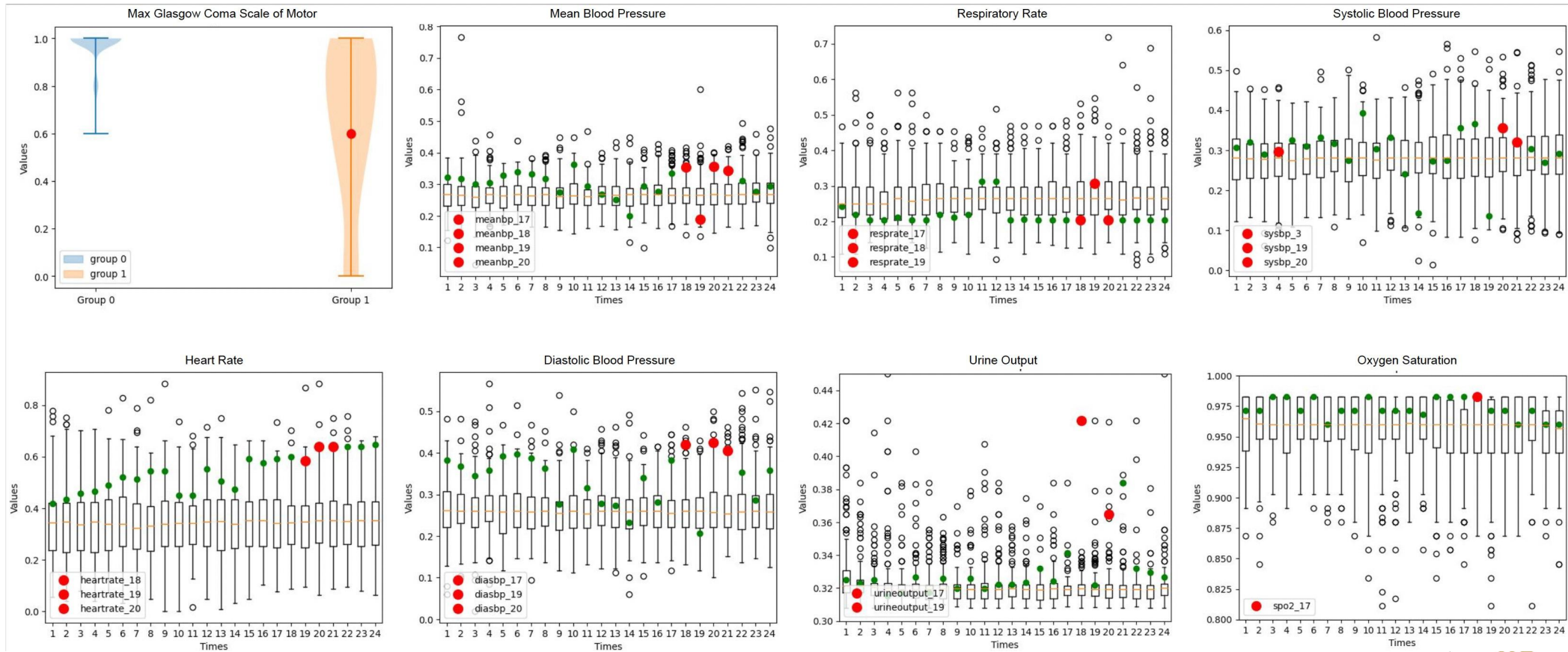
	Logistic	LSTM	Attention
AUROC	0.8215 $\pm 9.37e-3$	0.8170 $\pm 2.45e-3$	<b>0.8487</b> $\pm 2.75e-2$
AUPRC	0.9675 $\pm 2.31e-3$	0.9723 $\pm 1.42e-3$	<b>0.9726</b> $\pm 9.65e-3$

**Table 4.** Explanation and Efficiency Performance Result

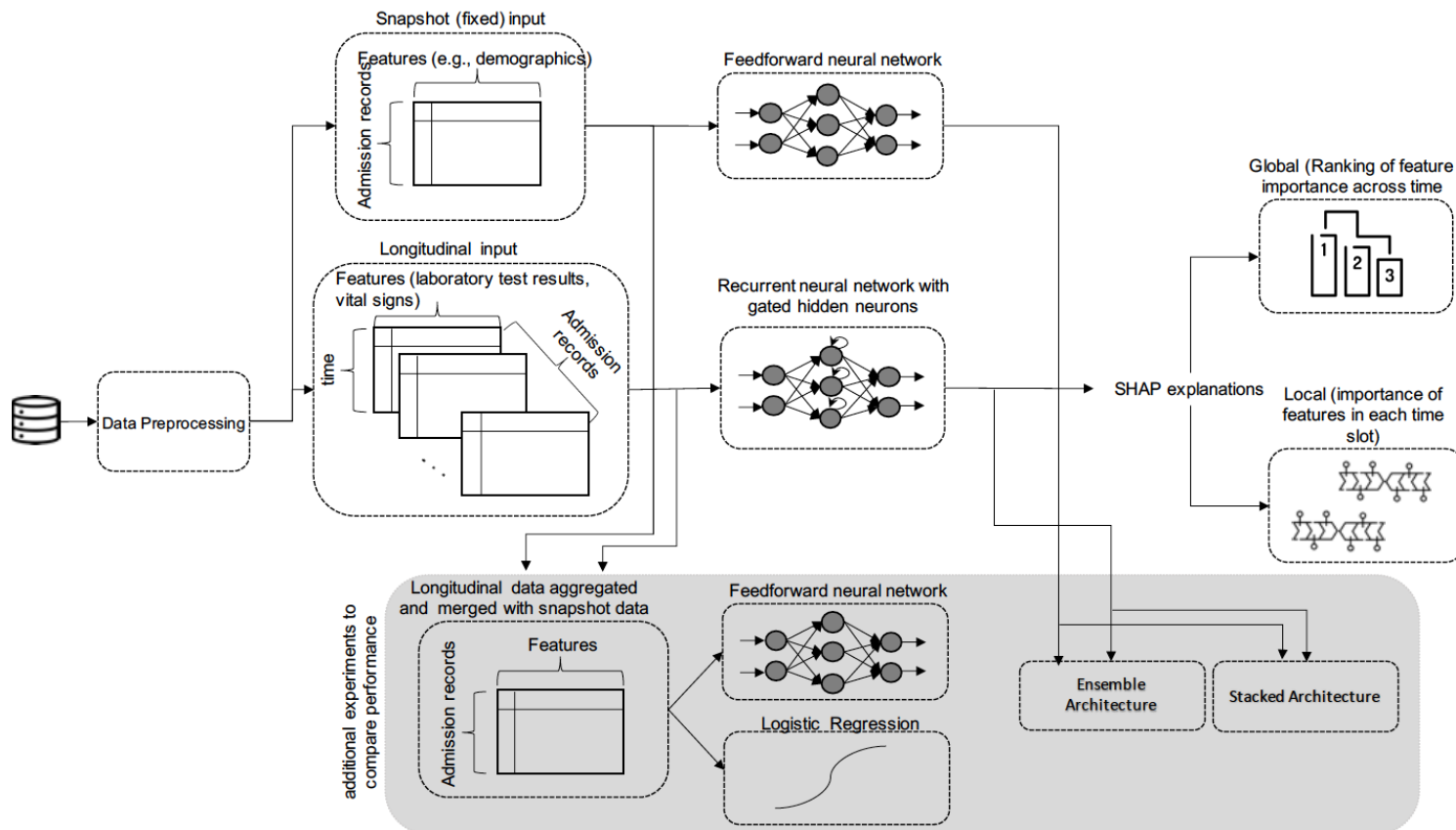
		Logistic		LSTM		Attention	
		Weight	SHAP	Weight	SHAP	Atten	SHAP
Fidelity+	AUROC ( $\downarrow$ )	-0.0681	-0.2217	-0.2547		<b>-0.3617</b>	-0.4016
	AUPRC ( $\downarrow$ )	-0.0702	-0.0507	-0.0600		<b>-0.0854</b>	-0.0725
	Prob. ( $\downarrow$ )	-0.2411	-0.3992	-0.2427		<b>-0.2544</b>	-0.2893
Fidelity-	AUROC ( $\uparrow$ )	-0.0007	-0.0887	-0.0059		<b>0.0000</b>	-0.0990
	AUPRC ( $\uparrow$ )	-0.0001	-0.0170	-0.0316		<b>0.0001</b>	-0.0312
	Prob. ( $\uparrow$ )	-0.0090	-0.0709	-0.0180		<b>-0.0017</b>	-0.0745
Time (ms, $\downarrow$ )		0.0216	10.9335			0.5715	502.174



# Local Explanation of a Positive Instance (Died)



# Interpretable DL Models for Predicting Short-Term Mortality



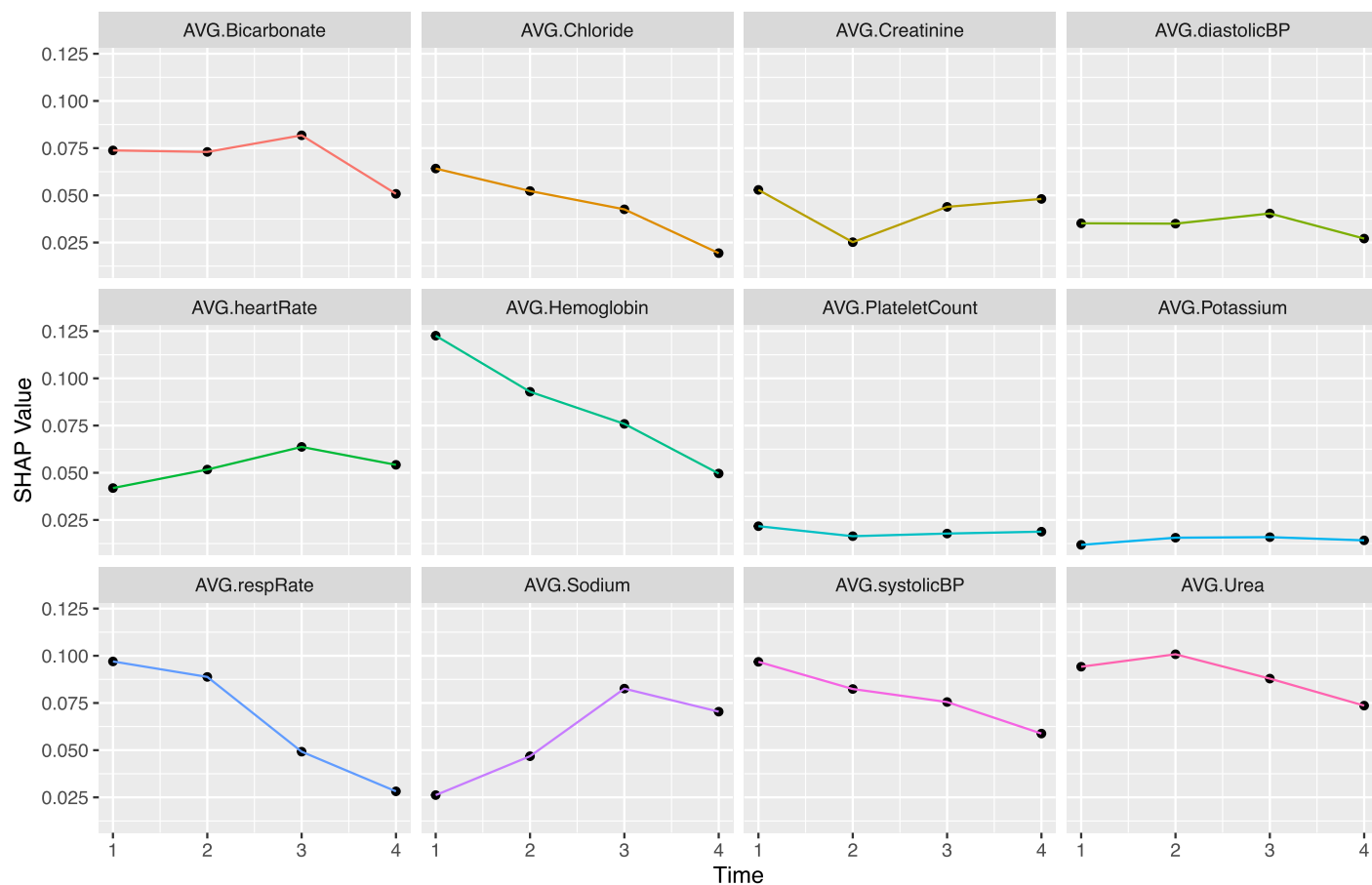
# Performance of the Models

Model	Data	AUROC±stdev	Precision±stdev	Recall±stdev	F±stdev
RNN	longitudinal (lab tests and vital signs)	0.7587±0.02	0.7370±0.01	0.7516±0.01	0.7330±0.00
DNN	fixed (demographics and comorbidities)	0.8052±0.00	0.8406±0.00	0.8312±0.00	0.8140 ±0.00
Ensemble (RNN, DNN)	longitudinal and fixed data	<b>0.8635±0.01</b>	<b>0.8505±0.00</b>	0.8380±0.00	0.8212±0.01
Stacked (RNN, DNN)	longitudinal and fixed data	0.8615±0.01	0.8394±0.00	<b>0.8418±0.00</b>	<b>0.8345±0.00</b>
Merged (DNN)	aggregated lab tests and vital signs merged with demographic and comorbidities data	0.8435±0.01	0.8441±0.00	0.8413±0.00	0.8295±0.01
Logistic Regression	aggregated lab tests and vital signs merged with demographic and comorbidities data	0.8460±0.01	0.8337±0.00	0.8345±0.00	0.8240±0.00
Logistic Regression	aggregated lab tests and vital signs	0.7315±0.01	0.7302±0.01	0.7443±0.00	0.7042±0.00



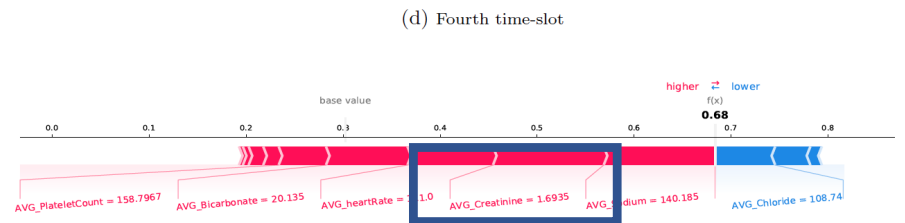
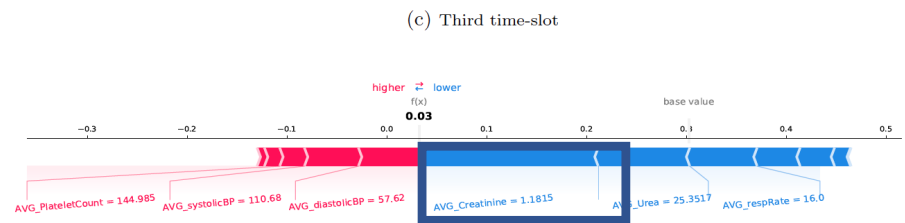
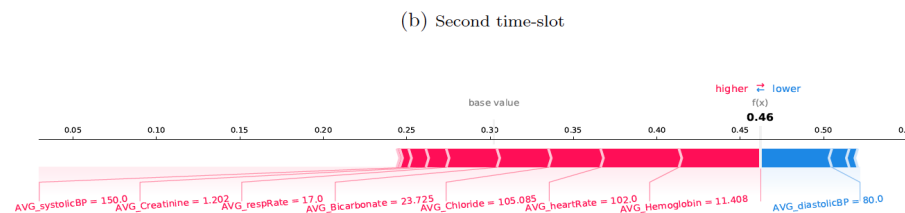
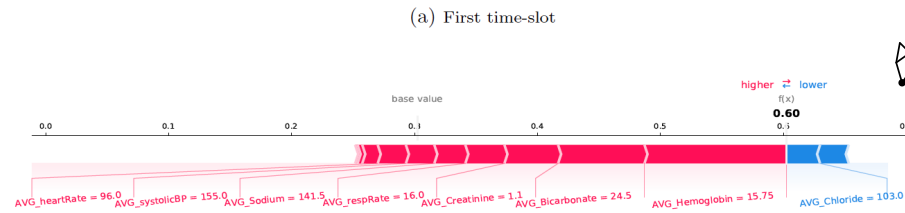


# Global Interpretation of the RNN model: SHAP Values

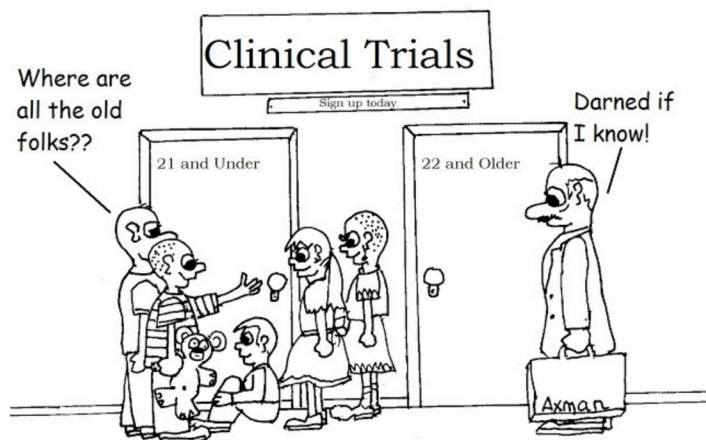


# Local Interpretation of the RNN model

Positive instance  
(died in 7 days)



# Clinical Trial Generalizability



- Stringent eligibility criteria
  - Low accrual
  - Early termination
  - Low generalizability
- Increased likelihood of adverse events

## Generalizability of Clinical Trial Results for Adolescent Major Depressive Disorder

Carlos Blanco, MD, PhD,<sup>a</sup> Nicolas Hoertel, MD, MPH,<sup>b,c,d</sup> Silvia Franco, MD,<sup>e</sup> Mark Olfson, MD, MPH,<sup>f</sup> Jian-Ping He, MSc,<sup>g</sup> Saioa López, MD,<sup>h</sup> Ana González-Pinto, PhD,<sup>f</sup> Frédéric Limosin, MD, PhD,<sup>i,j,k</sup> Kathleen R. Merikangas, PhD<sup>l</sup>

**BACKGROUND:** Although there have been a number of clinical trials evaluating treatments for adolescents with major depressive disorder (MDD), the generalizability of those trials to samples of depressed adolescents who present for routine clinical care is unknown. Examining the generalizability of clinical trials of pharmacological and psychotherapy interventions for adolescent depression can help administrators and frontline practitioners determine the relevance of these studies for their patients and may also guide eligibility criteria for future clinical trials in this clinical population.

**METHODS:** Data on nationally representative adolescents were derived from the National Comorbidity Survey: Adolescent Supplement. To assess the generalizability of adolescent clinical trials for MDD, we applied a standard set of eligibility criteria representative of clinical trials to all adolescents in the National Comorbidity Survey: Adolescent Supplement with a *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* diagnosis of MDD ( $N = 592$ ).

abstract

## Participation of Older People in Preauthorization Trials of Recently Approved Medicines

Erna Beers, MD,<sup>a,b</sup> Dineke C. Moerkerken, BSc,<sup>a,b</sup> Hubert G.M. Leufkens, PharmD, PhD,<sup>c,d</sup> Toine C.G. Egberts, PharmD, PhD,<sup>c,e</sup> and Paul A.F. Jansen, MD, PhD<sup>a,b,d</sup>

**OBJECTIVES:** To investigate the inclusion of older people in clinical trials of recently authorized medicines, evaluating adherence to the 20-year-old International Conference of Harmonisation (ICH) guideline on geriatrics (E7).

**DESIGN:** Observational.

**SETTING:** European public assessment reports, published clinical trials, World Health Organization International Clinical Trials Registry Platform.

**PARTICIPANTS:** Individuals aged 65 and older and 75 and older.

**MEASUREMENTS:** Number and proportion of randomized participants and all exclusion criteria of studies involving diseases characteristically associated with aging (venous thromboembolism, osteoporosis, atrial fibrillation) and diseases not unique to older adults (type 2 diabetes mellitus, depression, bipolar disorder, epilepsy).

criteria correlated with age (60.5%) were applied more frequently in larger trials ( $P < .02$ ).

**CONCLUSION:** Studies of diseases not uniquely associated with old age included an unacceptably low proportion of older people, contrary to the recommendations of the ICH E7 guideline. Although the proportion of older participants in trials of diseases characteristically associated with aging was appropriate for certain medicines, the use of age-sensitive exclusion criteria limits the representativeness of the trial population for daily practice. *J Am Geriatr Soc* 62:1883–1890, 2014.

**Key words:** older participant; clinical trial; exclusion; regulatory agency; geriatric medicine

Moodie et al. *Trials* (2016) 17:242  
DOI 10.1186/s13063-016-1368-3

Trials

RESEARCH

Open Access



## A case study of SMART attributes: a qualitative assessment of generalizability, retention rate, and trial quality

Erica E. M. Moodie<sup>1\*</sup>, James C. Karan<sup>1</sup> and Susan M. Shortreed<sup>2</sup>

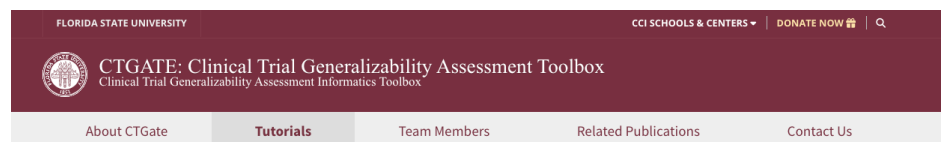
**Abstract**

**Background:** Personalizing medical care is becoming increasingly popular, particularly mental health care. There is growing interest in formalizing medical decision making based on evolving patient symptoms in an evidence-based manner. To determine optimal sequencing of treatments, the sequences themselves must be studied; this may be accomplished by using a sequential multiple assignment randomized trial (SMART). It has been hypothesized that SMART studies may improve participant retention and generalizability.

**Methods:** We examine the hypothesis that SMART studies are more generalizable and have better retention than



# ctGATE: Clinical Trial Generalizability Assessment Toolbox



## Home

### Generalizability assessment papers

Source Disease Category A Priori/A Posteriori Score/Non-Score

Show 10 entries

Search:

Source	PMID	Title	Disease Category	A Priori/A Posteriori	Score/Non-Score	Tutorial/Method
PubMed	24926156	The use of propensity scores to assess the generalizability of results from randomized trials	Not specified	A Posteriori	Score	Propensity score weighting/matching
PubMed	11870014	Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice?	Mental disorder	A Priori	Non-score	What is the percentage of patients who would be excluded by each exclusion criterion?
Cochrane	12153370	Representation of the elderly, women, and minorities in heart failure clinical trials	Cardiovascular diseases	A Posteriori	Non-score	How many elderly, women, and minorities are excluded by trials?
PubMed	14628985	How many subjects with major depressive disorder meet eligibility requirements of an antidepressant efficacy trial?	Mental disorder	A Priori	Non-score	List the major reasons for exclusion
PubMed	14770052	Exclusion criteria used in antidepressant efficacy trials: consistency across studies and representativeness of samples included.	Mental disorder	A Priori	Non-score	List the commonly used psychiatric exclusion criteria

## Tutorials

### Tutorials for generalizability assessment

Tutorials are listed here:

- GIST 2.0 / Interactive tutorial for GIST 2.0** (Sen A, Chakrabarti S, Goldstein A, Wang S, Ryan PB, Weng C. GIST 2.0: A scalable multi-trait metric for quantifying population representativeness of individual clinical studies. *Journal of biomedical informatics*. 2016 Oct 1;63:325-36. PMID: 27600470)
- Comparing populations in clinical studies** (Example publication: de C Williams AC, Nicholas MK, Richardson PH, Pither CE, Fernandes J. Generalizing from a controlled trial: the effects of patient preference versus randomization on the outcome of inpatient versus outpatient chronic pain management. *Pain*. 1999 Oct 1;83(1):57-65. PMID: 10506672)
- Standardized mean difference (SSMD) of propensity scores** (Example publication: Susukida R, Crum RM, Stuart EA, Mojtabai R. Generalizability of the findings from a randomized controlled trial of a web-based substance use disorder intervention. *The American journal on addictions*. 2018 Apr;27(3):231-7. PMID: 29569403)
- Propensity score weighting/matching** (Example publication: Stuart EA, Cole SR, Bradshaw CP, Leaf PJ. The use of propensity scores to assess the generalizability of results from randomized trials. *J R Stat Soc Ser A Stat Soc*. 2001 Apr 1;174(2):369-386. doi: 10.1111/j.1467-985X.2010.00673.x. PMID: 24926156; PMCID: PMC4051511.

We also provided the GitHub repository and an interactive tutorial for GIST 2.0:

All the tutorials and data can be downloaded on [Github](#)

<https://ctgate.cci.fsu.edu/>



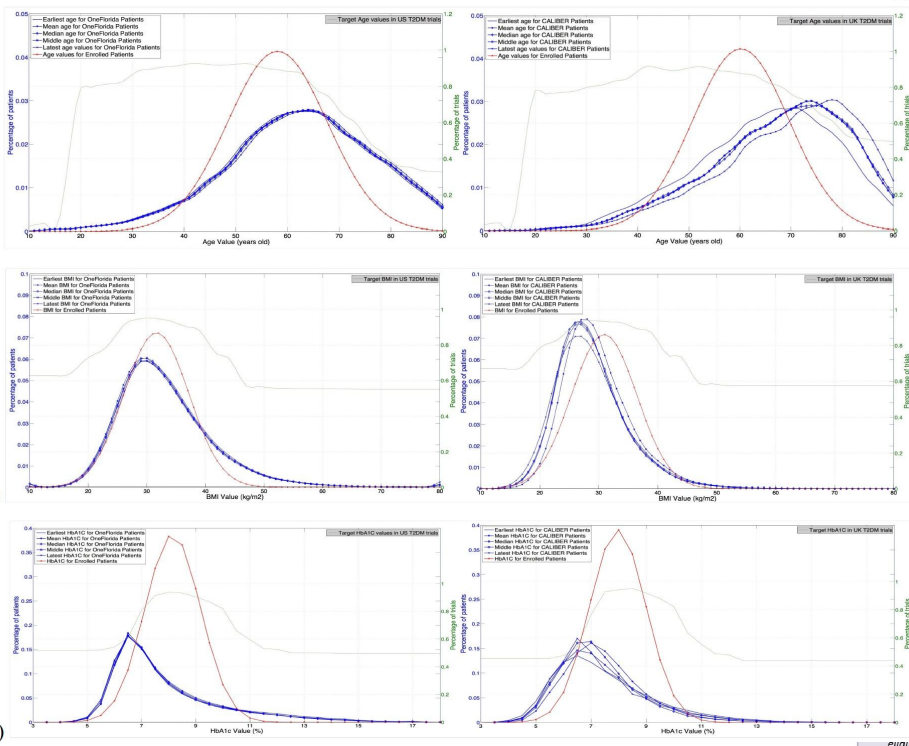
# Score-Based a priori Generalizability Assessments

## Assessing the Validity of a *a priori* Patient-Trial Generalizability Score using Real-world Data from a Large Clinical Data Research Network: A Colorectal Cancer Clinical Trial Case Study

Qian Li, MS<sup>1#</sup>, Zhe He, PhD<sup>2#</sup>, Yi Guo, PhD<sup>1#</sup>, Hansi Zhang, MS<sup>1</sup>, Thomas J George Jr, MD, FACP<sup>1</sup>, William Hogan, MD, MS<sup>1</sup>, Neil Charness, PhD<sup>2</sup>, Jiang Bian, PhD<sup>1\*</sup>  
<sup>1</sup>University of Florida, Gainesville, FL, USA; <sup>2</sup>Florida State University, Tallahassee, FL, USA;

### Abstract

Existing trials had not taken enough consideration of their population representativeness, which can lower the effectiveness when the treatment is applied in real-world clinical practice. We analyzed the eligibility criteria of vacizumab colorectal cancer treatment trials, assessed their a priori generalizability, and examined how it affects patient outcomes when applied in real-world clinical settings. To do so, we extracted patient-level data from a large collection of electronic health records (EHRs) from the OneFlorida consortium. We built a zero-inflated negative binomial model using a composite patient-trial generalizability (cPTG) score to predict patients' clinical outcomes (i.e., number of serious adverse events, [SAEs]). Our study results provide a body of evidence that 1) the cPTG scores can predict patient outcomes; and 2) patients who are more similar to the study population in the trials that were used to develop the treatment will have a significantly lower possibility to experience serious adverse events.



Li Q, He Z, Guo Y, et al. Assessing the Validity of a a priori Patient-Trial Generalizability Score using Real-world Data from a Large Clinical Data Research Network: A Colorectal Cancer Clinical Trial Case Study. AMIA Annu Symp Proc. Mar 4;2019:1101-1110

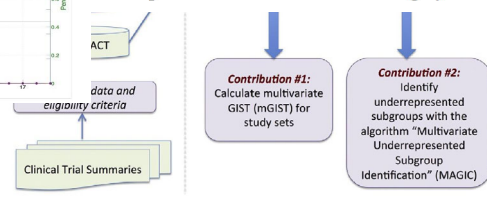
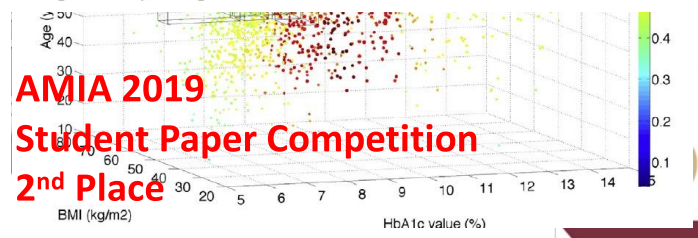
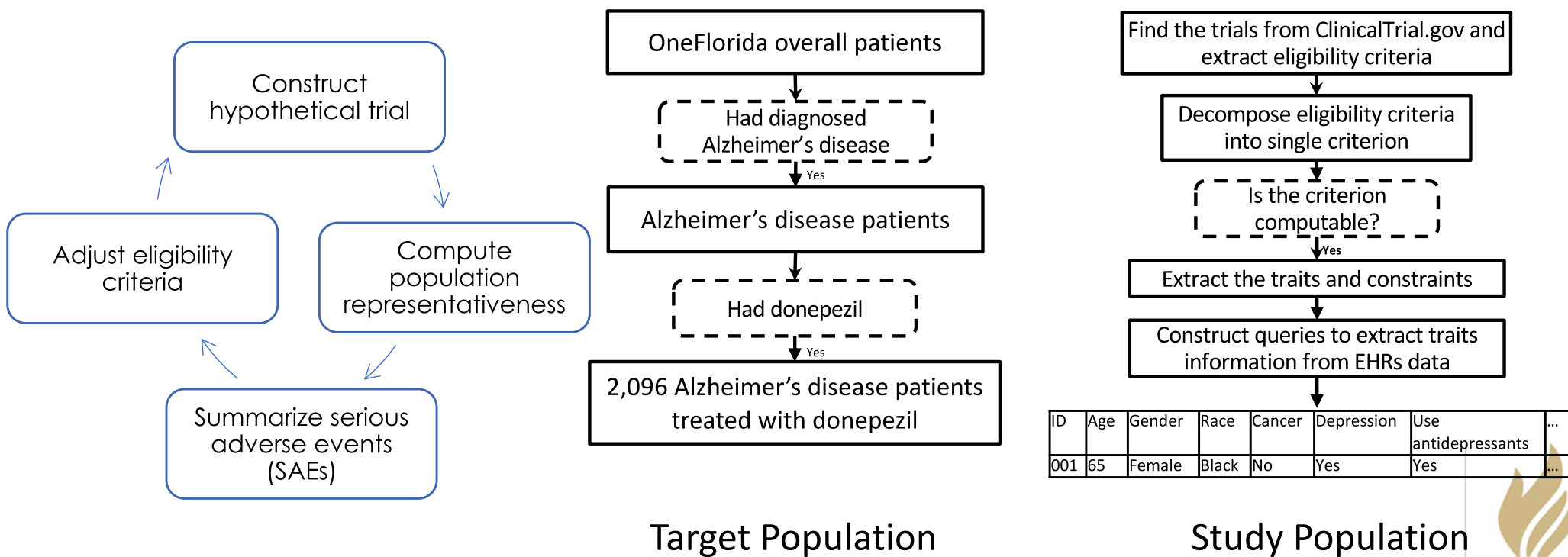


Fig. 1. The workflow for multivariate analysis of population representativeness of related clinical studies.

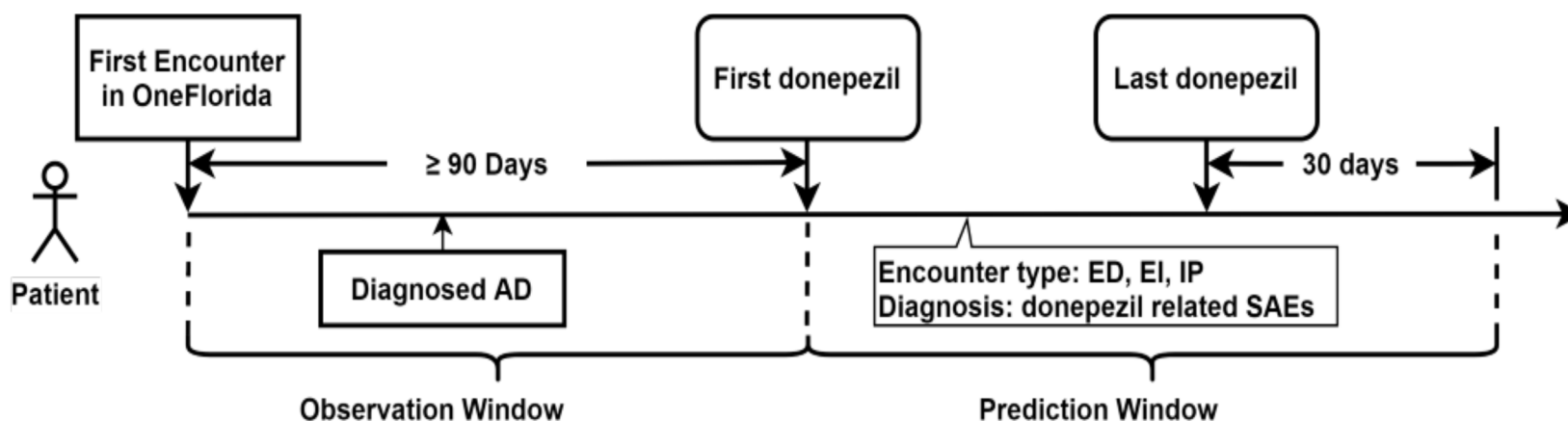


# Optimizing Eligibility Criteria Design with Informatics

- How broadening trial eligibility criteria will simultaneously impact trial generalizability and clinical outcomes in real-world data?



## Define Serious Adverse Events (SAEs)



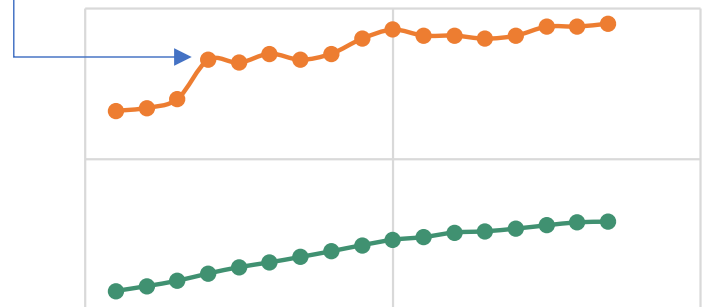
- Summarize SAEs from drug labels and completed trials
- Convert to ICD-9/10 codes
- Calculate number of SAEs for each patient



# Optimizing Criteria Design

Subsequently dropping criterion	Study population		mGIST
	Population Size (N)	# of SAEs	
00.Original	373	0.66	0.062
01.Drop Cardiac disease	503	0.67	0.078
02.Drop Depression	603	0.70	0.096
03.Drop Uncontrolled Hypertension	744	0.83	0.120
04.Drop Dementias other than AD	865	0.82	0.141
05.Drop Sleep disorder	972	0.85	0.157
06.Drop Cancer	1101	0.83	0.175
07.Drop Visual/Hearing Impairment	1206	0.85	0.194
08.Drop fecal/urinary incontinence	1320	0.90	0.214
09.Drop Conditions affect absorption	1413	0.93	0.232
10.Drop Parkinson's disease	1499	0.91	0.241
11.Drop Other AD treatments	1575	0.91	0.256
12.Drop Schizophrenia	1645	0.90	0.260
13.Drop Drug use disorders	1703	0.91	0.269
14.Drop Alcohol use disorders	1778	0.94	0.280
15.Drop Uncontrolled diabetes	1846	0.94	0.290
16.Drop Use antidepressants	1856	0.95	0.292

Determine whether a binary criterion should be included or removed



—●— # of SAEs —●— mGIST





# Outline

- Promises of Data Science in Healthcare
- Challenges of Data Science in Healthcare and Possible Pathways Forward
- What Makes Me an Ideal Candidate for this Position?



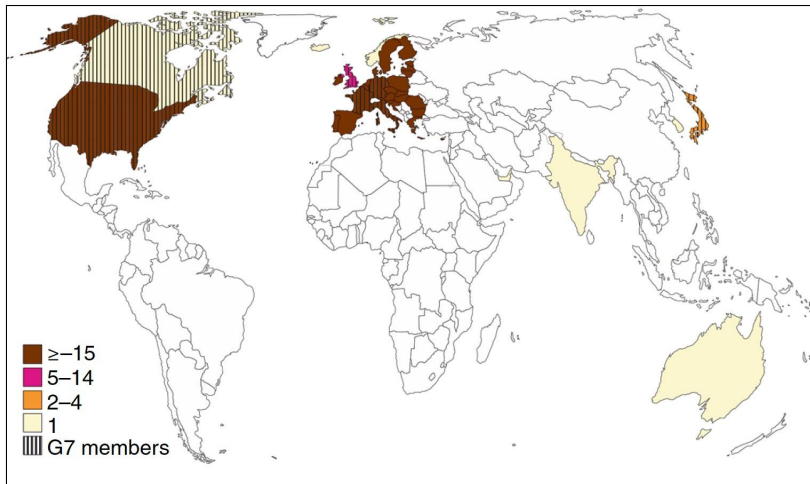


# Challenges in Health Data Science

- Issues with RWD like EHR
  - Biases
  - Coding of information
  - Missing information
    - 80% of clinical information is locked in free-text narratives
    - A lot of data types not accessible for research: imaging, genomics, microbiome, etc.
- Issues with models like deep learning
  - Explainability
  - ...



# Ethical AI Guidelines



**Fig. 2 | Geographic distribution of issuers of ethical AI guidelines by number of documents released.** Most ethics guidelines are released in the United States ( $n = 21$ ) and within the European Union (19), followed by the United Kingdom (13) and Japan (4). Canada, Iceland, Norway, the United Arab Emirates, India, Singapore, South Korea and Australia are represented with 1 document each. Having endorsed a distinct G7 statement, member states of the G7 countries are highlighted separately. Map created using [https://d-maps.com/carte.php?num\\_car=13181](https://d-maps.com/carte.php?num_car=13181).

- Most prevalent principles
  - **Transparency**
  - **Justice, Fairness, and Equity**
  - Non-maleficence
  - **Responsibility and Accountability**
  - Privacy
  - Beneficence
  - Freedom and autonomy
  - **Trust**
  - Sustainability
  - Dignity
  - Solidarity
- **WHO report on AI in health:**
  - Protecting human autonomy
  - Promoting human well-being and safety
  - Ensuring **transparency, explainability, and intelligibility**
  - Fostering responsibility and accountability
  - Ensuring **inclusiveness and equity**

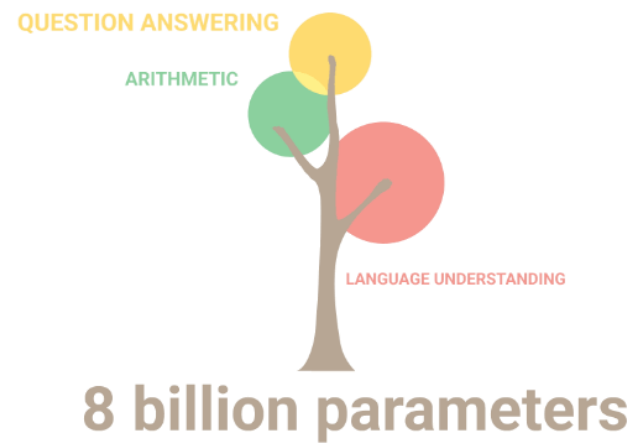


# Medical AI Reporting Guidelines

- MINIMAR (Hernandez-Boussard 2020)
  - **Study population and setting** (population, study setting, data source, cohort selection)
  - **Demographic characteristics** (age, sex, race, ethnicity. Socioeconomic status)
  - **Model architecture** (model output. Target user, data splitting, gold standard, model task, features, missingness)
  - **Model evaluation** (internal model validation, external model validation, transparency)
- TRIPOD-AI
  - **Methods** (Source of data, Participants, Data Preparation, Outcomes, Predictors, Sample size, Analytical methods. Model development)
  - **Results** (Participants, Model specification, Model performance, Usability of the model, Sensitivity analysis )
  - **Discussion** (Limitations, Interpretation, Implications)



# Generative AI for Healthcare



(Google Blog)



# Generative AI for Healthcare

nature

Explore content ▾ About the journal ▾ Publish with us ▾

nature > articles > article

Article | [Open Access](#) | [Published: 12 July 2023](#)

## Large language models encode clinical knowledge

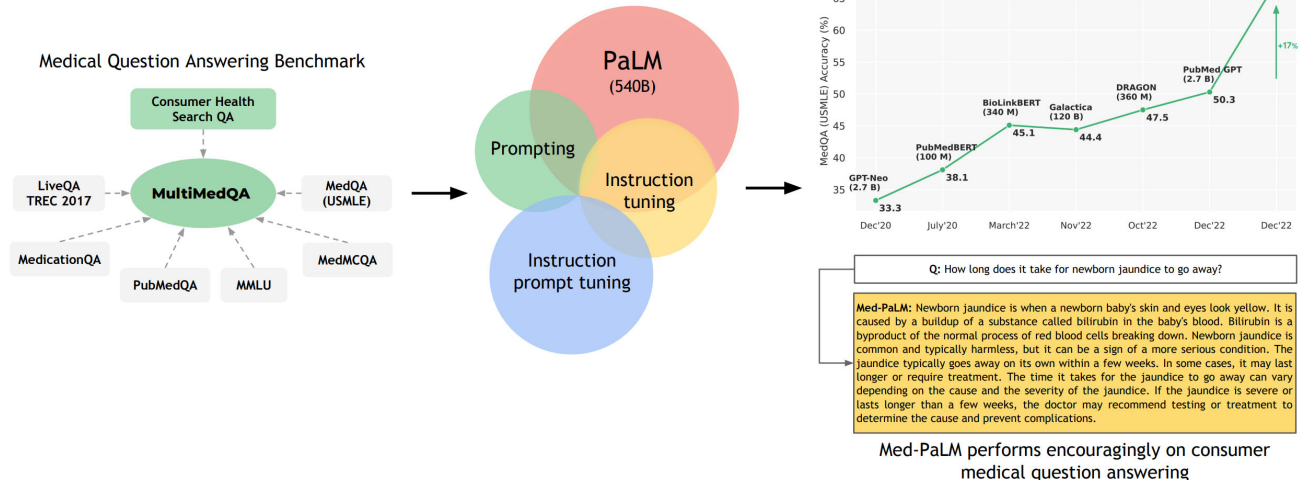
[Karan Singhal](#) ✉, [Shekoofeh Azizi](#) ✉, [Tao Tu](#), [S. Sara Mahdavi](#), [Jason Wei](#), [Hyung Won C Scales](#), [Ajay Tanwani](#), [Heather Cole-Lewis](#), [Stephen Pfohl](#), [Perry Payne](#), [Martin Senevirat Gamble](#), [Chris Kelly](#), [Abubakr Babiker](#), [Nathanael Schärli](#), [Aakanksha Chowdhery](#), [Philip M Demner-Fushman](#), [Blaise Agüera y Arcas](#), [Dale Webster](#), [Greg S. Corrado](#), [Yossi Matias](#), I ... [Vivek Natarajan](#) ✉ [+ Show authors](#)

[Nature](#) (2023) | [Cite this article](#)

44k Accesses | 707 Altmetric | [Metrics](#)

(Singhal Nature 2023)

- Med-PaLM (PaLM tuned for medical domain)
  - Obtained a passing score (60%) on US medical licensing-style questions
- Med-PaLM 2
  - Achieved “expert” doctor level on medical exam questions (85%)
- AI for cancer radiotherapy planning with Mayo Clinic



# How to Check Compliance with AI Reporting Guidelines?

**User:** For validation, describe how the predictions were calculated

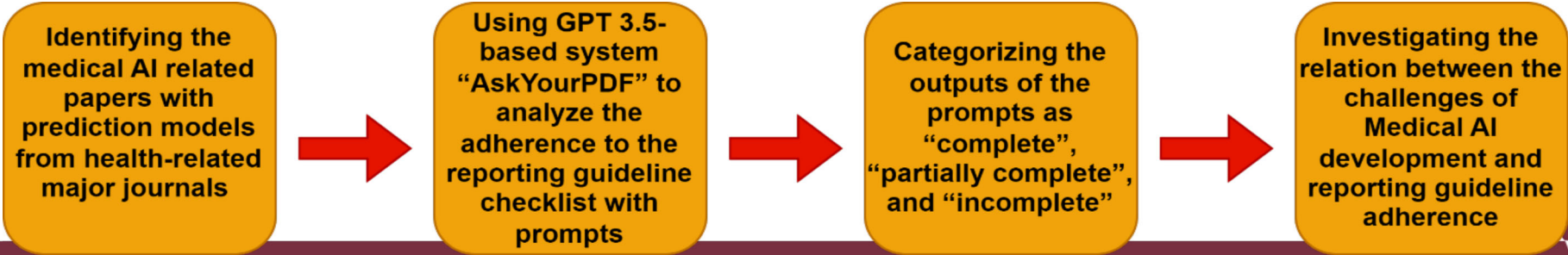
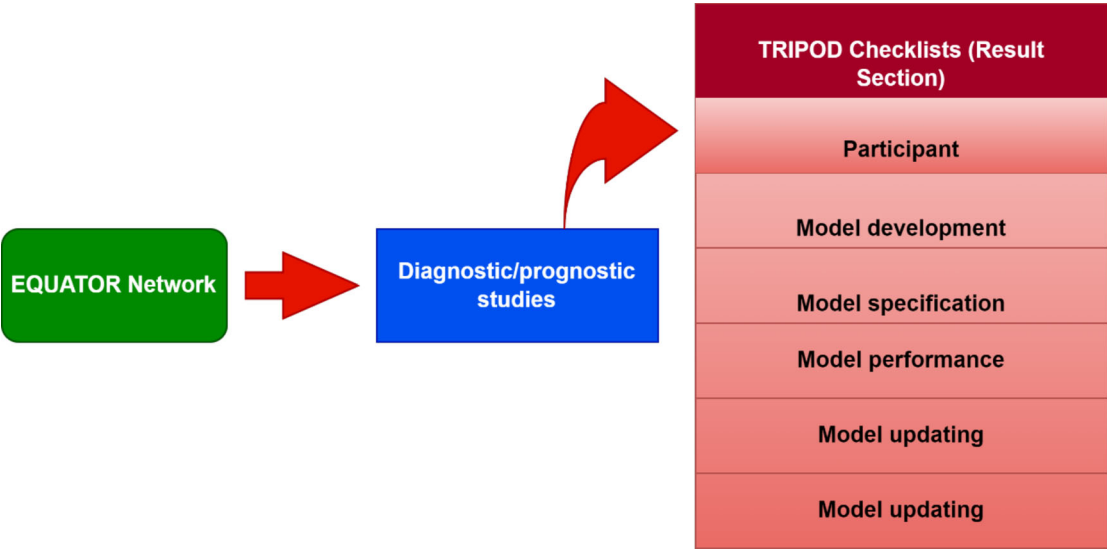
**ChatBOT:** The study used a variety of models, including a Naive Bayesian Classifier (NBC) and Balanced Random Forest Classifiers (BRFs), to predict suicide risk based on electronic health record (EHR) data. The models were trained on a subset of the data and validated on a separate subset using metrics such as area under the receiver operating curve (AUC) and positive predictive value (PPV). The specific details of how the predictions were calculated for validation are not provided in the given context.

**User:** Provide details on how risk groups were created, if done

**ChatBOT:** The study did not create risk groups. Instead, the study used logistic regression and random forest classifiers to predict suicide risk based on temporal variables and other features in electronic health records. The models were trained on a dataset of individuals with and without suicide attempts, and the performance of the models was evaluated using metrics such as area under the receiver operating curve (AUC) and positive predictive value (PPV).



# Evaluating AI Papers Against TRIPOD





# Identifying AI Papers to Evaluate

- Identified 28 medical AI related papers from literature search using keywords such as “prediction model”, “electronic health records”, “EHR”, and “machine learning” in health-related major journals (e.g., JAMIA and JAMIA Open) from PubMed.
- Among these 28 papers, we closely investigated 10 papers with one-shot and few-shot prompting techniques to analyze their adherence to the reporting guideline checklist.
- We uploaded one PDF-version paper at a time, and used the one or zero-shot prompts for each reporting guideline item into a GPT 3.5-based system known as “AskYourPDF” as input. Finally, we categorized the outputs of the prompts as “complete”, “partially complete”, and “incomplete”, where appropriate.



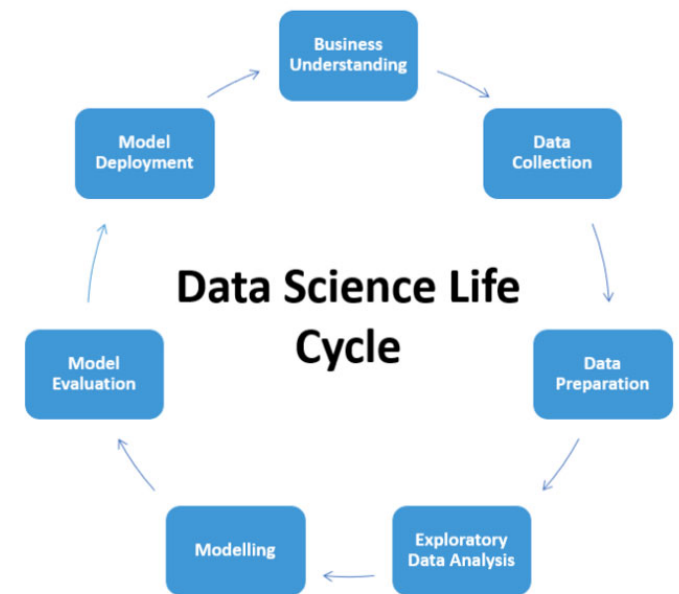
# Evaluation Results

Challenges	TRIPOD Checklists	Complete	Partially Complete	Incomplete
Fairness	Participant's Demographics	9		1
Accuracy	Model Performance	8		2
Explainability	Interpretation	10		
	Model Specification		1	9
Reproducibility	Model's Internal Validation	5		5
	Model's External Validation	2		8



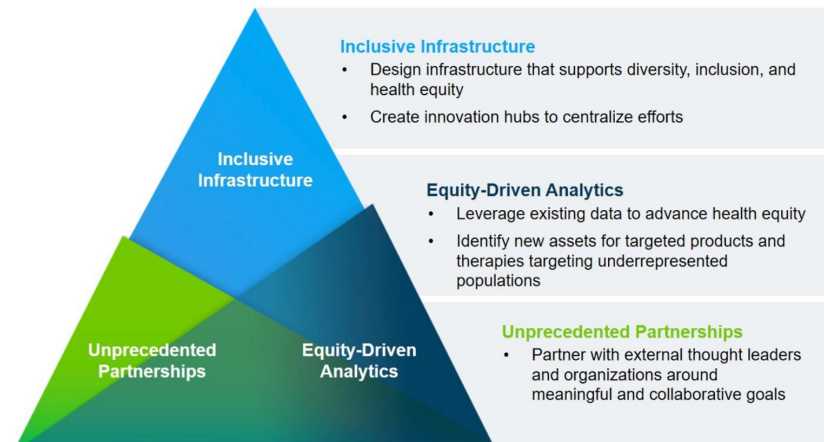
# What are missing in the Guidelines?

- The extent to which domain experts (e.g., MDs) are involved in the life cycle of data science
  - Data modeling
  - Feature engineering
  - Model selection
  - Result interpretation



# Achieving Equity, Diversity, and Inclusion in Health Data Science

- Inclusive data infrastructure
  - Bridge2AI: promoting EDI and FAIR
  - AIM-AHEAD
- Inclusive design
  - Accurate representation of knowledge
  - Explainable AI / Causality
  - Federated learning
- Social determinants of health
- Partnership

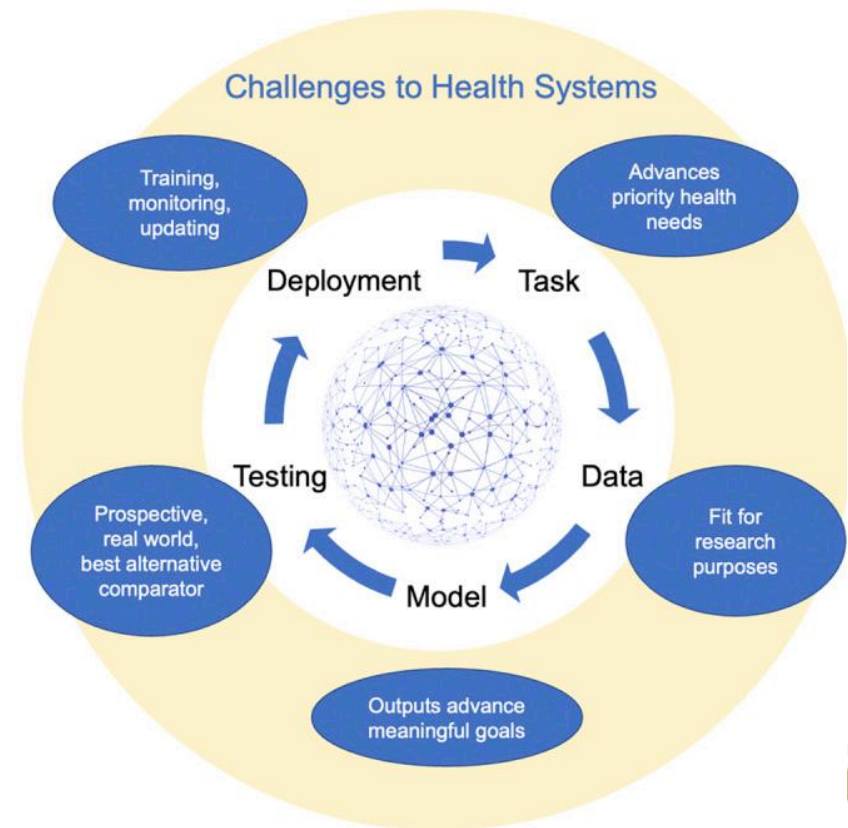


(Avalere.com)



# Strategies to Engage Clinicians and Basic Scientists in Data Science

- Train clinicians and basic scientists to be AI literate
- Using seed grants to incentivize collaboration
- Start from priority health needs
- Employ user-centered and human-in-the-loop approach
- Improve model explainability and trustworthiness
- Show real-world impact of data science

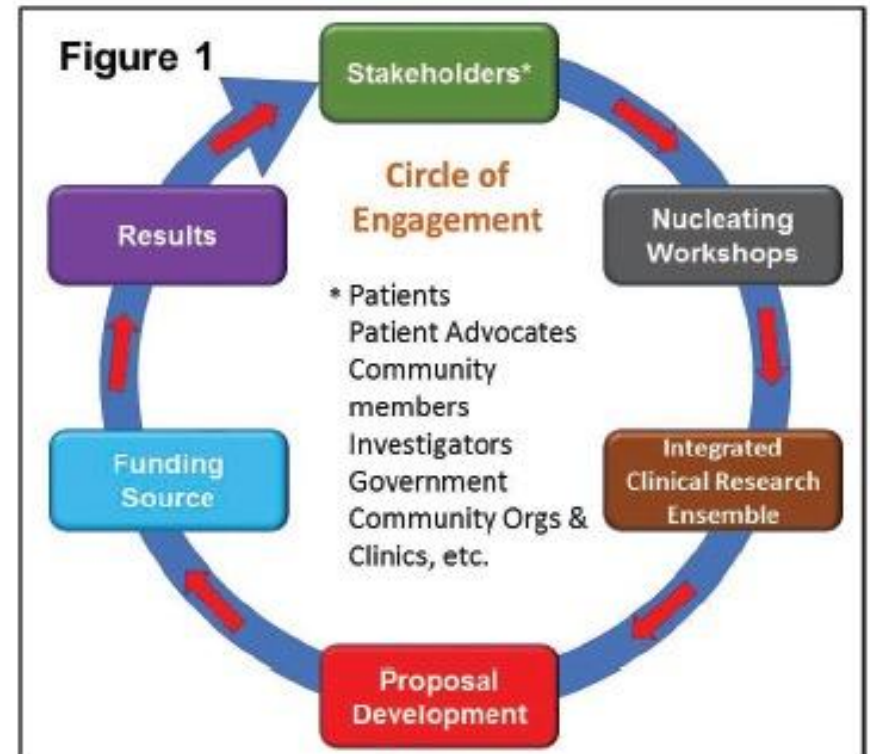


London AJ. Artificial intelligence in medicine: Overcoming or recapitulating structural challenges to improving patient care? Cell Rep Med. 2022 May 17;3(5):100622.



# Community Engagement

- Public education
- Citizen scientists
- CTSI Discovery Radio and Podcast programs
- Benefit the community with research findings and tools
- Prioritize disadvantaged regions and underserved populations



CTSI Circle of Engagement



# Future Directions

- Fundamental Research
  - Enhancing the interpretability of AI models: causality, transparency, trustworthiness
  - Standardizing SDOH and patient-reported measures (e.g., pain experience, function and health status, mental health issues) in EHR
  - Evaluating how the research community follows the guidelines for reporting AI models (TRIPOD-AI, MINIMAR, etc.)
- Applied AI Research
  - XAI-based risk prediction and decision support
  - Clinical trial patient selection and optimization with informatics





# FLORIDA STATE UNIVERSITY NEWS

THE OFFICIAL NEWS SOURCE OF FLORIDA STATE UNIVERSITY

HOME NEWS TOPICS ▾ MULTIMEDIA ▾ ABOUT FSU ▾ FOR JOURNALISTS ▾

HOME / NEWS / SCIENCE & TECHNOLOGY / FSU CONTINUES TO ADVANCE TRANSLATIONAL RESEARCH WITH NEW GRANT FROM THE NATIONAL LIBRARY OF MEDICINE

## FSU continues to advance translational research with new grant from the National Library of Medicine

BY: KELSEY KLOPFENSTEIN ✉ | PUBLISHED: MAY 11, 2022 | 3:56 PM | SHARE: [f](#) [t](#)

### RELATED STORIES

'Top Gun' research: FSU's FCAAP innovates at high speed for Florida's aerospace industry

FSU researchers expand understanding of vortex spread in superfluids

Shine a light: New research shows how low-energy light can bend plastic

FSU chemistry doctoral student wins ACS award for research on recyclability

FAMU-FSU Engineering students design four-legged lunar robot for NASA's BIG Idea Challenge

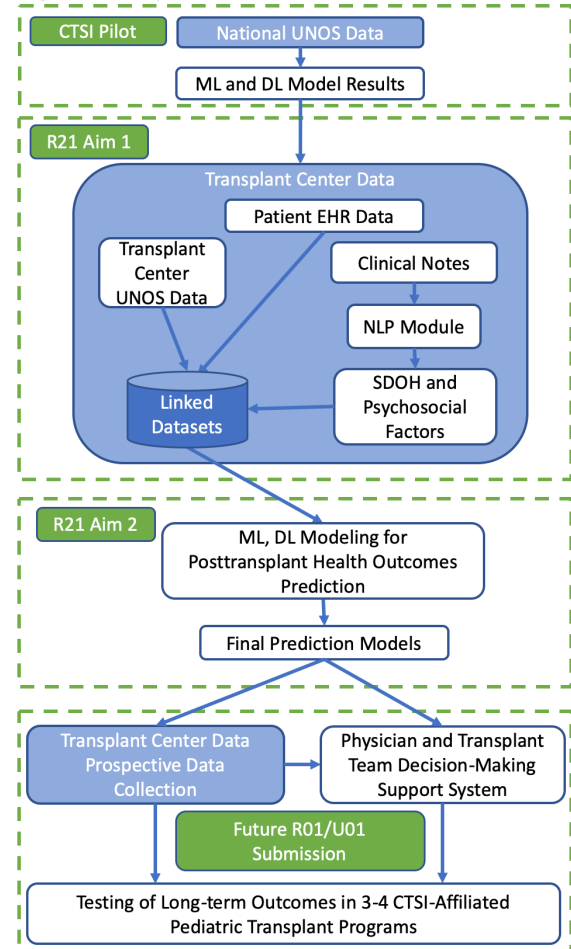
### RECOMMENDED STORIES

'Top Gun' research: FSU's FCAAP innovates at high speed for Florida's aerospace industry

FSU researchers expand understanding of vortex spread in superfluids



Figure 1: Analytic Plan





# Interdisciplinary Team Led by Zhe He Receives ISL Planning Grant

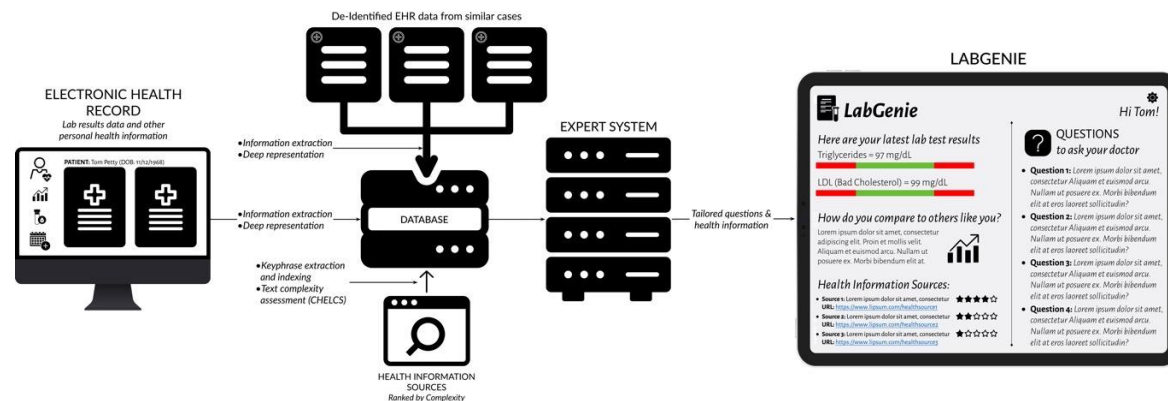
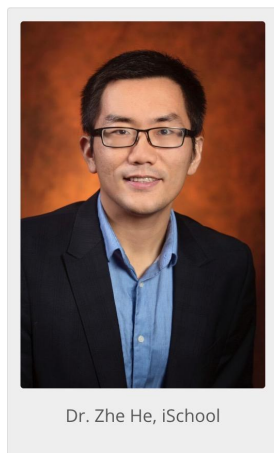
Posted by Michael Walsh on May 4, 2022

iSchool Professor Dr. Zhe He and his interdisciplinary team recently received a grant from the Institute for Successful Longevity (ISL) for their project, “Towards an AI-Assisted Application for Lab Result Comprehension for Older Adults with Multiple Chronic Conditions,” which aims to make health information more accessible to patients, especially older adults, by developing a portal catered to patients’ information needs. The study team will use this planning grant to collect pilot data for the developing of an AI-assisted tool to help older adults interpret their lab results by providing them with tailored recommendations and health information sources.

“Our tool will enhance older adults’ understanding of medical information and facilitate patient-provider communication and shared decision making,” said He. “With older adults more engaged in the care, they are more likely to follow the treatment plan, adhere to the medications, and live a healthy lifestyle, which will support their successful longevity.”

The ISL conducts research into how to live longer, stay active, and be fully engaged in life. An ISL Planning Grant offers \$25,000 to support the collection of pilot data; it is seen as a first step toward gaining federal funding.

“We hope to use this planning grant to generate strong pilot data for our federal grant application,” said He. “In the long term, we hope it will have a sustained and positive impact on patient engagement, self-management, and ultimately health outcomes.”



## Zhe He Receives the 2022 Lois Lunin Award



The Association for Information Science & Technology (ASIS&T) is delighted to announce that Zhe He of Florida State University, Tallahassee, FL is the 2022 recipient of the Lois Lunin Award. This award recognizes individuals who have made noteworthy contributions to the practice of Information Science and Technology through leadership, mentoring, and innovation.

Dr. Zhe He is an Associate Professor at Florida State University School of Information. He also holds courtesy appointments with the Department of Behavioral Sciences and Social Medicine of the College of Medicine and Department of Computer Science. He is Informatics Lead of University of Florida-Florida State University Clinical and Translational Science Award. His research focuses on the practice of information science and technology in health and biomedicine. He has been funded by the National Institutes of Health, Eli Lilly and Company, Amazon, NVIDIA, FSU Council on Research and Creativity, and Institute for Successful Longevity. Dr. He obtained Ph.D. in Computer Science from New Jersey Institute of Technology, an MS in Computer Science from Columbia University, and a BS in Computer

Science from Beijing University of Posts and Telecommunications. Before joining FSU, he was a Postdoctoral Research Scientist in the Department of Biomedical Informatics at Columbia University. He has published over 100 peer-reviewed papers in leading health informatics and information science venues. He is an Associate Editor of BMC Medical Informatics and Decision Making. In 2022, he was inducted as a Fellow of the American Medical Informatics Association.

In nominating He for the award, Jinxuan Ma from the School of Library and Information Management at Emporia State University stated, "Dr. He has strived to make an impactful contribution to medical/biomedical fields with his research expertise in biomedical ontologies, machine learning, natural language processing, knowledge representation, and big data analytics. With the aim of improving the population health and advance biomedical research through the collection, analysis, and application of health and clinical data, such passion to improve the lives of people has majority impacted on Dr. He's interdisciplinary research, innovative teaching, outstanding leadership, and influential mentorship."

Upon learning of his selection as the 2022 Lois Lunin Award winner, He said, "I'm thrilled to receive Lois Lunin Award from ASIS&T this year. This incredible honor recognizes my commitment and dedication to promoting the practice of information science and technology in the biomedical and health domain through leadership, mentoring, and innovation. In the era of COVID-19, like many educators and researchers in our field, I have a paramount responsibility to educate students on important issues around health information and data science and to innovate information technologies that positively impact our society. After learning about Dr. Lois Lunin's incredible professional life designing and implementing medical information and imaging systems, I have an even stronger aspiration for continuing Dr. Lunin's legacy and my contributions to the education and research in information science and technology for the betterment of our world."

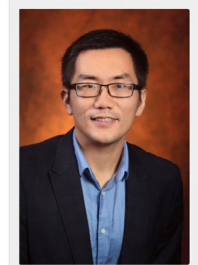


## CCI Faculty Member Dr. Zhe He Recognized as Fellow of AMIA

Posted by Emma Bonney on October 8, 2021

iSchool Associate Professor and Researcher **Dr. Zhe He** was recognized as a Fellow of the American Medical Informatics Association (FAMIA) on October 5th, 2021.

The FAMIA Applied Recognition program was established in 2018 and recognizes professionals who have made significant contributions to the field of biomedical informatics and demonstrate an interdisciplinary understanding of the field. Dr. He has worked on a number of projects, both internally and externally, dedicated to advancing health science data and biomedical informatics. "My past and on-going applied informatics research on clinical trial design, biomedical ontology quality assurance, and consumer health information seeking well prepared me for this," said Dr. He. "Moreover, as Informatics Lead of UF-FSU Clinical and Translational Science Award, I have been leading a team to provide consulting services on informatics, biostatistics, epidemiology, and research design to the FSU community. This allowed me to gain valuable leadership experience and make strong connections with other researchers."



Dr. Zhe He

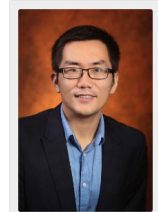
## iSchool Professor Elected as AMIA KDDM Vice Chair

Posted by Lauren Scala on November 30, 2022

Associate Professor in the School of Information Dr. Zhe He was recently elected as the Vice Chair for the American Medical Informatics Association's Knowledge Discovery and Data Mining (AMIA KDDM) Working Group.

AMIA connects a broad community of professionals and students interested in informatics and is the bridge for knowledge and collaboration across a continuum, from basic and applied research to the consumer and public health arenas. AMIA Working Group serves as networks in which AMIA members can exchange information on a particular area of special interest in biomedical and health informatics.

KDDM focuses on the process of extracting meaningful patterns from biomedical data (knowledge discovery), and using automated computational and statistical tools and techniques on large datasets (data mining). As such, KDDM Working Group has been a nexus of researchers and practitioners who develop and use AI and data science techniques in the biomedical and clinical domain to share their experiences, best practices, and perspectives to move the field forward. Dr. He has been a member of AMIA since 2009, and served as the Secretary of the KDDM Working Group from 2020-2022. Earlier this year, Dr. He was inducted as a Fellow of AMIA.



Dr. Zhe He



# Acknowledgements

## Students:



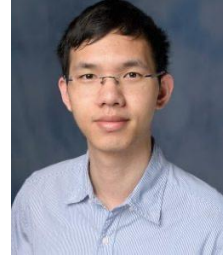
Neelufar Payrovnaziri



Laura Barrett



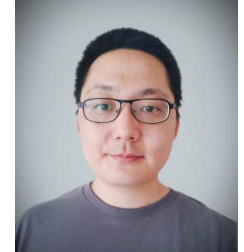
Shubo Tian



Qian Li



Hansi Zhang



Arslan Erdengasileng

## Collaborators:



Xiuwen Liu, PhD



Jiang Bian, PhD



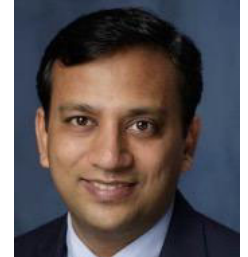
Pablo Rengifo, MD



Jonathan Chen, MD



Michael Killian, PhD



Dipankar Gupta, MD

## Funding:

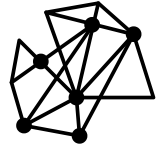
R21LM013911  
R21AG061431  
R21AG061431-02S1  
R01AG064529  
P01AA029547



# References

- Payrovnaziri SN, Chen Z, Rengifo-Moreno P, Miller T, Bian J, Chen JH, Liu X, He Z. Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review. *Journal of the American Medical Informatics Association*. 2020 Jul;27(7):1173-85.
- Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. *Journal of Biomedical Informatics*. 2021 Jan 1;113:103655.
- Miller RJ, Sabovčik F, Cauwenberghs N, Vens C, Khush KK, Heidenreich PA, Haddad F, Kuznetsova T. Temporal shift and predictive performance of machine learning for heart transplant outcomes. *The Journal of Heart and Lung Transplantation*. 2022 Jul 1;41(7):928-36.
- Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*. 2019 Sep;1(9):389-99.





**eHealth Lab**  
ehealthlab.cci.fsu.edu

**Thank you!**

**Zhe He, PhD, FAMIA**  
zhe@fsu.edu

